Lunchtime Data Talk IV

Home Price Indices: Appreciating the Differences

Mark Fleming, CoreLogic
Stan Humphries, Zillow

September 16, 2013

URBAN INSTITUTE

Measuring House Prices

Valuing a Thinly Traded Heterogeneous Good

September 16th, 2013

Mark Fleming, Chief Economist

CoreLogic's "HPI Ecosystem"

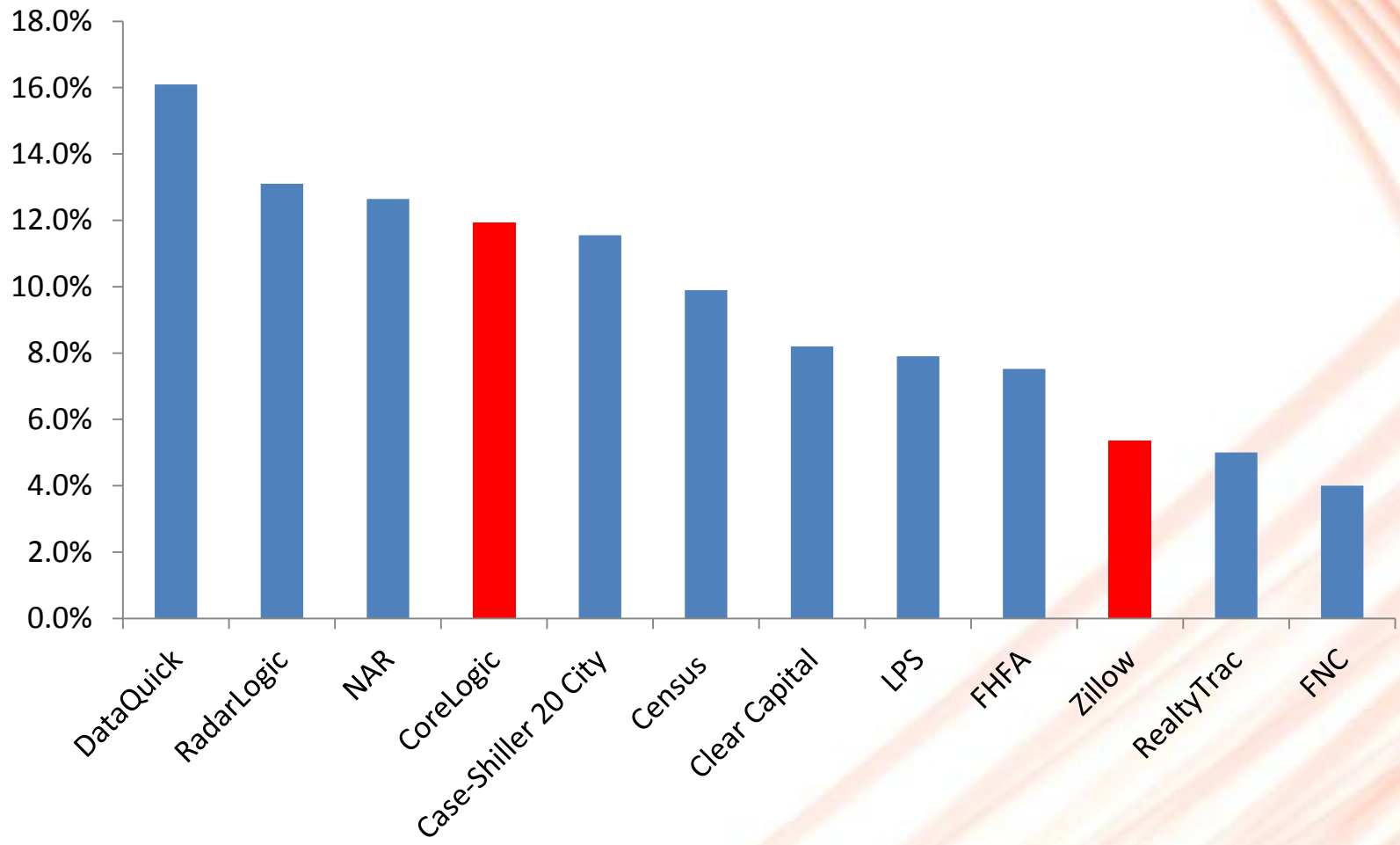
	CoreLogic HPI	CoreLogic Case-Shiller Indexes	Moody's Analytics	S&P
Frequency	Monthly	Quarterly	Monthly forecast of all CSI's, + Quarterly history	Monthly
Data Source	Public Record, CL Securities, CL Servicing	Public Record	Public Record	Public Record, plus one MLS (Dallas SP only)
Methodology	Repeat Sales, arithmetic Mean	Repeat sales, arithmetic mean. (note FHFA interval weighted, geometric mean seasonally adjusted)	Repeat sales, arithmetic mean. (note FHFA interval weighted, geometric mean seasonally adjusted)	Repeat sales, arithmetic mean
Revisions	Monthly revisions no stop date	Revisions to 2 years only	Revisions to 2 years only	
History	35+ Years	30+ Years	30+ Years	7/20 Years
Transaction Type	All sales regardless of financing	All sales regardless of financing	All sales regardless of financing	All sales regardless of financing
Geographic Coverage	►Zips: 6700+ ►CBSA: 519+ ►County:1100+ ►States 50+, DC ►National	*Includes FHFA Data blended into the index. Uses FHFA data where CS has no coverage ►(CSI's / FHFA's) – Total ►Zips: (5,337 / 0) – 5,337 ►MSA: (105 / 290) – 395 ►County:(344 / 83) – 427 ►States: (28 / 23) – 51 ►Composite National ►(96 condo st, MSA, cnty)	*Includes FHFA Data blended into the index. Uses FHFA data where CS has no coverage ►(CSI's / FHFA's) – Total ►Zips: (5,337 / 0) – 5,337 ►MSA: (105 / 290) – 395 ►County:(344 / 83) – 427 ►States: (28 / 23) – 51 ►Composite National ►Fcst all CSI's (FHFA + CSI)	20 MSA, 10/20/Nat'l Composite ►20/MSA Composite ►20 City MSA's, 6 City Condos ►National Index
Tiers	►12 Tiers (including distressed excluded tier), tier coverage varies by geography	►Tiers: Low, Mid, High, Aggregate ►State: 17 tiered ►MSA 56 tiered ►County 73 tiered ►Zip: 2075 tiered	►Tiers: Low, Mid, High, Aggregate ►State: 17 tiered ►MSA 56 tiered ►County 73 tiered ►Zip: 2075 tiered	►17 MSAs with tiers
Transaction counts	70+ Million	45–50 Million sales transactions (does not include FHFA)	45–50 Million sales transactions (does not include FHFA)	N/A
Non-Disclosure States/Counties	Yes	No	No	Dallas Metro

Source: CoreLogic August 2013

CoreLogic

Its Housing Data! Of Course There Are A Diversity of Right Answers

Year-Over-Year Growth As of May 2013-Indices Based on a Variety of Methods and Input Data

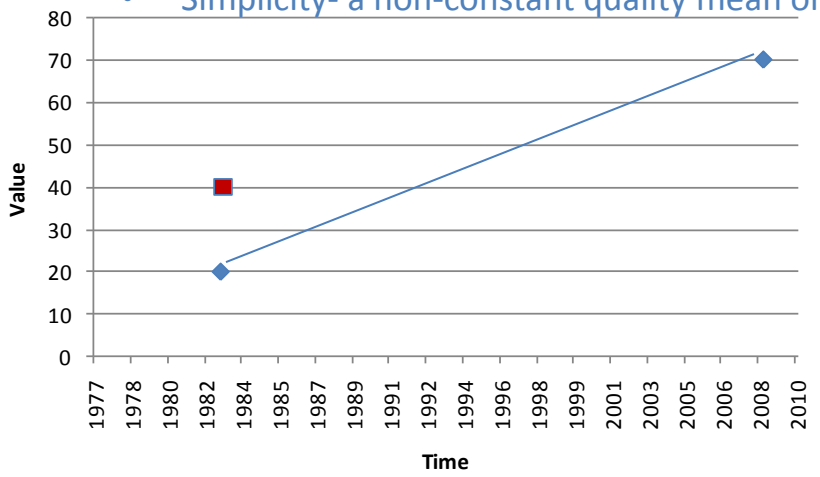


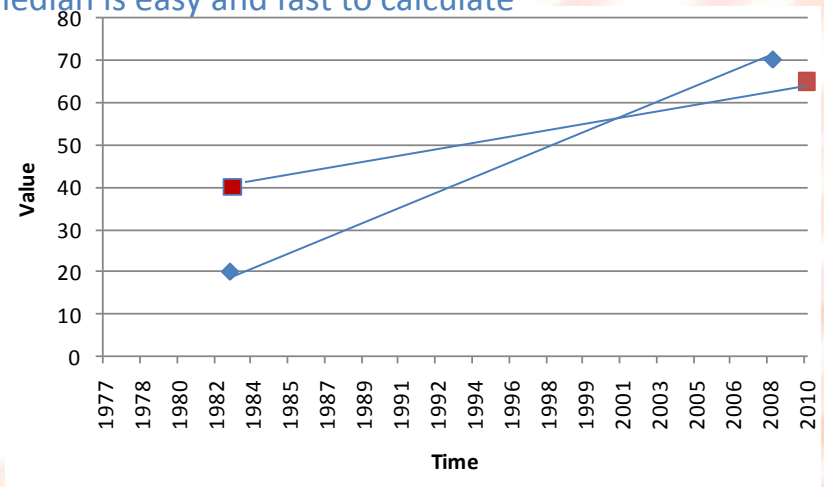
Source: CoreLogic Calculations August 2013

CoreLogic

Method Matters- Why Use Repeat Sales Techniques

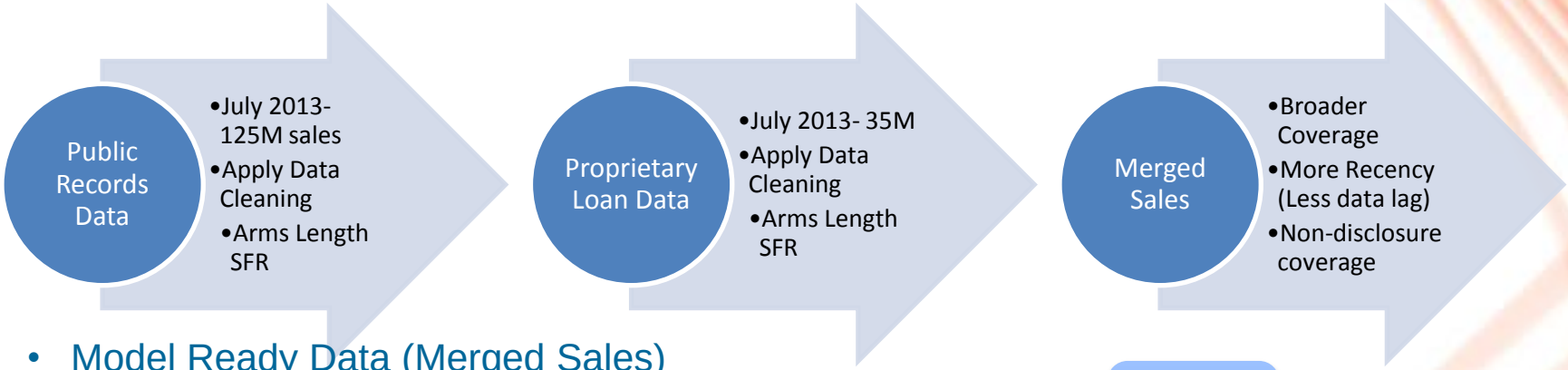
- Advantages
  - Constant Quality- Controlling for the changing mix of sales by measuring the change in price on the same housing unit
  - Less data required than hedonic methods-requires addresses, prices and dates
  - Housing characteristics not consistently available in all geographies
  - Long History of Repeat Sales Technique -beginning with art sales (another infrequently traded and heterogeneous good)
  - Uses flow of "new market information"- based in information theory and signal processing
    - Less than about .5% of the housing stock transacts every month
  - Based on actual prices as opposed to aggregations of price predictions
    - Price Predictions (AVMS)- >10% error ≈30%
- Disadvantages
  - New Homes (or any single sales) not included
  - Simplicity- a non-constant quality mean or a median is easy and fast to calculate



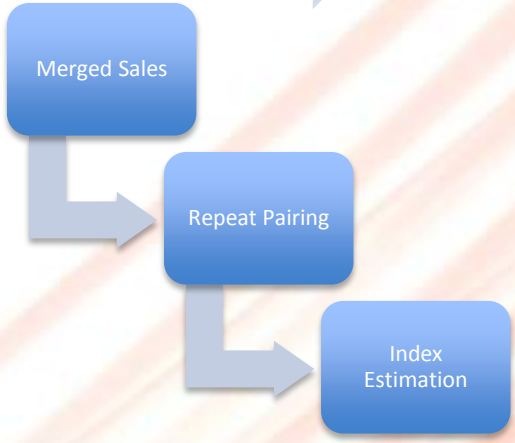


CoreLogic

Repeat Sales Model Architecture-Data, Data, Data



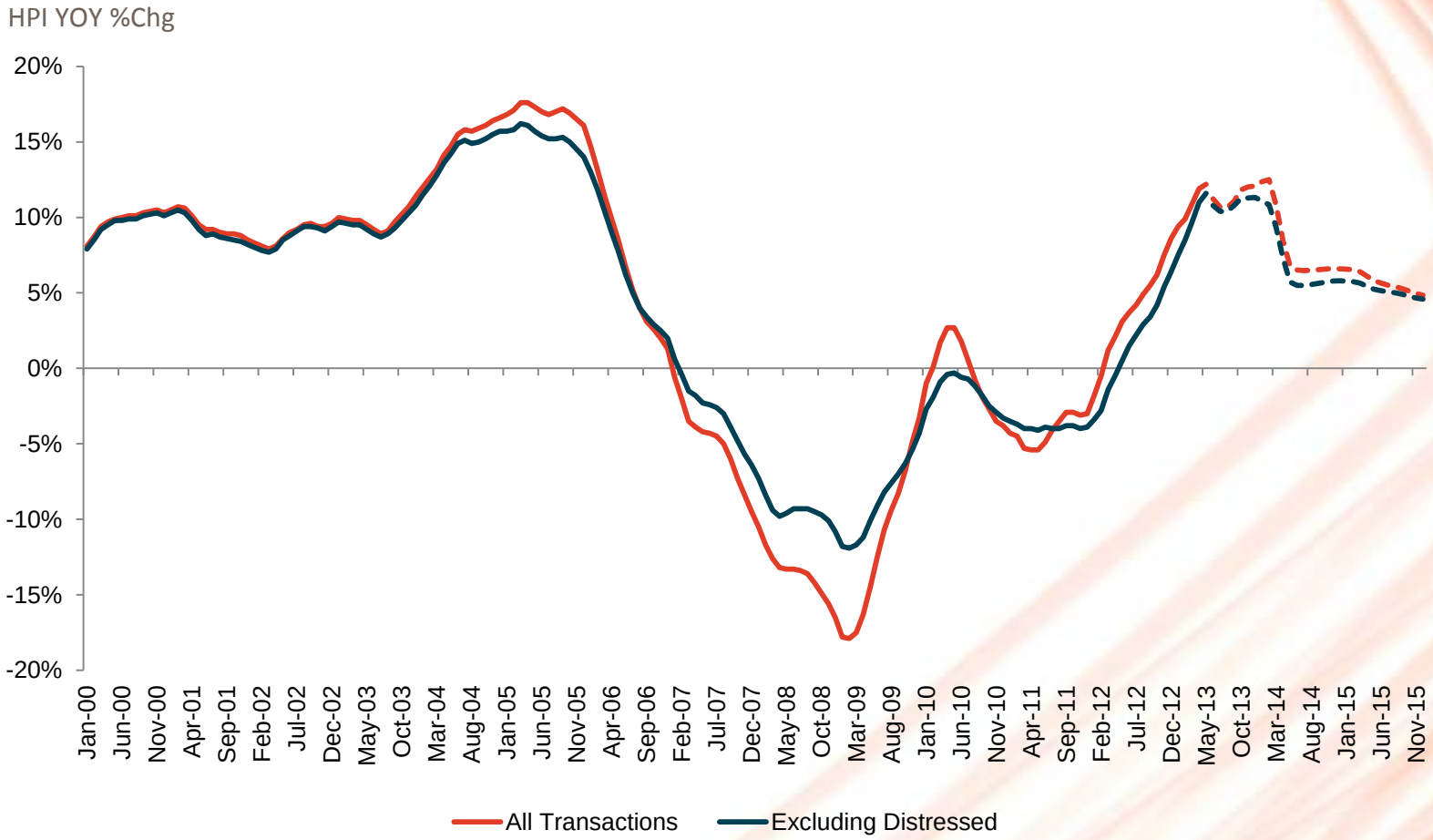
- Model Ready Data (Merged Sales)
  - Data cleaning critical to remove “noise” from the signal
  - Control for remodels (constant quality)
  - Targeting single family residential arms length transactions
    - Include Reo and short sales (handled in index construction)
- Repeat Pairing Process
  - Data cleaning critical to remove “noise” from the signal
- Index Estimation by Tier and Geography
  - Model estimation by tier (12) and all geographic levels
  - Index exclusion and Infilling
  - Final Index Output



