

# Measuring Aggregate Housing Wealth: New Insights from Automated Valuation Models

Joshua Gallin, Eric Nielsen, Raven Molloy, Paul Smith, and Kamila Sommer

Federal Reserve Board  
NBER CRIW Summer Institute

July 17, 2018

# Disclaimer and Thanks

The analysis and conclusions set forth here are those of the authors and do not indicate concurrence by other members of the research staff, the Board of Governors, or the Federal Reserve System.

We thank Zillow for providing the Zestimate data and for very helpful discussions about their construction. Max Miller and Hannah Hall provided excellent research assistance.

# Why care about aggregate household housing wealth?

## Housing wealth:

- Is a major component of total household wealth.
  - About two-thirds of a typical household's total assets (SCF 2013).
  - About 25% of aggregate wealth, or roughly \$25 trillion (FAUS).
- Affects many aspects of household financial decision-making.
  - E.g., consumption, savings, small business formation.
- Was a key driver of wealth changes during the Great Recession.

# What is aggregate household housing wealth?

- National aggregate of housing owned for *personal* use:
  - Owner-occupied.
  - Vacation/seasonal properties.
  - Vacant-for-sale, vacant for other purposes.
- Corresponds to concepts in the Financial Accounts of the U.S (FAUS).
  - Frequent input into empirical and quantitative macro models/analyses.
- Easy to define, but hard to estimate.
  - Need both prices and quantities.
  - Prices are difficult to measure.

# Two main current measurement methods—neither perfect.

- *Owner self-reports:*

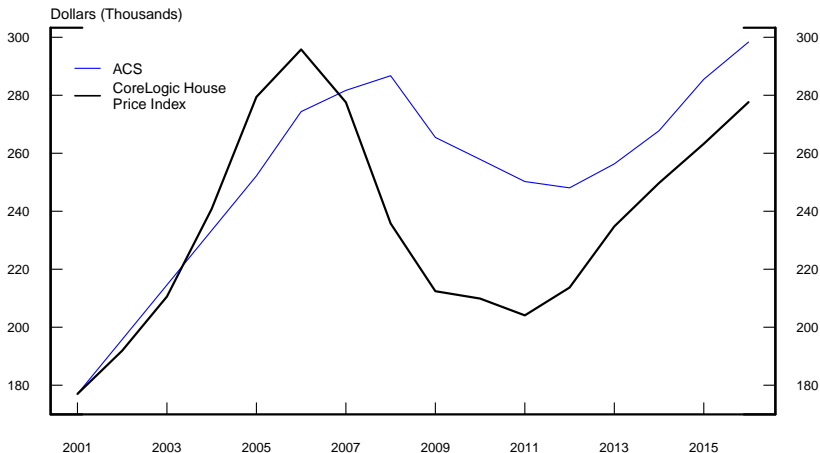
- (+) Value entire stock of owner-occupied homes.
- (+) Captures quality changes.
- (−) Behavioral biases – lagged recognition of changing market conditions + overoptimism.
  - Ihlanfeldt and Martinez-Vasquez (1986), Goodman and Ittner (1992), Kiel and Zabel (1999), Bucks and Pence (2006), Henriques (2013), Chan, Datrup, and Ellen (2014), Benitez-Silva et al. (2016), Davis and Quintin (2016), ...

- *Repeat-sales HPI:*

- (+) Market-price driven.
- (−) Transacting homes only (includes invest. properties).
- (−) Holds quality constant.
  - Case, Pollakowski, and Wachter (1997), Gatzlaff and Haurin (1997), Dreiman and Pennington-Cross (2004), Glennon, Kiefer, and Mayock (2016), ...

# These two methods disagree over the Great Recession.

## ACS Average Home Values vs. CoreLogic



Source: American Community Survey (U.S. Census Bureau), CoreLogic, and Zillow

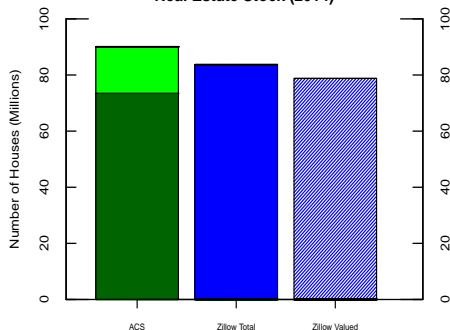
# A new contender: Zillow's AVM.