CoreLogic

National HPI Dynamically Responds to Distribution of Transactions

National index is not constant transaction weighted- Based on transaction volumes in each market

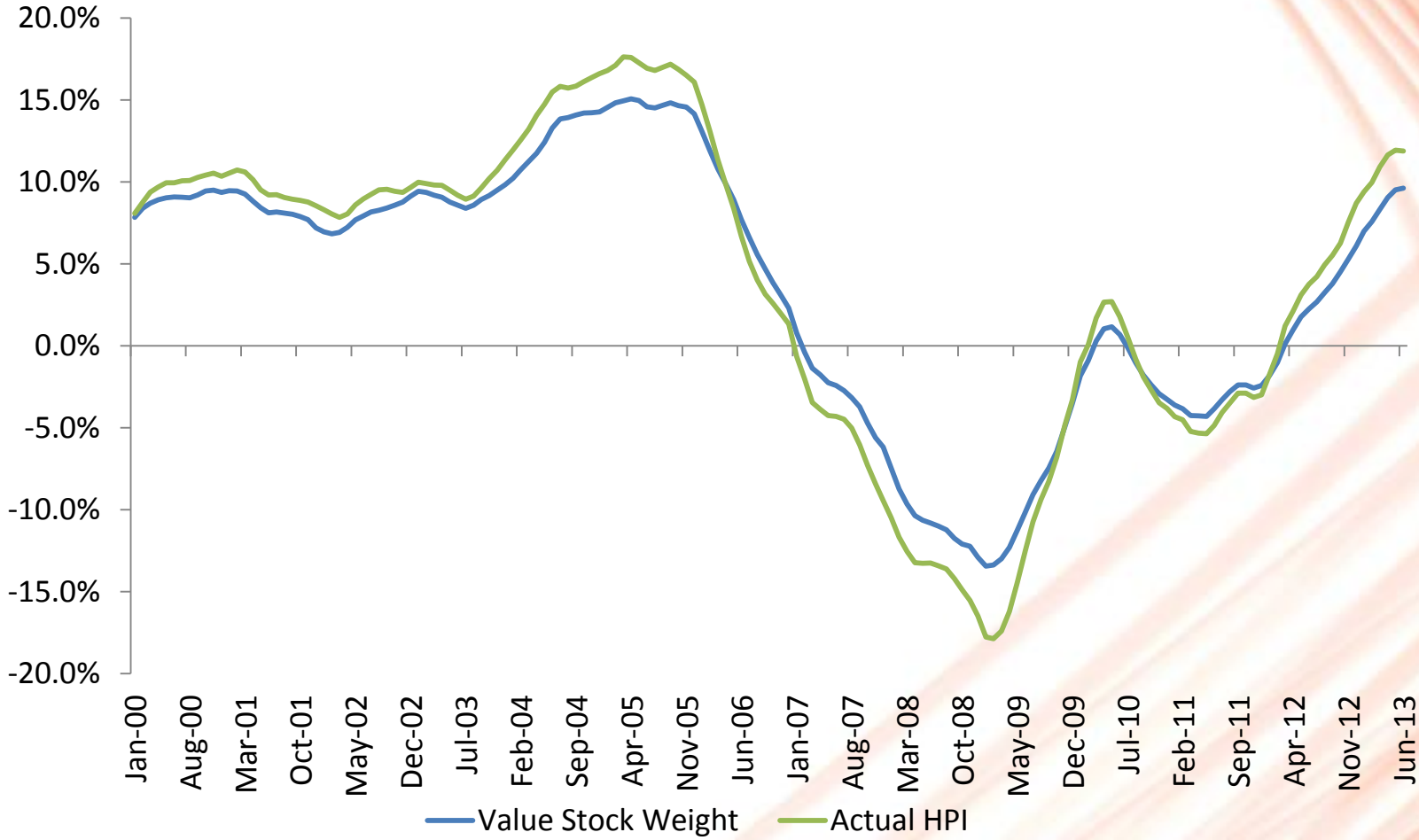


Source: CoreLogic HPI, HPI Forecast May 2013

CoreLogic

Different Weighting Doesn't Yield Large Quantitative Differences

National CoreLogic HPI Year-Over-Year Growth By Different Weight Types

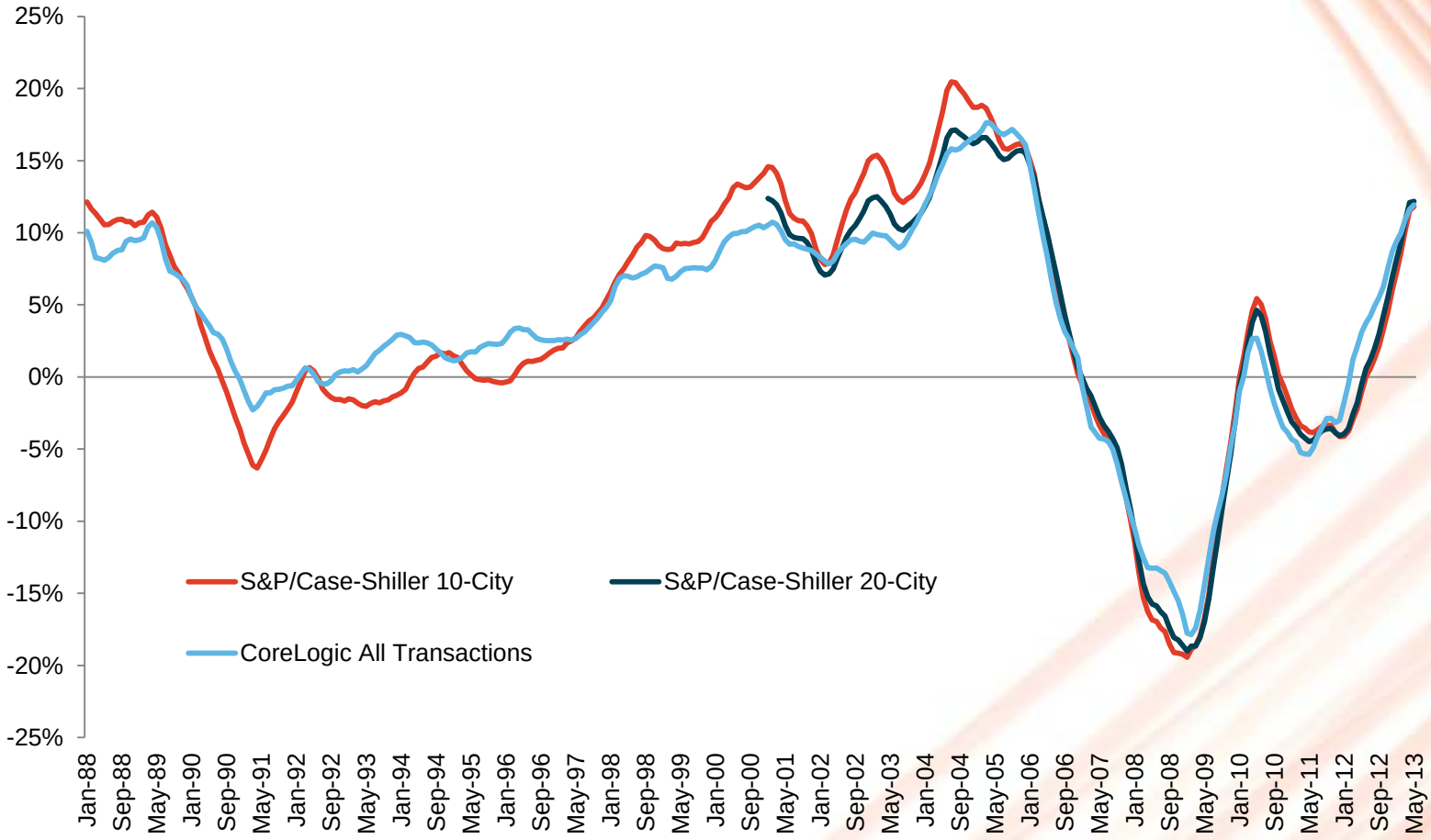


Source: CoreLogic July 2013

CoreLogic

Composite Repeat Sale and Dynamic Distribution Indices Are Similar

HPI Year-Over-Year Growth- Selected Repeat Sale Indices

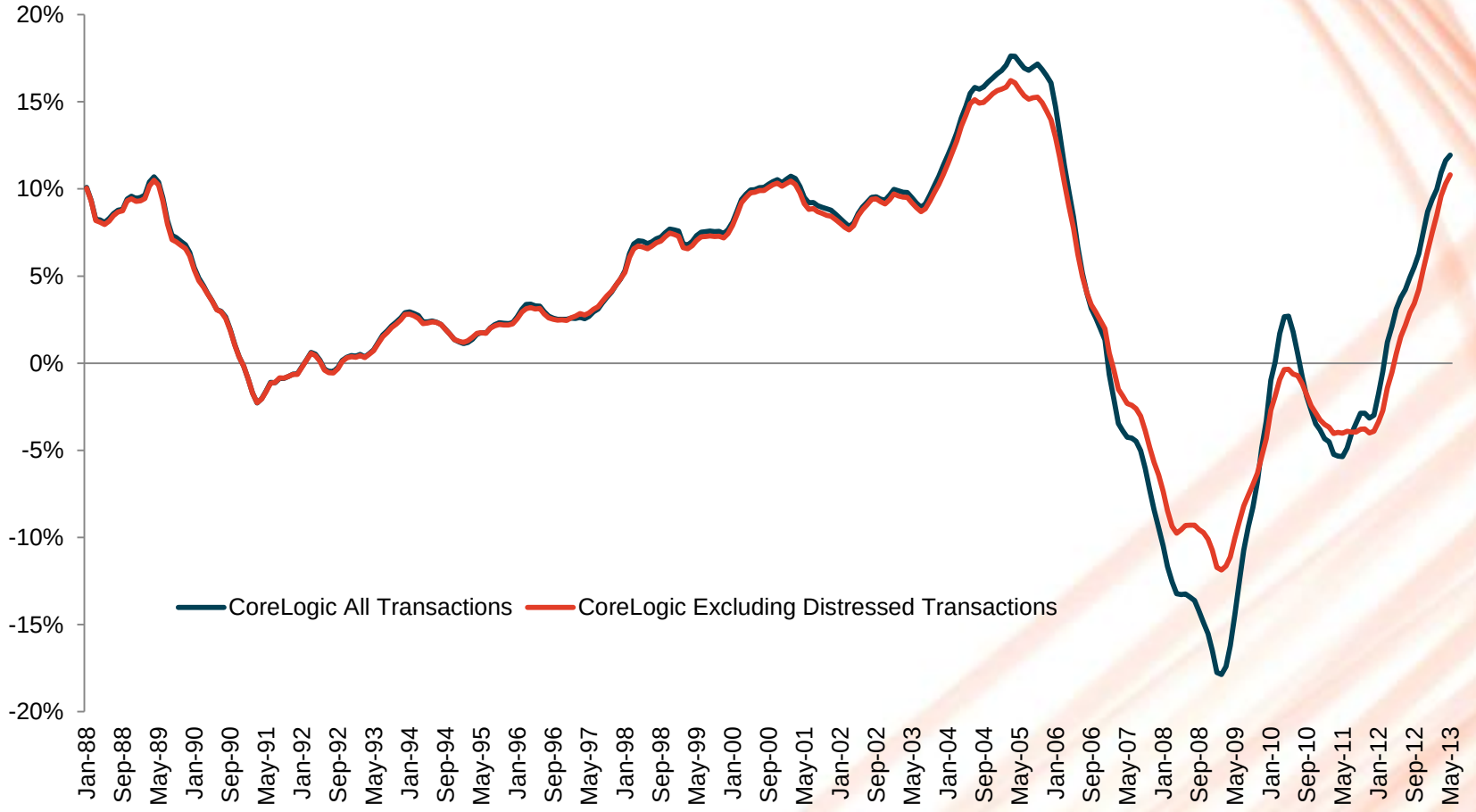


Source: CoreLogic July 2013

CoreLogic

Distressed Sales Impact Turning Points

Year-Over-Year Change



Source: CoreLogic July 2013

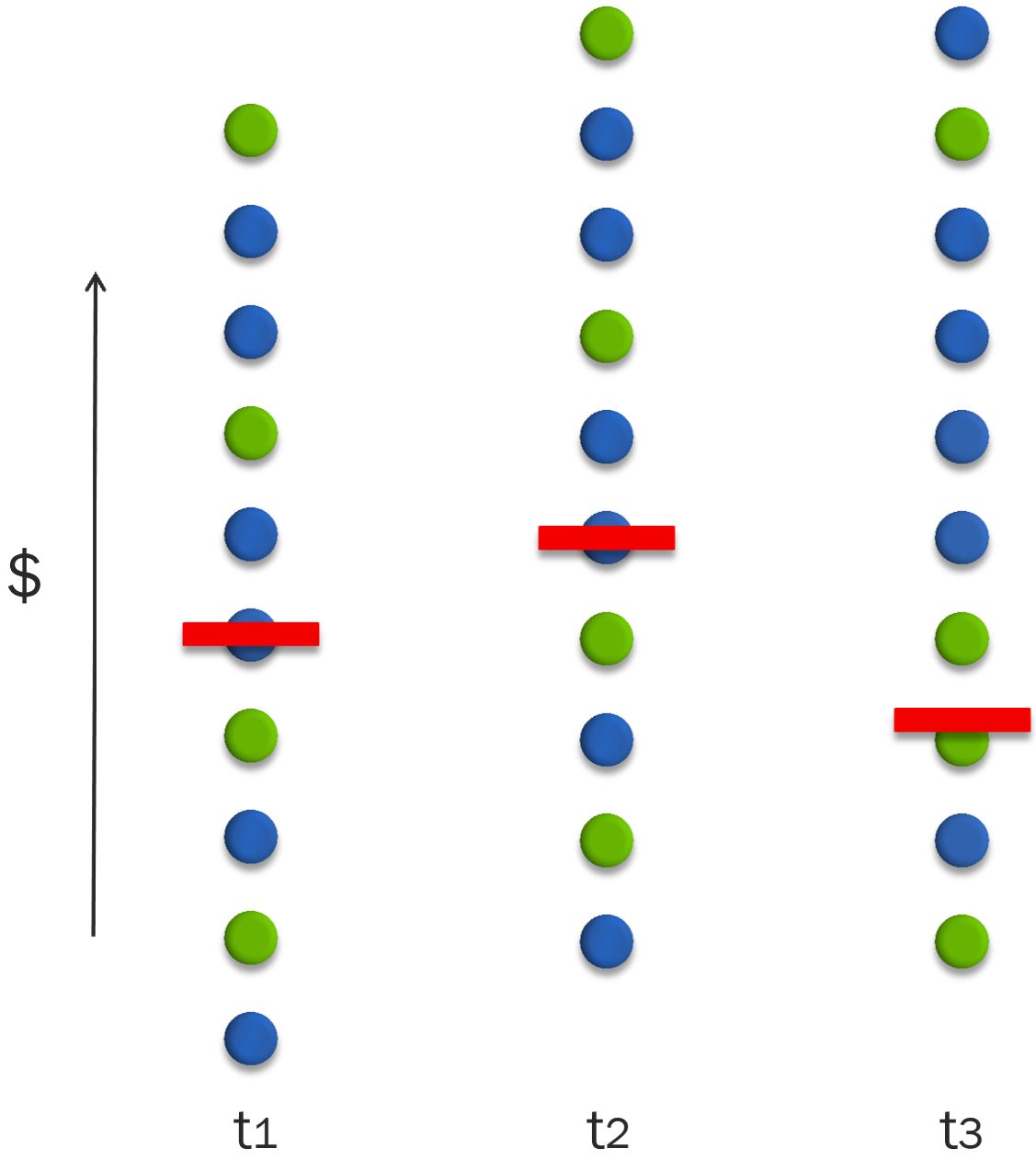
CoreLogic

Zillow®

Understanding the Zillow Home Value Index

Stan Humphries, Chief Economist

In the beginning...

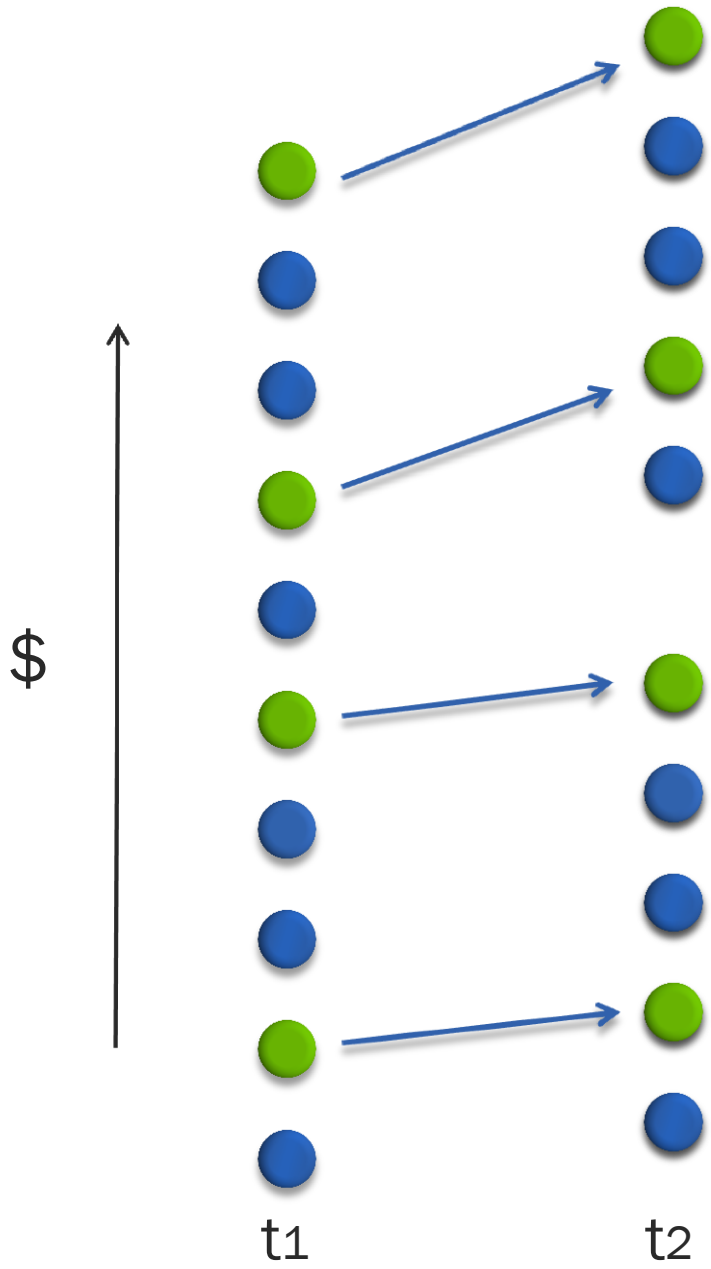


Median Sales Price

Unsold home

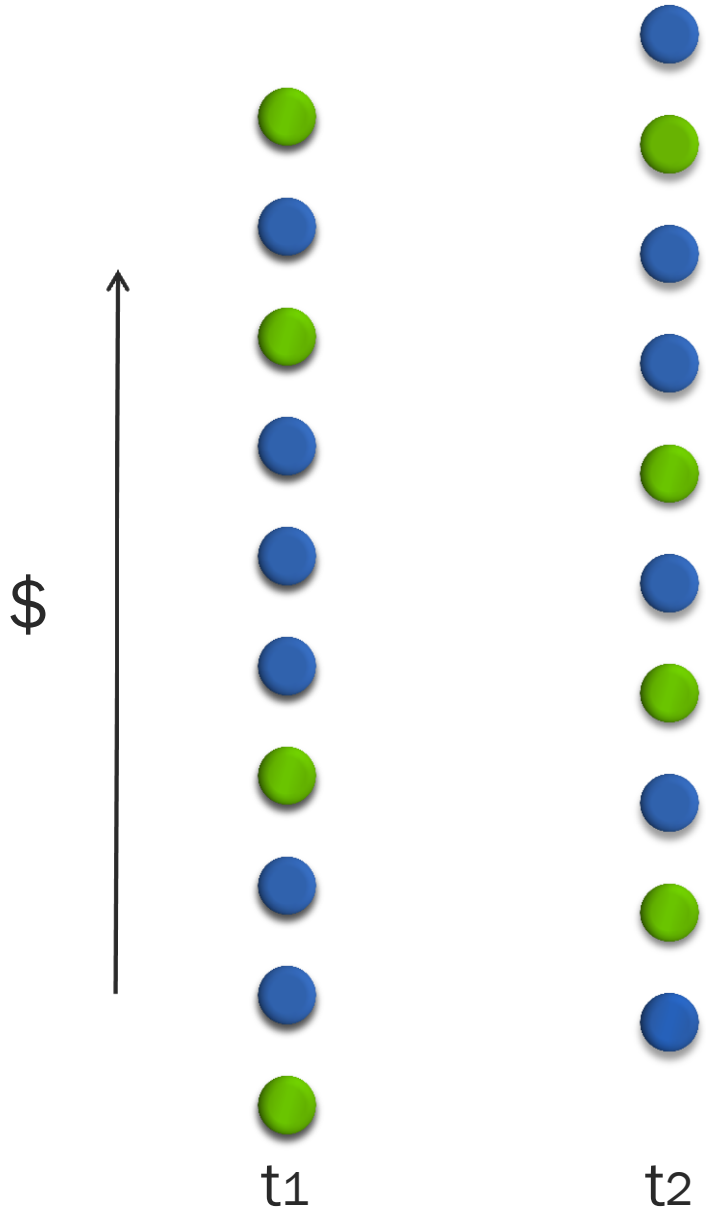
Sold home

Repeat Sales



Unsold home

Sold home

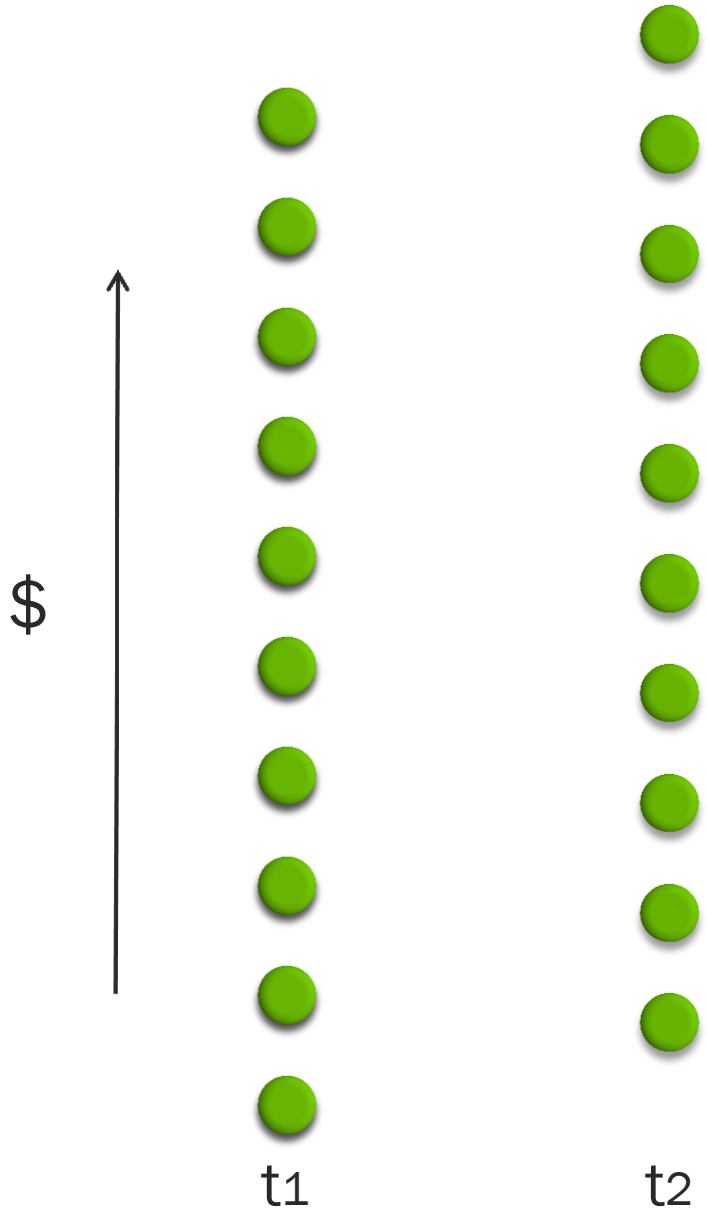


Theoretical Example:

Sales are uniform random sample

Unsold home

Sold home



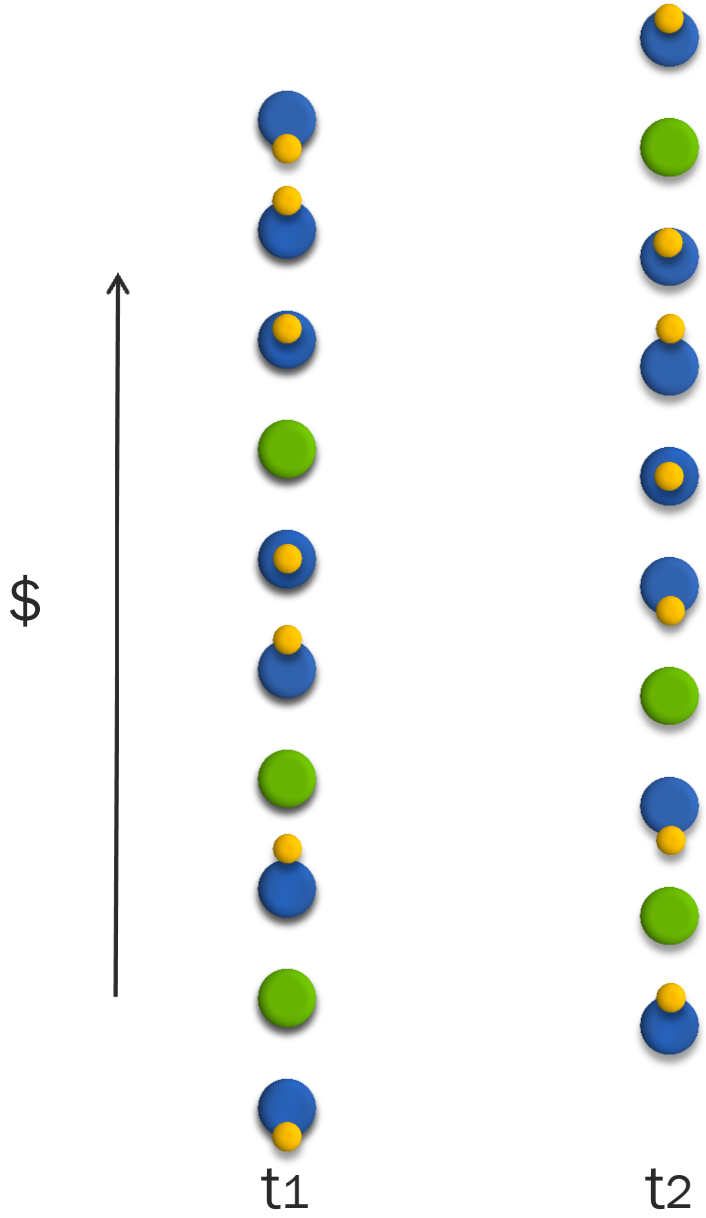
Theoretical Example:

All homes sell every time period

Unsold home

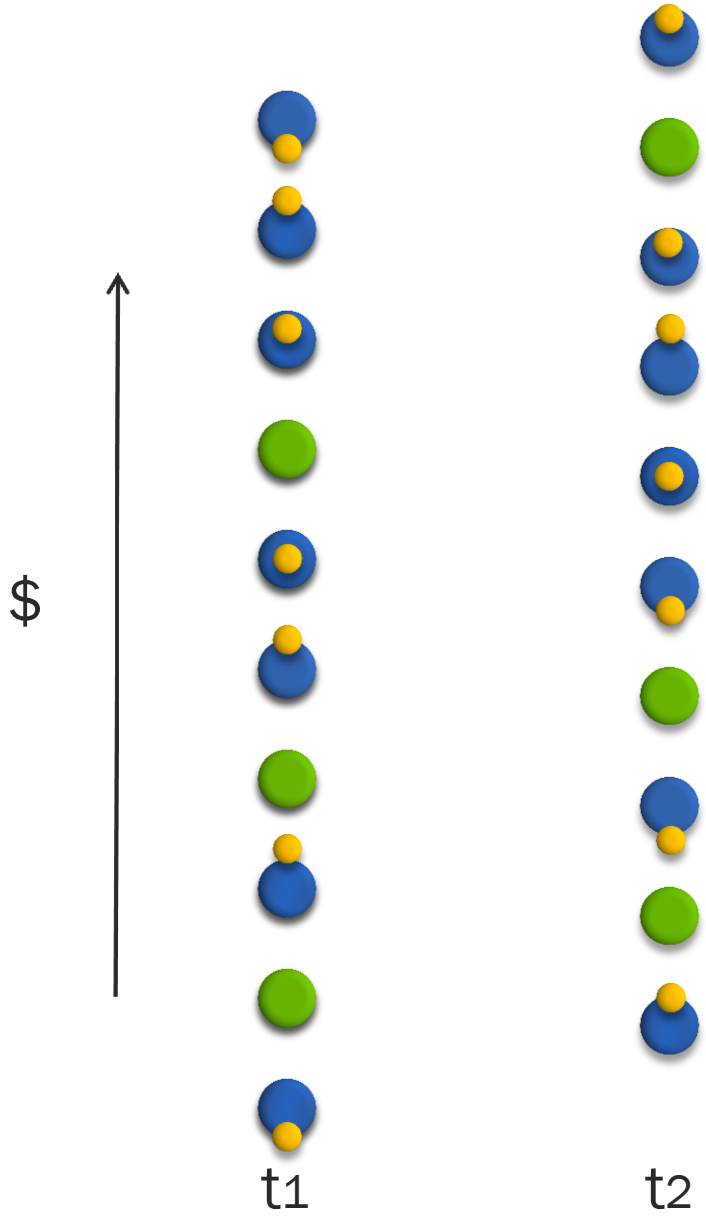
Sold home

Origins of the Zillow Home Value Index



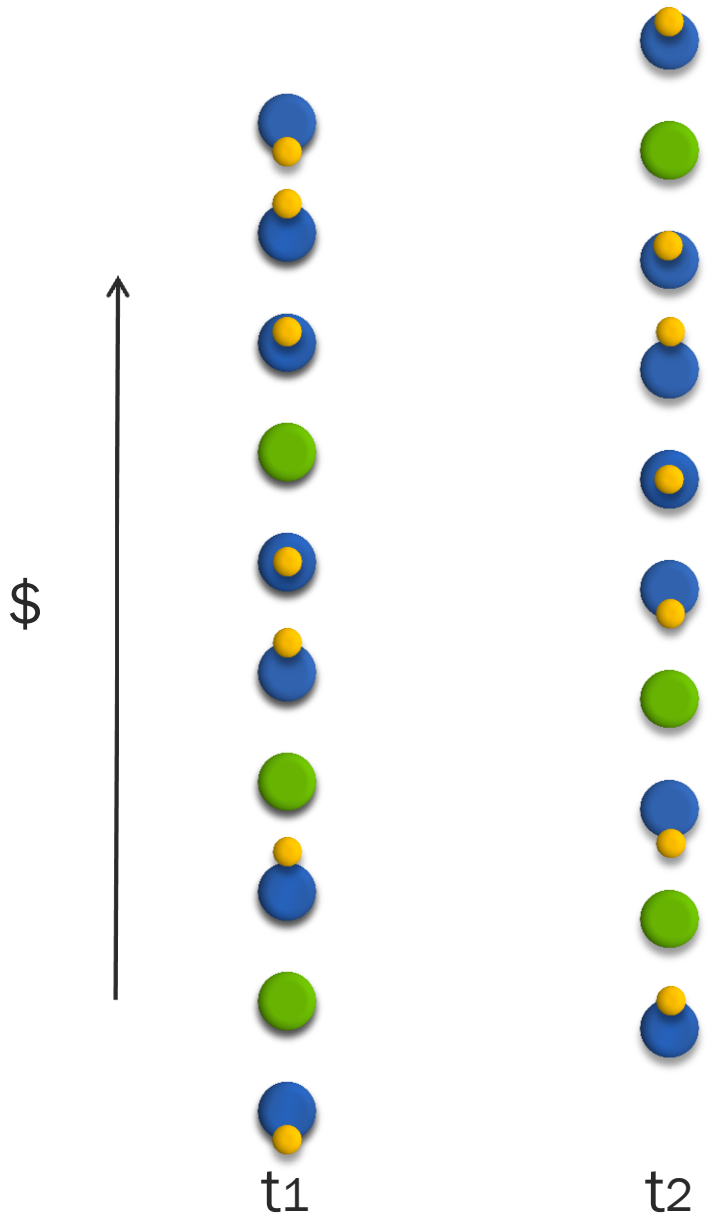
- Unsold home
- Sold home
- Home value estimate

Origins of the Zillow Home Value Index



- Unsold home
- Sold home
- Home value estimate

Median Absolute Error: 8.4%



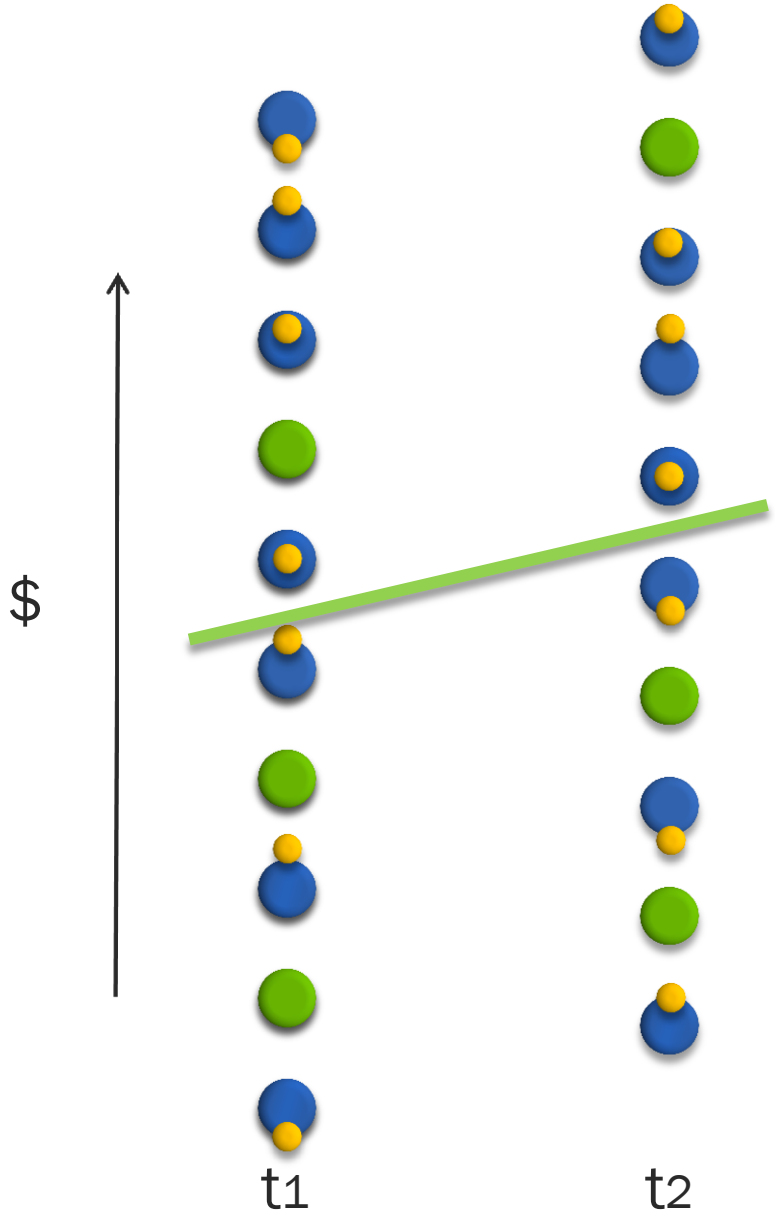
Origins of the Zillow Home Value Index

- Unsold home
- Sold home
- Home value estimate

Median Absolute Error: 8.4%

Median Error: 0.0%

Origins of the Zillow Home Value Index



Unsold home

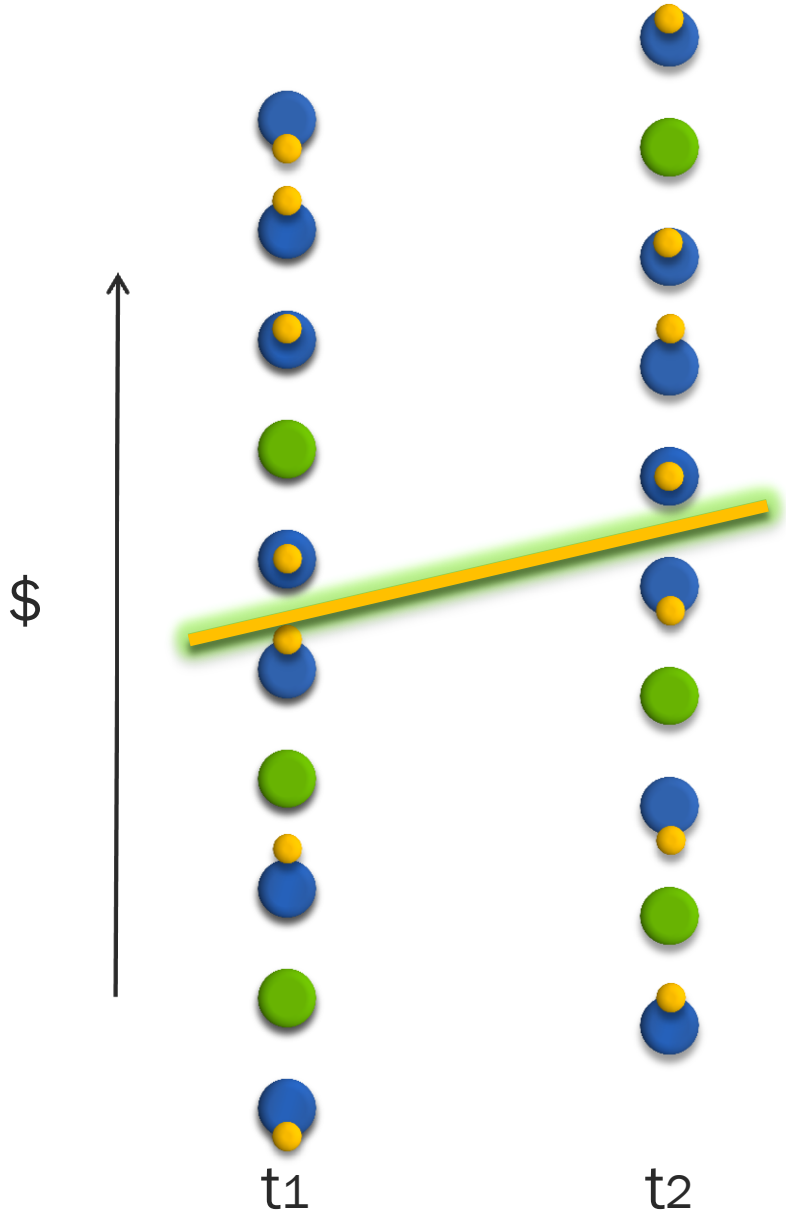
Sold home

Home value estimate

Median Absolute Error: 8.4%

Median Error: 0.0%

Origins of the Zillow Home Value Index



Unsold home

Sold home

Home value estimate

Median Absolute Error: 8.4%

Median Error: 0.0%

Data from multiple sources

Physical attributes

Tax assessments

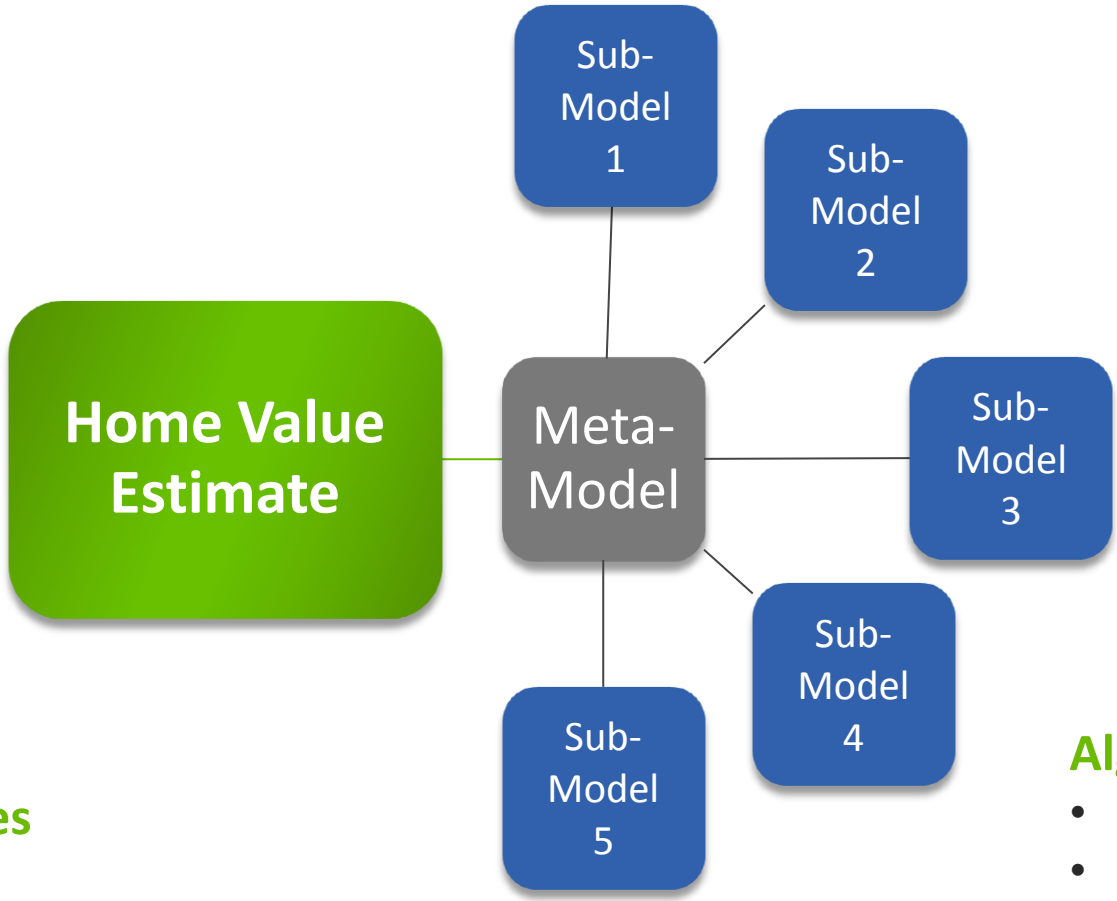
Prior sale prices

Home Value Estimate

Identifying only valid, arm's length transactions

Models trained with recently sold data

Models applied to all homes every day

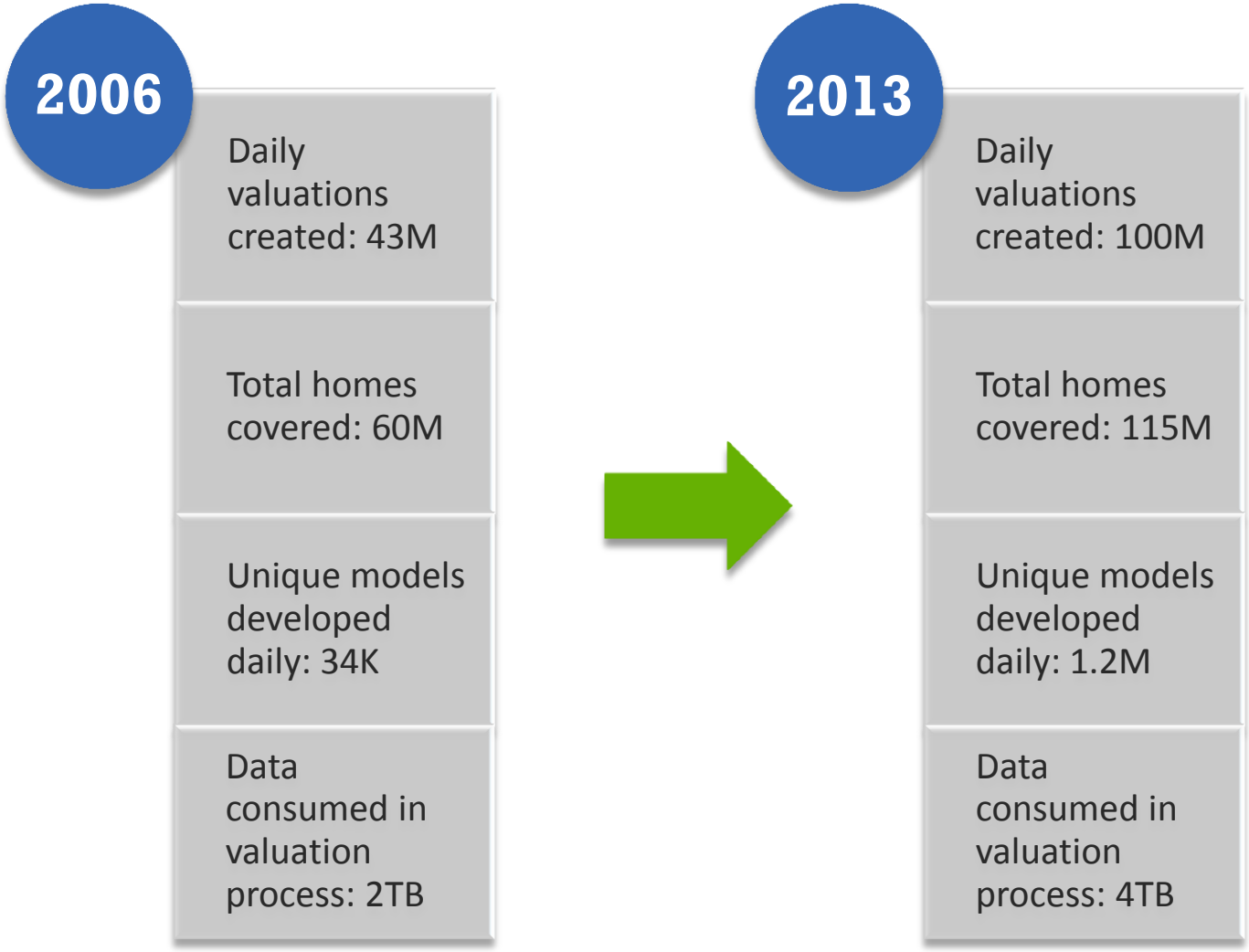


Modeling Approaches

- Hedonic
- Appraiser Emulation
- Sale Price Trending
- Tax Assessment
- Geospatial

Algorithms

- Decision Trees
- Random Forest
- SVM
- KNN
- Linear Regression


2006
Daily valuations created: 43M
Total homes covered: 60M
Unique models developed daily: 34K
Data consumed in valuation process: 2TB
2013
Daily valuations created: 100M
Total homes covered: 115M
Unique models developed daily: 1.2M
Data consumed in valuation process: 4TB

National median error: 8.4%

32% of Zestimates are within 5% of sale price

57% of Zestimates are within 10% of sale price

82% of Zestimates are within 20% of sale price

Errors are just as likely to be above the sale price as below the sale price.