- Suite of machine learning models:
  - Property/location characteristics + sales data = value estimates.
  - Very rich data – water views, local geographic amenities, etc.
- AVM advantages:
  - Captures changing characteristics (like self reports).
  - Disciplined by market prices (like repeat-sales indexes).
  - Estimates cover a substantial fraction of housing stock.
- AVM disadvantages (more on these next):
  - Data/estimates not designed to be nationally representative.
  - AVM is a black box (to us).
- Our data (custom delivery from Zillow):
  - Average and total value by property type (mf, sf) by county/month.
  - Error distributions, average errors by price decile.

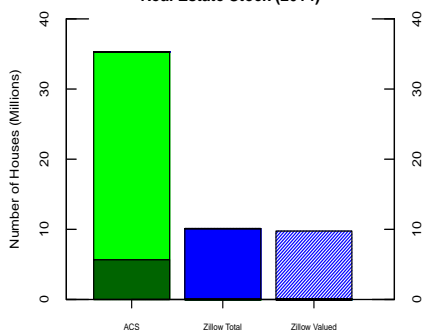
# Zillow universe is different from national *own-use* universe.

- Estimates missing for units with missing data, high forecasted model error, or with too few comparable properties.
- Zillow does not distinguish own-use from rental units.

Comparison of Single Family Residential Real Estate Stock (2014)



Comparison of Condo/Coop Residential Real Estate Stock (2014)



ACS bar includes owned properties (dark green) and rented properties (light green)



# Our method combines Zillow's AVM with Census property counts.

- Coverage issues prevent us from simply summing up Zillow's estimates to a national total.
- Solution: Use American Community Survey (ACS) property counts.
  - Nationally representative, consistent over time, and can break out rental units.
- For each county  $i$ , year  $t$ , and property type  $c$  (sf and mf):

$$\text{Aggregate Value}_{i,c,t} = \text{Property Counts}_{\text{ACS},i,c,t} \times \text{Avg. Value}_{\text{Zillow},i,c,t}$$

- Sum across counties/states and property types to get the national aggregate.

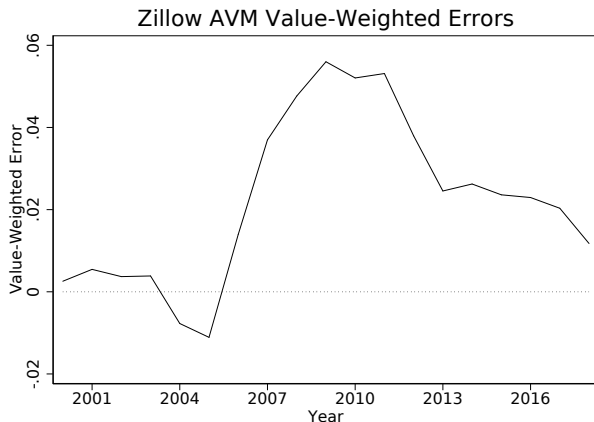
ACS splits

ACS/Zillow mapping

# Assumption 1: Zillow's average is an unbiased estimate of the true average.

- Need the average to be unbiased for each:
  - Property type (mf and sf).
  - Geography (county and state).
  - Time period.
- We can compare model predictions against market transactions.
  - Want errors to be close to 0, on average.
  - Want property-level errors to be uncorrelated.
- From Zillow, we have average errors by transaction price decile.
  - Can estimate value-weighted average error.
  - Can see spread/symmetry of errors.
  - Do not have access to individual-level errors.

# Assumption 1: Zillow's value-weighted errors are small.



- Modest positive bias, particularly during downturn.
- We adjust the AVM averages to reflect this known bias.

## Assumption 2: Average value of the valued and non-valued units are equal.

- Not directly testable with our data.
  - Missing characteristics and/or too few “comparable” sales.
- Benign for sf, where the hit rate is high.
  - Mf hit rate is low, but mf is a small share of the aggregate.
- May hold approximately within counties.
  - 2014 ACS merged at the property-level with a similar AVM.
  - Missing and non-missing properties differ by  $\approx 1\%$  in owner-reported ACS value within a county.
  - Unconditionally, missings are  $\approx 20\%$  less valuable.
- We therefore aggregate from the county level where possible.