Data Coverage and Zestimate Accuracy Table

Choose a location type below to change data:

Top Metro Areas
States/Counties*
National

	Zestimate Accuracy	Homes on Zillow	Homes With Zestimates	Within 5% of Sale Price	Within 10% of Sale Price	Within 20% of Sale Price	Median Error
Atlanta, GA	★★★	2.1M	1.9M	31.0%	56.5%	81.5%	8.5%
Baltimore, MD	★★★★	968.1K	911.0K	38.3%	65.1%	85.8%	6.9%
Boston, MA	★★★★	1.5M	1.4M	36.3%	65.1%	89.8%	7.1%
Chicago, IL	★★★	3.3M	3.2M	34.0%	59.7%	81.1%	7.9%
Cincinnati, OH	★★★	794.4K	709.8K	32.9%	57.6%	82.0%	8.2%
Cleveland, OH	★★★★	818.3K	727.2K	35.9%	63.7%	83.7%	7.1%
Dallas-Fort Worth, TX	★★	2.1M	1.9M	20.8%	42.8%	76.2%	11.6%
Denver, CO	★★★★	924.7K	856.0K	41.3%	72.2%	93.8%	6.2%
Detroit, MI	★★★	1.8M	1.7M	27.8%	53.0%	79.8%	9.3%
Houston, TX	★	2.1M	1.8M	--	--	--	--
Kansas City, MO	★	743.6K	691.2K	--	--	--	--
Las Vegas, NV	★★	667.5K	660.7K	20.2%	39.4%	76.0%	12.6%
Los Angeles, CA	★★★	3.1M	2.9M	26.5%	52.2%	83.5%	9.6%
Miami-Fort Lauderdale, FL	★★★	2.5M	2.4M	30.4%	56.6%	82.6%	8.6%

From data updated on August 28, 2013 and shown on Zillow.com.

	Zillow National Home Value Index ("ZHVI")
Primary Purpose	Housing market analysis and research
Methodology	Hedonic Imputation · Median of actual and estimated market values of all homes within a market (or market segment) · 3-month smoothed, using a Henderson Filter · Seasonally-adjusted only
Underlying Data	Actual and estimated values of 83 million individual single-family homes, condos, and co-ops: · Actual, non-distressed sale prices recorded during the index reporting period · Estimated non-distressed market values for every home in the Zillow database that does not sell during the reporting period · Includes newly constructed homes · Index data history to 1997
Coverage	@95% of US housing stock by market value
Release Frequency	Monthly
Reporting Lag	18-23 days

About Zillow Real Estate Research

In-depth research and monthly reports

www.zillow.com/research

www.zillowblog.com/research/local-market-reports

All data is freely available on

www.zillow.com/research/data

Real Estate Metrics

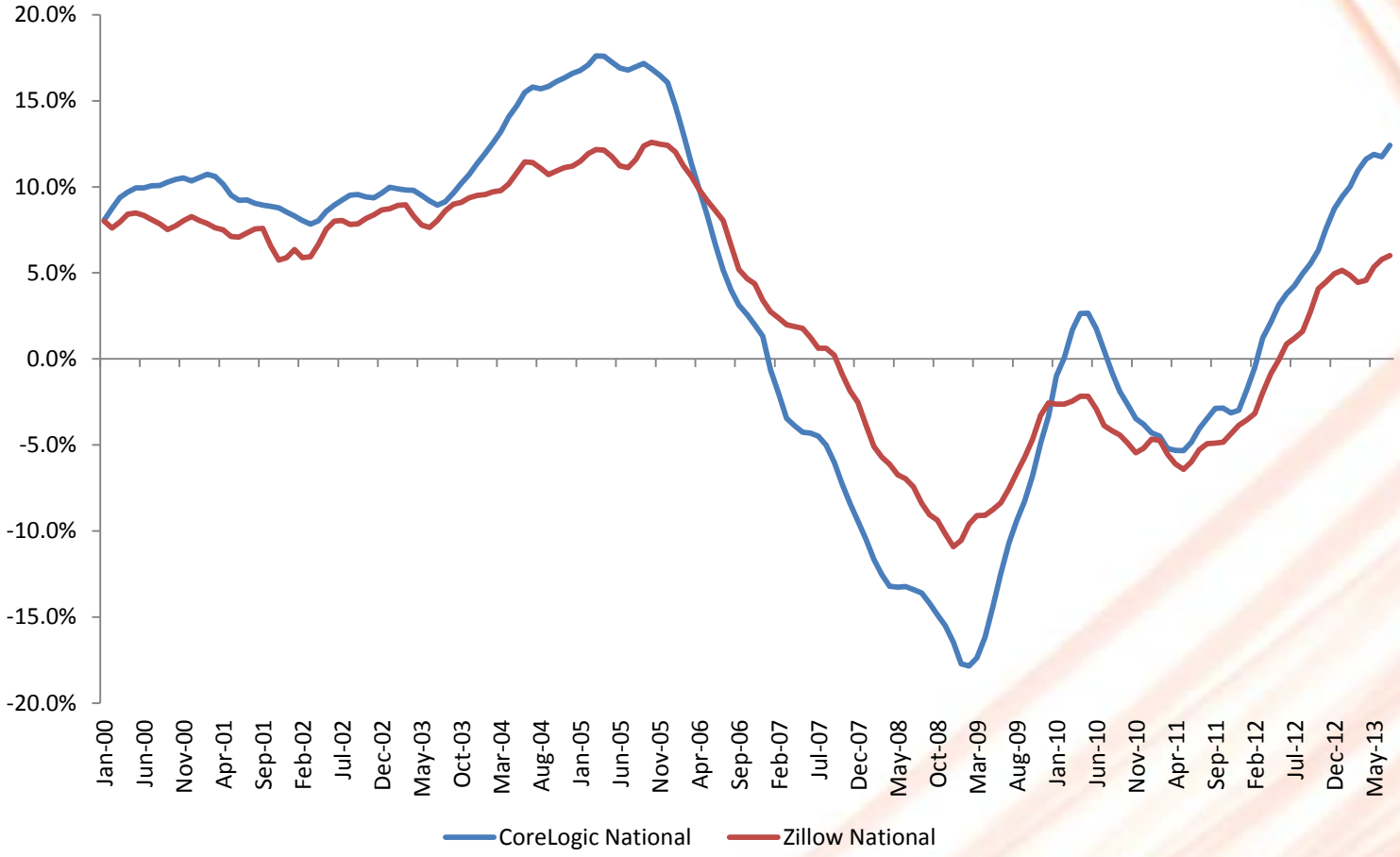
- Zillow Home Value Index
- Zillow Rent Index
- Zillow Home Value Forecast
- Negative equity
- List prices
- Sale prices
- Rental prices
- Home sales
- $ value/square foot
- $ price/square foot
- Sale-to-list price ratio
- % listings with price cuts
- $ amount of listing price cuts
- % homes sold for loss/gain
- % homes foreclosed
- % sales that are foreclosure re-sales
- % homes increasing/decreasing in value
- % homes sold in the past year
- Price-to-rent ratios
- Price-to-income ratios
- Affordability

Comparative Results

September 16th, 2013

HPI Methods Create Differentiated View of Growth

Year-Over-Year HPI Growth

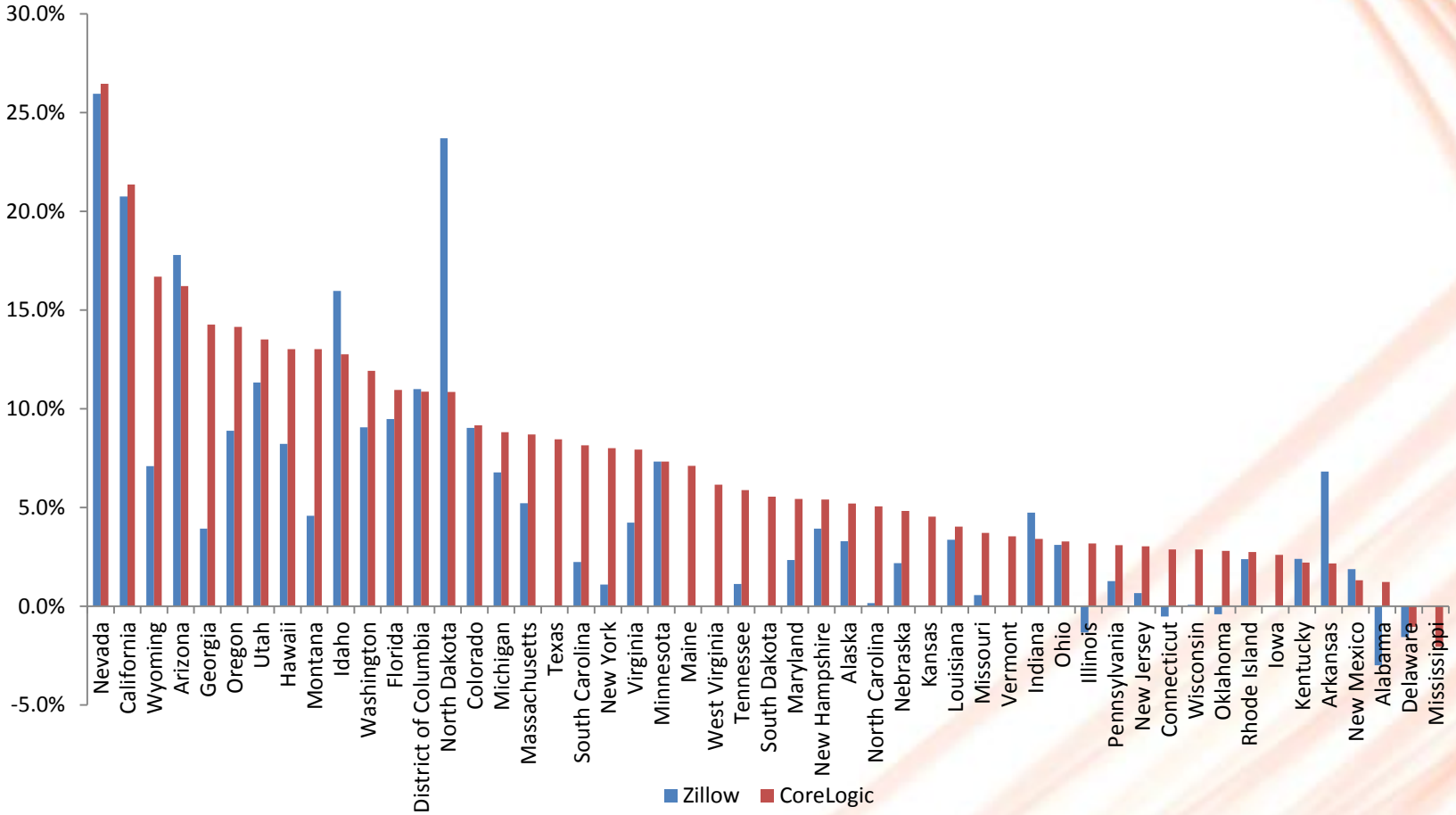


Source: ZIllow, CoreLogic September 2013

CoreLogic

HPI Methods Create Differentiated View of Growth

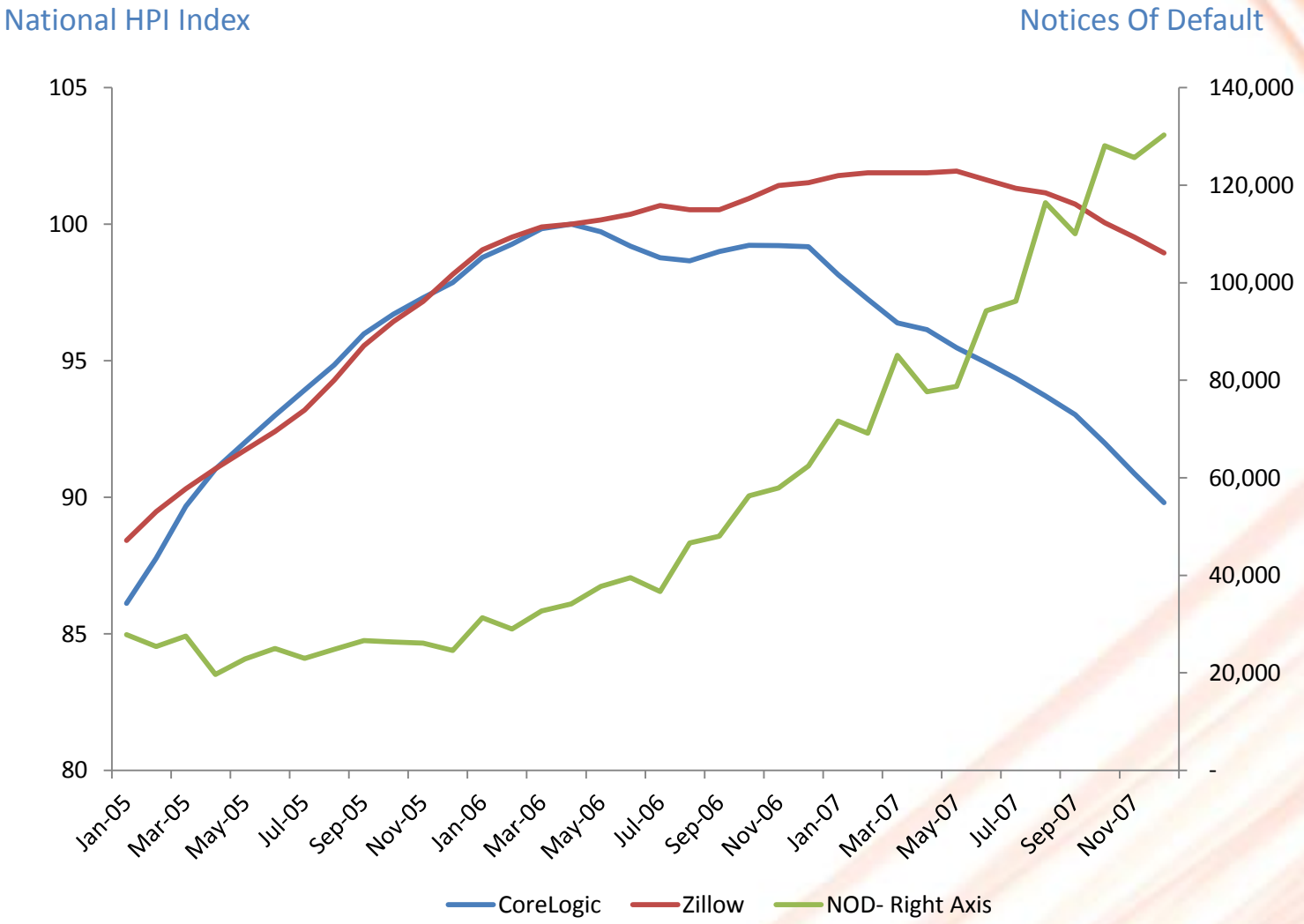
Year-Over-Year HPI Growth, June 2013



Source: ZIllow, CoreLogic September 2013

CoreLogic

As Prices Fall Distress Rises- The Relationship Between HPI and NOD

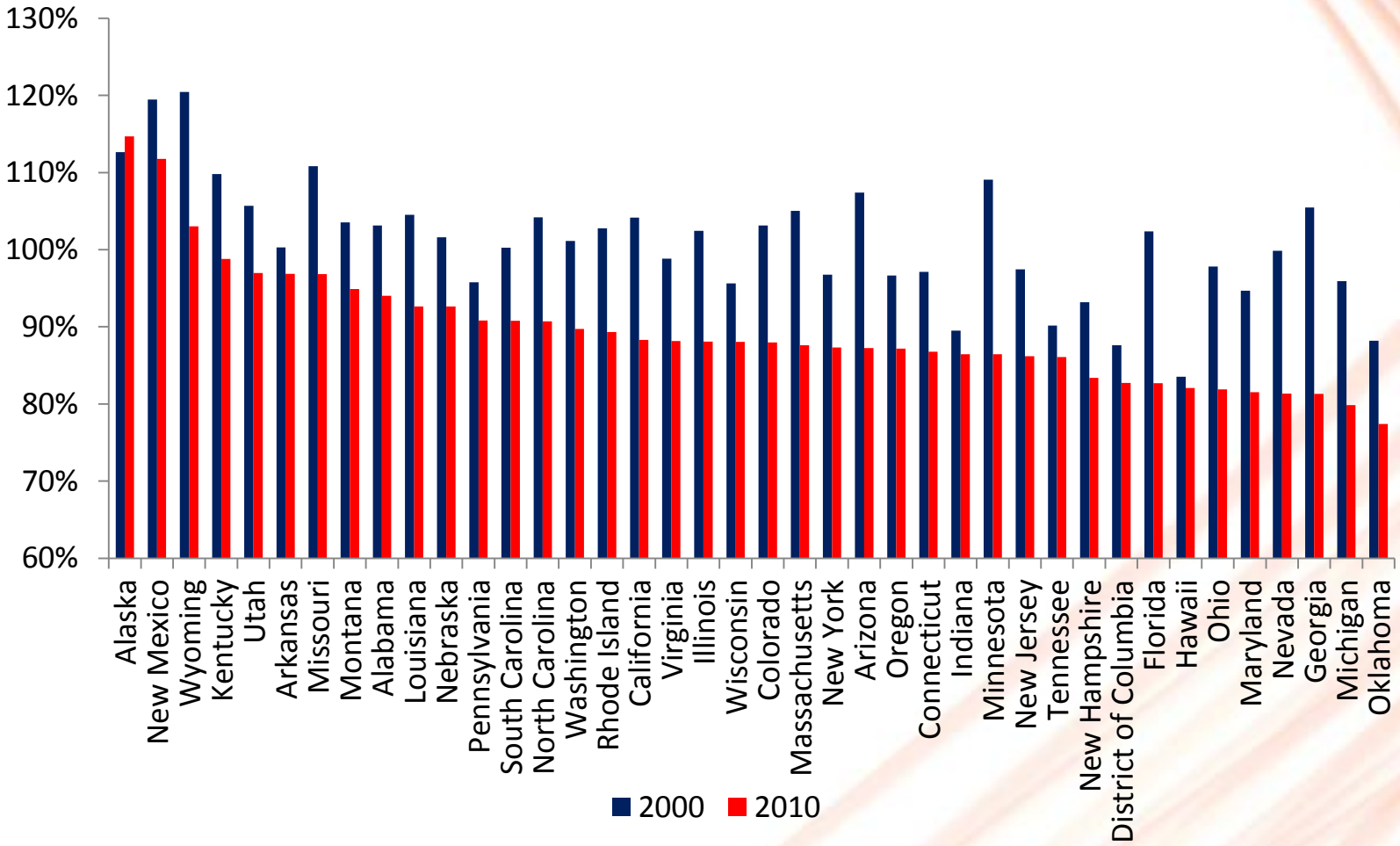


Source: CoreLogic August 2013

CoreLogic

Stock to Stock- Predicted Value Relative to Homeowner's Assessment

Zillow Median Values as a Share of Census Median Values



Source: Census, CoreLogic August 2013

CoreLogic

CoreLogic Office Of The Chief Economist

CoreLogic Economics

@corelogicecon

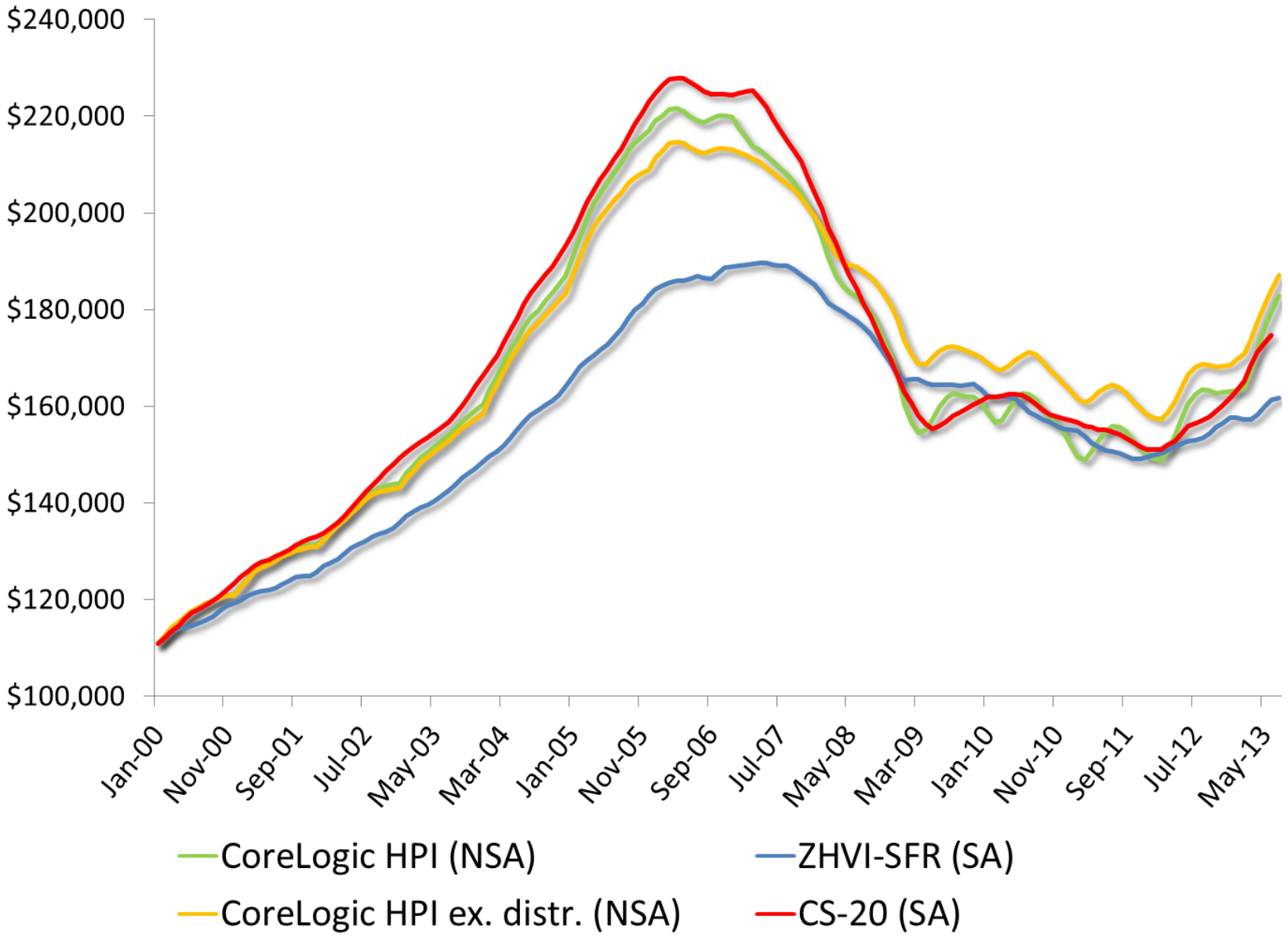
CoreLogic Insights

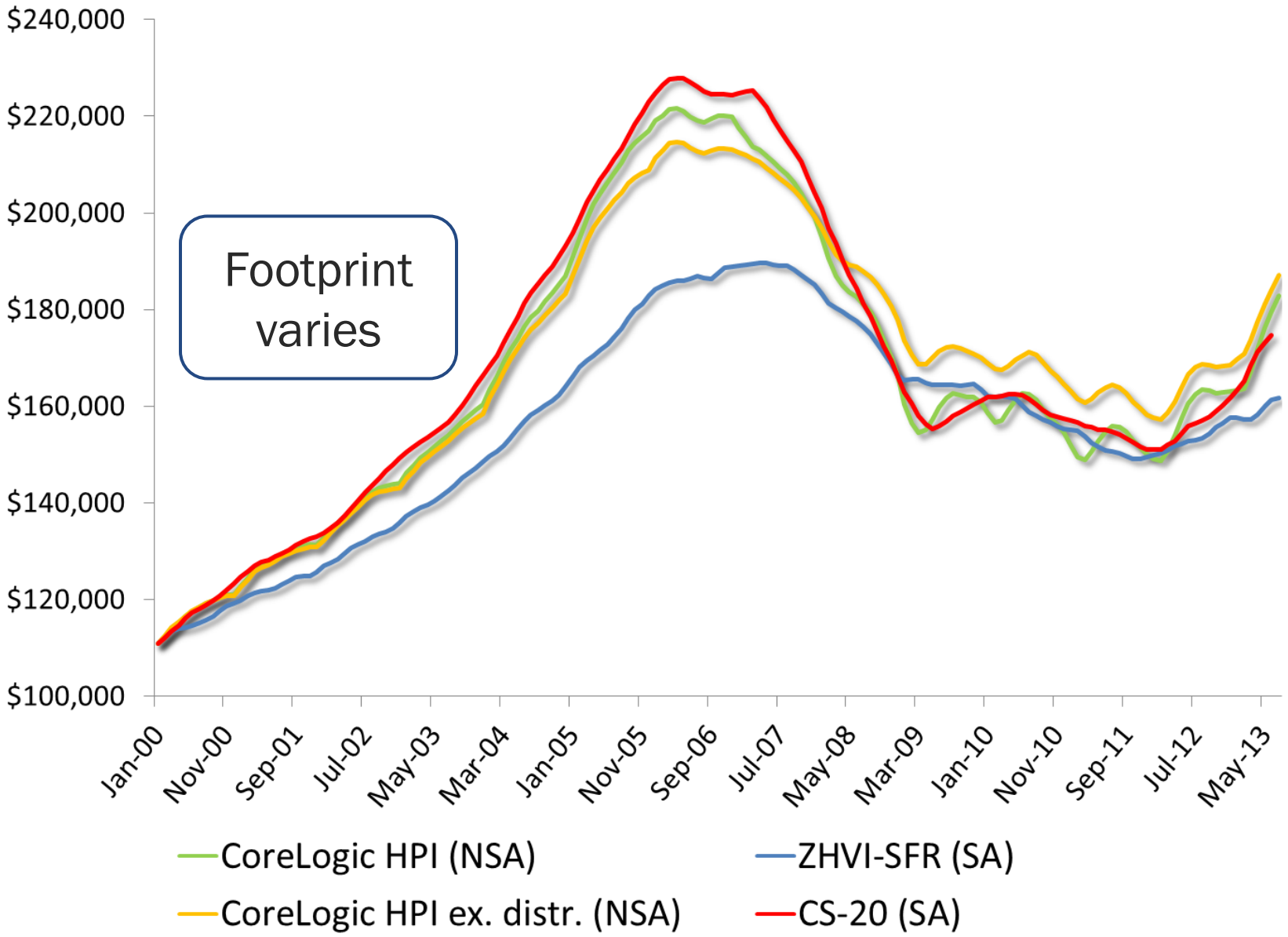
http://www.corelogic.com/insights

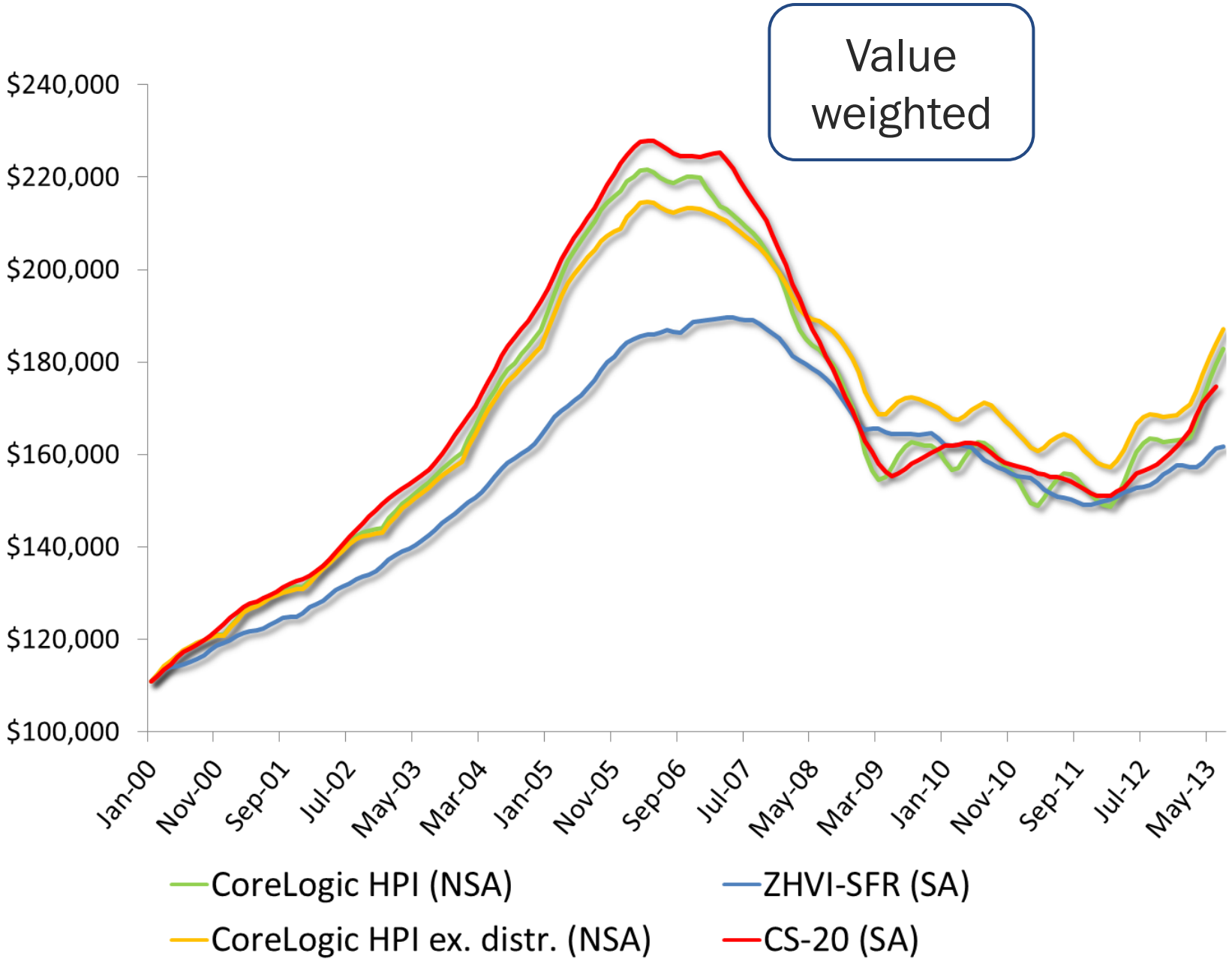
CoreLogic Academic Research Council

http://www.corelogic.com/clarc

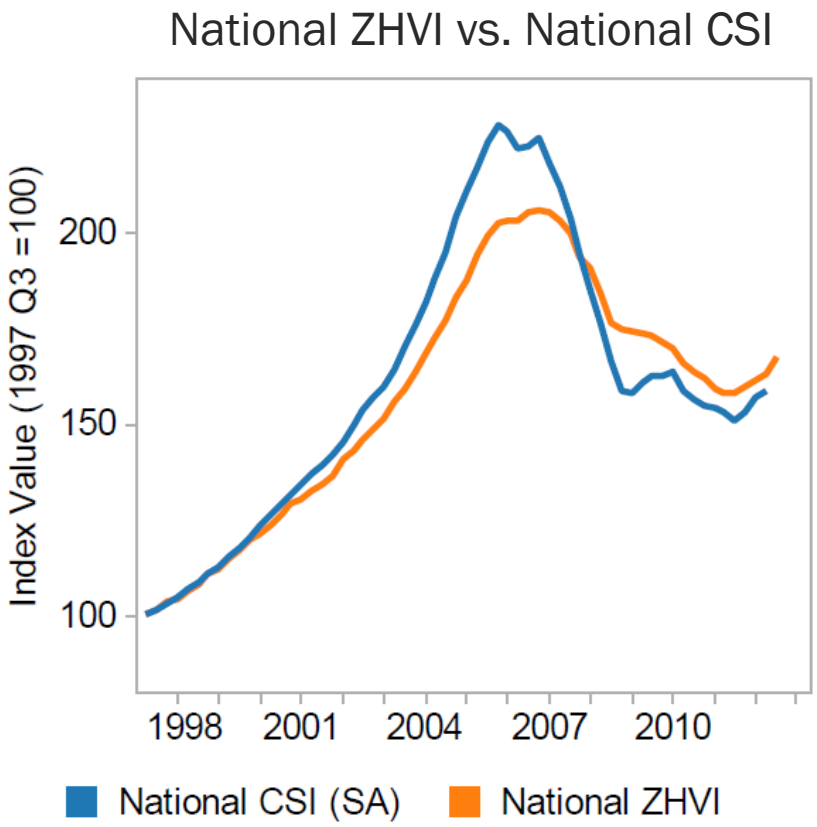
Comparing the Zillow Home Value Index to CoreLogic and Case-Shiller

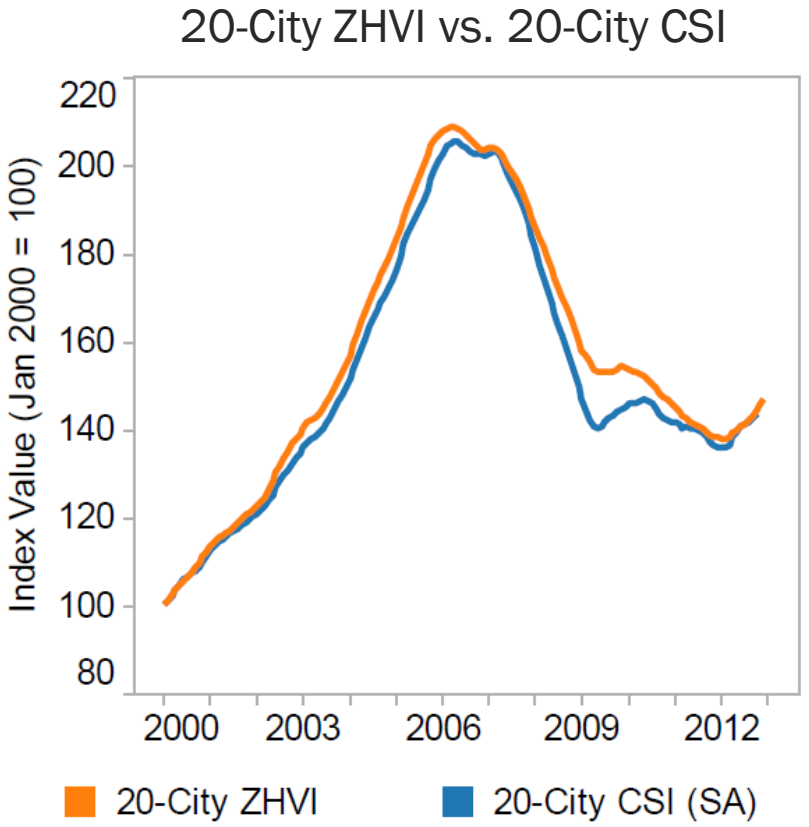

$240,000
$220,000
$200,000
$180,000