## Assumption 3: Average value of the Zillow units equals the average value of the own-use units.

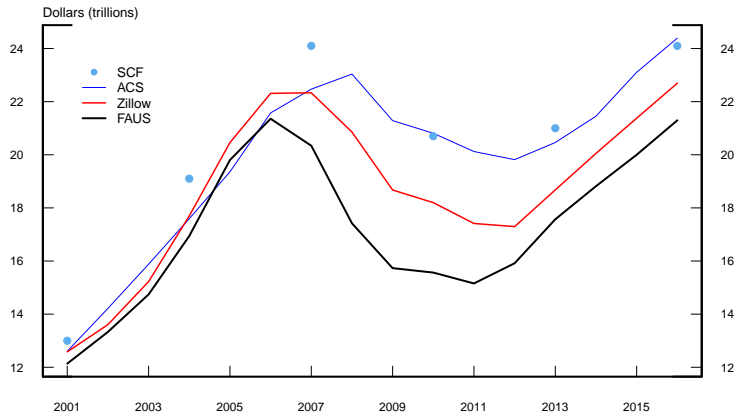
- Zillow averages include rental properties.
- Rentals may be lower quality but in more desirable locations.
  - SF rentals are 30% smaller than owner-occupied units in the AHS.
  - Tract-level rental share is correlated with density within metros in the 2000 Census.
- ACS/AVM property-level merge suggests rentals are 20% (mf) to 35% (sf) less valuable.
- We do not adjust for “rental bias” in our baseline estimates.
  - Adjustment would raise our 2014 aggregate by 6%.
  - Do not have data to make adjustment for other years.

## Assumption 4: The average value for counties not valued by Zillow equals the average value in the state.

- Zillow does not have average values for every county.
- We use the state-level average to impute missing counties.
- Missing counties are less populous, more rural than average.
  - Missing counties likely have lower average values.
- Violations of (4) are likely immaterial because overall coverage is very high.
  - Zillow's counties cover over 95% of housing stock.

# Main Results: Zillow yields different aggregates over the bust and recovery.

## Aggregate Housing Wealth



Source: American Community Survey (U.S. Census Bureau), Financial Accounts of the United States, Survey of Consumer Finance (triennial), and Zillow.

rental

# Main Results: Zillow yields different timing over the bust and recovery.

**Percent Change in Aggregate Housing Wealth**



Source: American Community Survey (U.S. Census Bureau), Financial Accounts of the United States, Survey of Consumer Finance (triennial), and Zillow.



# Main Results: Discussion.

- AVM methods represent a promising way forward.
  - Overcome reporting biases associated with surveys.
  - Overcome representativeness biases associated with repeat-sales.
- Zillow's AVM suggests a different path for housing wealth over the Great Recession.
  - Shallower than repeat-sales, sharper than self-reports.
  - Self reports lag other measures substantially.
  - All three show similar growth since 2012.
- These differences are substantial.
  - Zillow is about \$3.5 trillion higher than FAUS in 2008.
  - Zillow is over \$2.7 trillion below the ACS in 2011.

# Future Work.

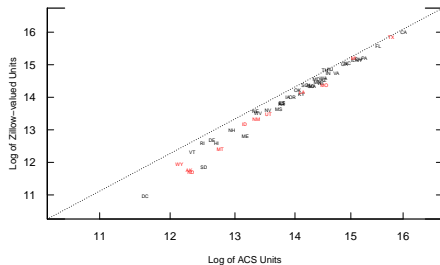
- Zillow makes geographically disaggregated wealth estimates possible.
  - Can disaggregate by state, county, or MSA.
  - Small/rural states have more uncertain estimates.
  - Difficult or impossible to do using price indexes + fixed investment.
- Assess variation in “regional” LTV and DTI.
  - Plausible estimates back to 2000, earlier than other sources.
  - Cannot create property-level estimates (currently).
- Revisit various macroeconomic studies of the Great Recession.
  - Consumption, monetary policy transmission, etc.
- Understand differences between self-reports, price indexes, and AVMs in greater depth.
  - Investigate regions where methods disagree most sharply.
  - Property-level comparisons.