$160,000
$140,000
$120,000
$100,000
Jan-00
Nov-00
Sep-01
Jul-02
May-03
Mar-04
Jan-05
Nov-05
Sep-06
Jul-07
May-08
Mar-09
Jan-10
Nov-10
Sep-11
Jul-12
May-13
CoreLogic HPI (NSA)
ZHVI-SFR (SA)
CoreLogic HPI ex. distr. (NSA)
CS-20 (SA)

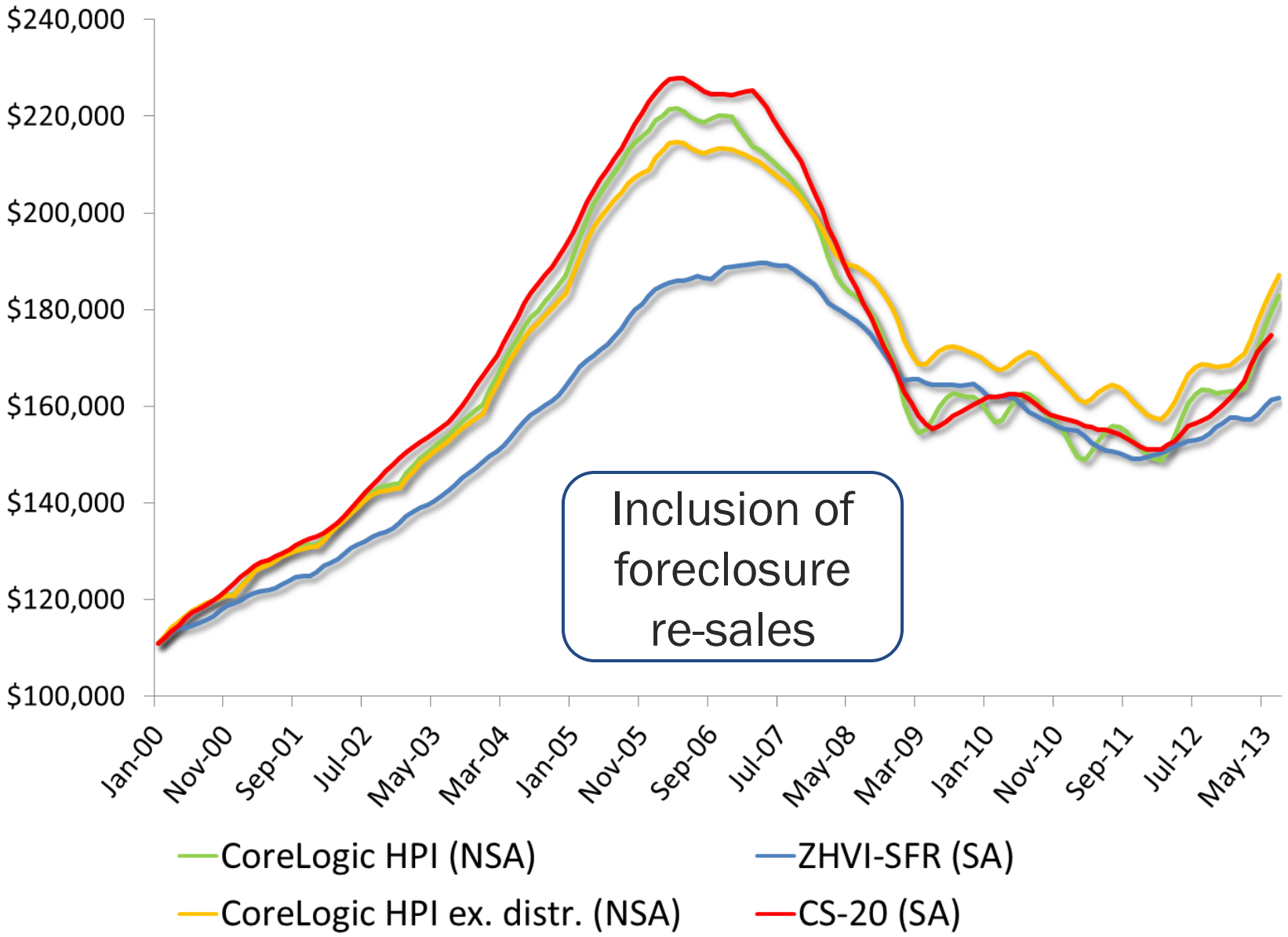

$240,000
$220,000
$200,000
$180,000
$160,000
$140,000
$120,000
$100,000
Footprint varies
Jan-00
Nov-00
Sep-01
Jul-02
May-03
Mar-04
Jan-05
Nov-05
Sep-06
Jul-07
May-08
Mar-09
Jan-10
Nov-10
Sep-11
Jul-12
May-13
CoreLogic HPI (NSA)
ZHVI-SFR (SA)
CoreLogic HPI ex. distr. (NSA)
CS-20 (SA)


Value weighted
$240,000
$220,000
$200,000
$180,000
$160,000
$140,000
$120,000
$100,000
Jan-00
Nov-00
Sep-01
Jul-02
May-03
Mar-04
Jan-05
Nov-05
Sep-06
Jul-07
May-08
Mar-09
Jan-10
Nov-10
Sep-11
Jul-12
May-13
CoreLogic HPI (NSA)
ZHVI-SFR (SA)
CoreLogic HPI ex. distr. (NSA)
CS-20 (SA)

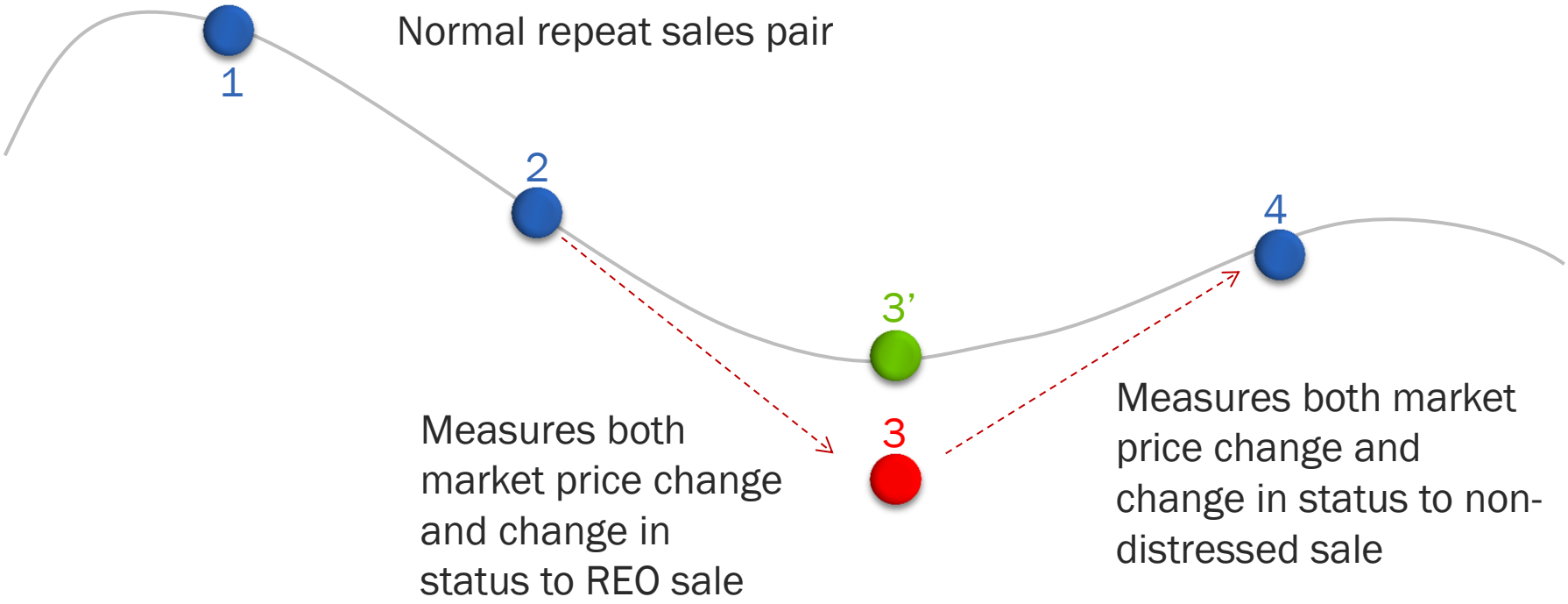
The impact of footprint and value weighting






$240,000
$220,000
$200,000
$180,000
$160,000
$140,000
$120,000
$100,000
Jan-00
Nov-00
Sep-01
Jul-02
May-03
Mar-04
Jan-05
Nov-05
Sep-06
Jul-07
May-08
Mar-09
Jan-10
Nov-10
Sep-11
Jul-12
May-13
Inclusion of foreclosure re-sales
CoreLogic HPI (NSA)
ZHVI-SFR (SA)
CoreLogic HPI ex. distr. (NSA)
CS-20 (SA)

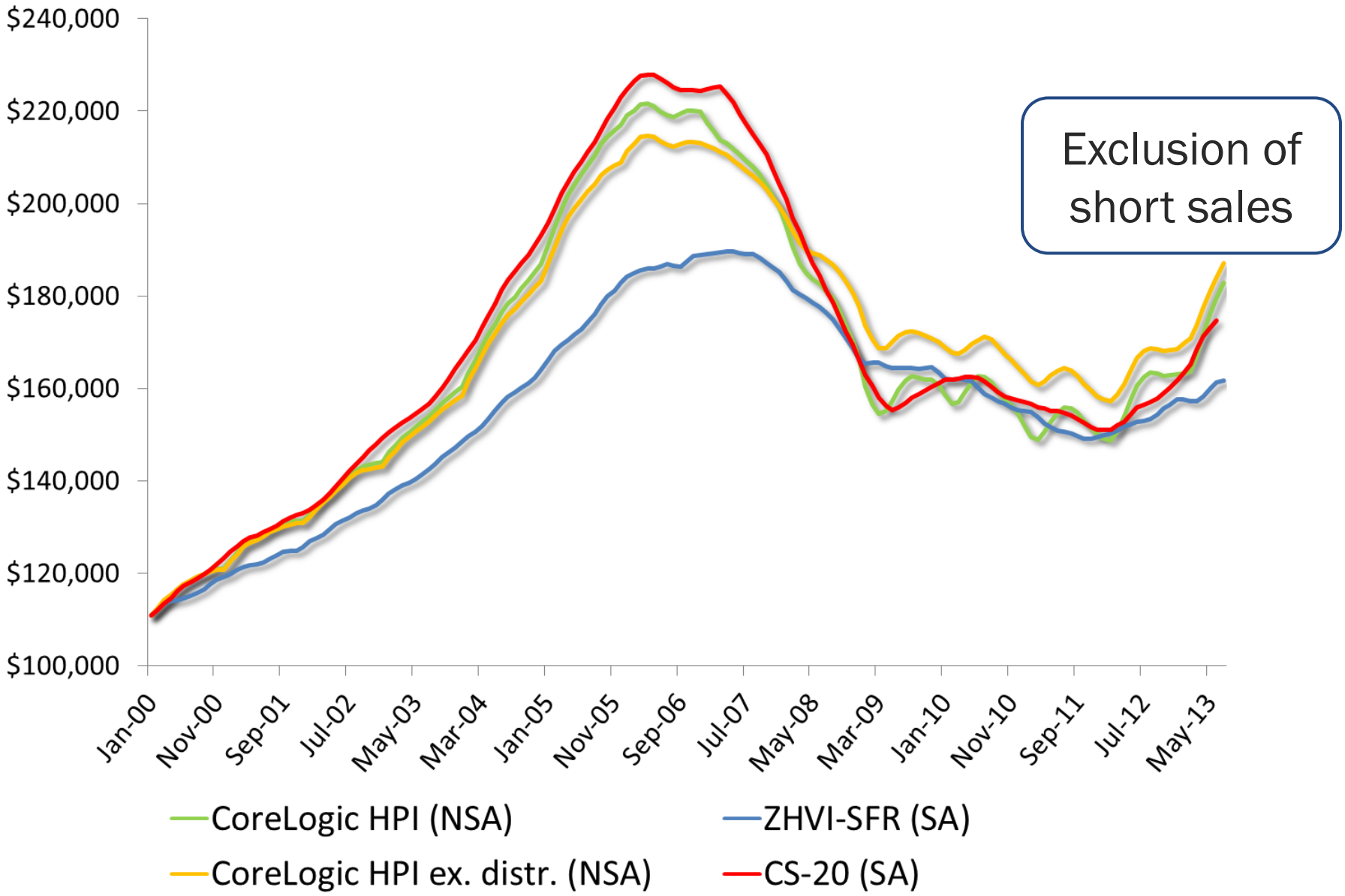
Problems with the inclusion of foreclosure re-sales



Non-distressed sale

Non-distressed value of home

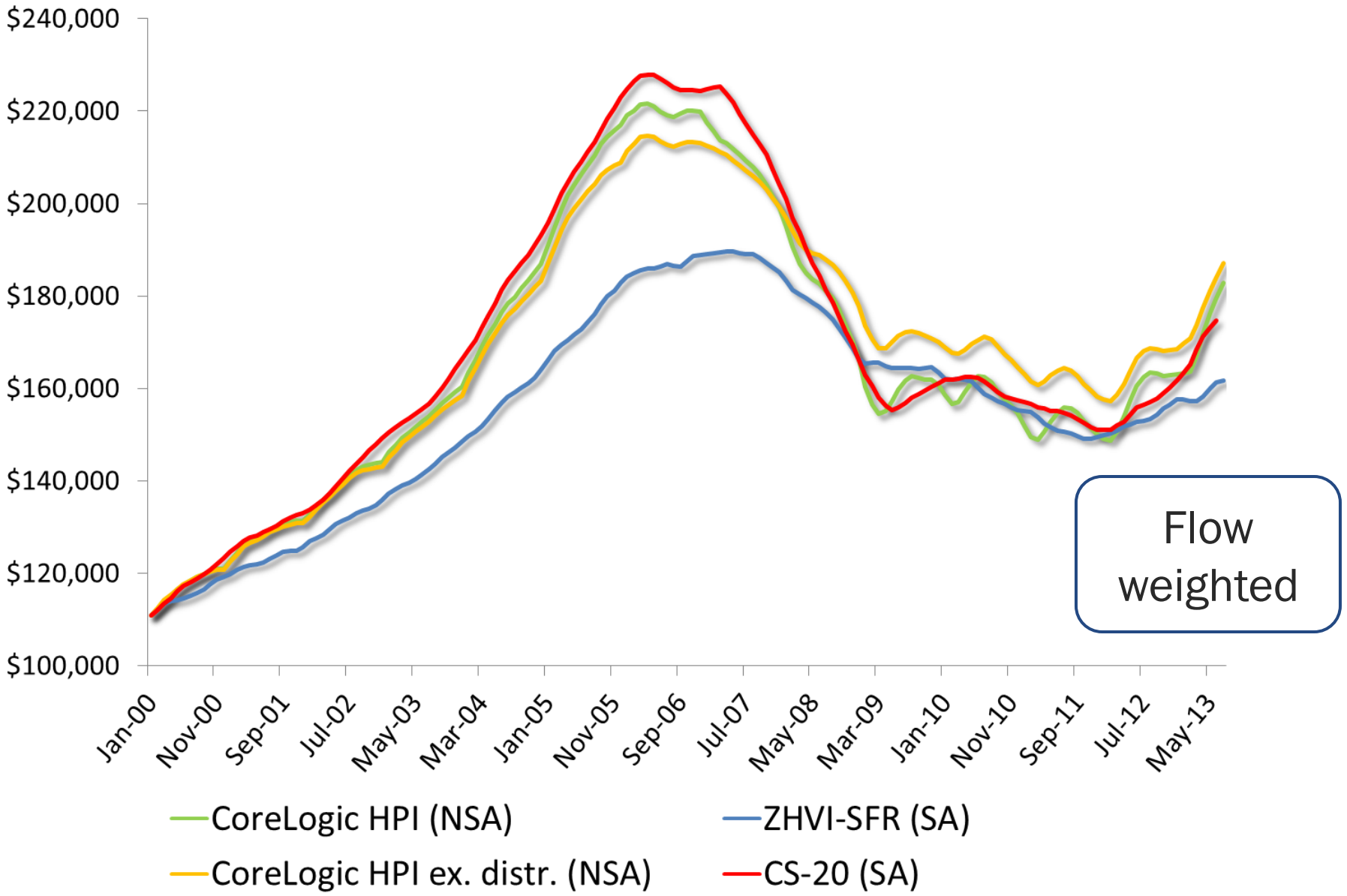
REO Sale


$240,000
$220,000
$200,000
$180,000
$160,000
$140,000
$120,000
$100,000
Jan-00
Nov-00
Sep-01
Jul-02
May-03
Mar-04
Jan-05
Nov-05
Sep-06
Jul-07
May-08
Mar-09
Jan-10
Nov-10
Sep-11
Jul-12
May-13
CoreLogic HPI (NSA)
ZHVI-SFR (SA)
CoreLogic HPI ex. distr. (NSA)
CS-20 (SA)
Exclusion of short sales

Excluding short sales:
A noble endeavor but does it lead to upward bias?

- **Current approach (as we understand it):**
  - For sales after 2006, the sale is removed from the pairing process if the sale price is less than the mortgage amount at time of origination (no assumptions made about remaining principal).
- **Example:**
  - 2007: Sale of home for $100K ($80K mortgage)
  - 2013: Home sells again for $75K (but mortgage balance is only $70K)

  → This home would be excluded from the index despite not being a short sale.
- **This can easily become a filter for homes that have declined in value but are not, in fact, short sales. The exclusion of such transactions will result in an upward bias in the estimate of home price appreciation.**


$240,000
$220,000
$200,000
$180,000
$160,000
$140,000
$120,000
$100,000
Jan-00
Nov-00
Sep-01
Jul-02
May-03
Mar-04
Jan-05
Nov-05
Sep-06
Jul-07
May-08
Mar-09
Jan-10
Nov-10
Sep-11
Jul-12
May-13
CoreLogic HPI (NSA)
CoreLogic HPI ex. distr. (NSA)
ZHVI-SFR (SA)
CS-20 (SA)
Flow weighted

Goldman Sachs

How much did house prices really increase in 2012?

April 4, 2013

percentage point difference in the 2012 house price growth rate is entirely driven by the weighting method.

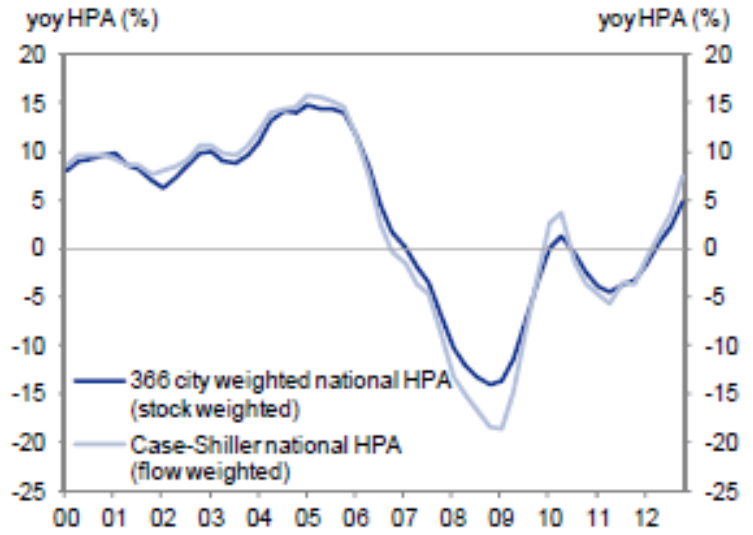
The analysis above shows that the national house prices, weighted by the housing stock of each location, probably increased 3-4% rather than 7-8% in 2012. The difference can be explained by a declining share of distressed sales and the weighting method used in repeat-sales index construction. This result is intuitive considering the large number of cities in the US where house prices tend to stay flat for years and rarely experience significant appreciation or depreciation. For example, out of the 366 metro areas in our

Exhibit 1: House prices appreciation rates turned from negative in 2011 to positive in 2012 across indices

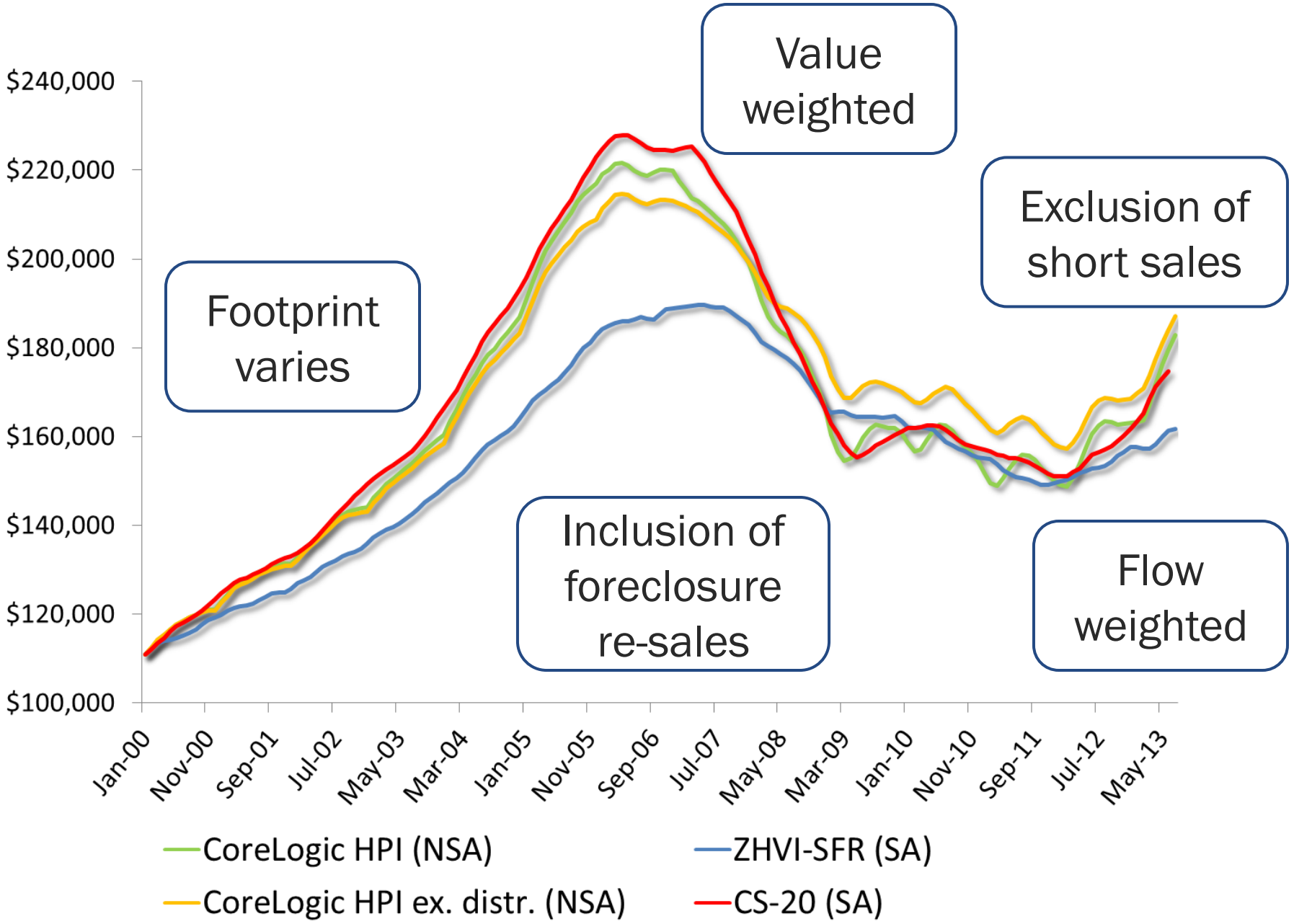
House Price Index	HPA (yoy %Chg) 2011	2012
S&P/Case-Shiller Composite 20	-4.1	6.8
S&P/Case-Shiller US National	-3.7	7.3
CoreLogic National incl Distressed	-3.1	8.3
CoreLogic National excl Distressed	-4.1	6.7
FHFA Purchase-Only Monthly Index	-1.3	5.6
Zillow.com Home Value Index	-3.4	5.6
Radar Logic 25 MSA Composite	-6.4	13.3
FNC Composite 30	-4.4	5.8
NAR Median Sales Price	-3.9	11.1

Source: Haver Analytics, GS Mortgage Strategy Research.

Exhibit 4: Flow weighted house price index shows a bigger decline in 2009 and a sharper recovery in 2012



Source: Fiserv, GS Mortgage Strategy Research.


Value weighted
Exclusion of short sales
Footprint varies
Inclusion of foreclosure re-sales
Flow weighted
$240,000
$220,000
$200,000
$180,000
$160,000
$140,000
$120,000
$100,000
Jan-00
Nov-00
Sep-01
Jul-02
May-03
Mar-04
Jan-05
Nov-05
Sep-06
Jul-07
May-08
Mar-09
Jan-10
Nov-10
Sep-11
Jul-12
May-13
CoreLogic HPI (NSA)
CoreLogic HPI ex. distr. (NSA)
ZHVI-SFR (SA)
CS-20 (SA)


Zillow®