# The end

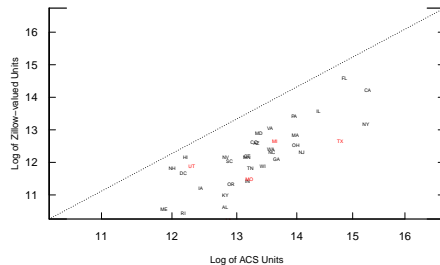
**Thank you! Questions?**

# Zillow universe is different from national universe.

Single Family, ACS Total Homes v. Zillow



Multifamily, ACS Total Homes v. Zillow



[back](#)

# Split of housing units in the ACS.

- Need total units intended for owner-use by structure type.
- Identifying such units is not trivial: must decide which vacant homes are intended for owner-use.

UNITS IN STRUCTURE	2014 ACS			
	% Own Use	% Rental Use	% Other Vacant	Total Homes (Millions)
Total housing units	60	34	5	134.0
1-unit, detached	79	16	5	82.4
1-unit, attached	59	36	4	7.7
2 units	21	71	8	5.1
3 or 4 units	14	80	6	6.0
5 to 9 units	12	83	5	6.5
10 to 19 units	11	85	5	6.1
20 or more units	17	79	4	11.6
Mobile home	66	25	9	8.5
Boat, RV, van, etc.	60	40	0	0.1

## 1. "Own Use" Units

- Unit is occupied and is owned outright or being bought with a mortgage.
- Unit is vacant and is for sale only.
- Unit is vacant and is for seasonal or recreational use.

## 2. "Rental Use" Units

- Unit is occupied and rented.
- Unit is vacant and for migrant farm work.
- Unit is vacant and for rent only.

## 3. "Other Vacant" Units

- Rented or sold, but not occupied (no way to determine which it is).
- Other vacant.

NB: "Other Vacant" units are assigned to owner-occupied use according to share:

$$share = \frac{\text{Own Use Units}}{(\text{Own Use Units} + \text{Rented Units})}$$

# Mapping ACS housing units to Zillow data.

- Zillow provides avg. prices for two broad property classes: single-family and condo/coop.
- Map these property classes to ACS data using the ACS “structure type” indicator.

ACS Structure Type	Valued Using Zillow Property Category	Land Use Codes Underlying Zillow Estimates
1-unit, detached	"Single-family"	Single-family residence, townhouse/rowhouse, bungalow, patio home, zero lot line, cluster home, miscellaneous residential
1-unit, attached		
2 units	"Condo/coop"	Coop, condo
3 or 4 units		
5 to 9 units		
10 to 19 units		
20 or more units		
Mobile home	Not Available	
Boat, RV, van, etc.		

[back](#)

# ACS/Zillow counts by property type.

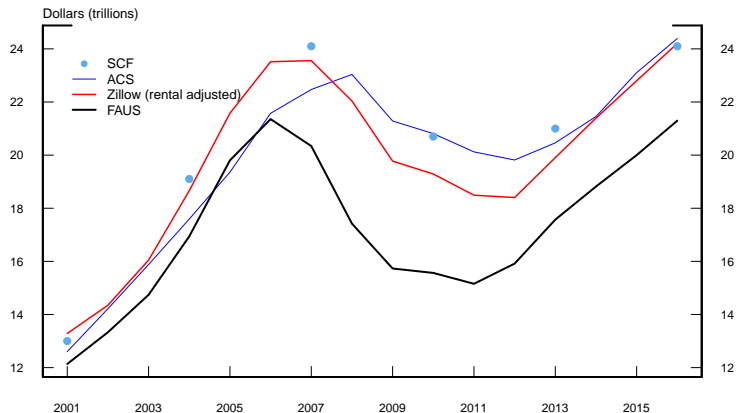
**Table 1**  
**Property Counts in 2015 (millions)**

	Owner-Occupied Units in ACS	Units with a Zillow AVM	Total Units in ACS
Single Family	70.6	79.4	90.6
Multifamily	5.4	9.8	35.6
Total	75.9	89.3	126.2

[back](#)

# Main Results: Rental Adjusted Aggregates.

## Aggregate Housing Wealth



Source: American Community Survey (U.S. Census Bureau), Financial Accounts of the United States, Survey of Consumer Finance (triennial), and Zillow.

[back](#)