An Agent-Based Model of the Housing Market

Rob Axtell (George Mason University)
Ernesto Carrella (George Mason University)
Ben Conlee (Ellington Capital Management)
Doyne Farmer (Oxford University)
John Geanakoplos (Yale University)
Jon Goldstein (George Mason University)
Matthew Hendrey (George Mason University)
Peter Howitt (Brown University)
Philip Kalikman (Yale University)
David Masad (George Mason University)
Nathan Palmer (George Mason University)
Chun-Yi Yang (George Mason University)

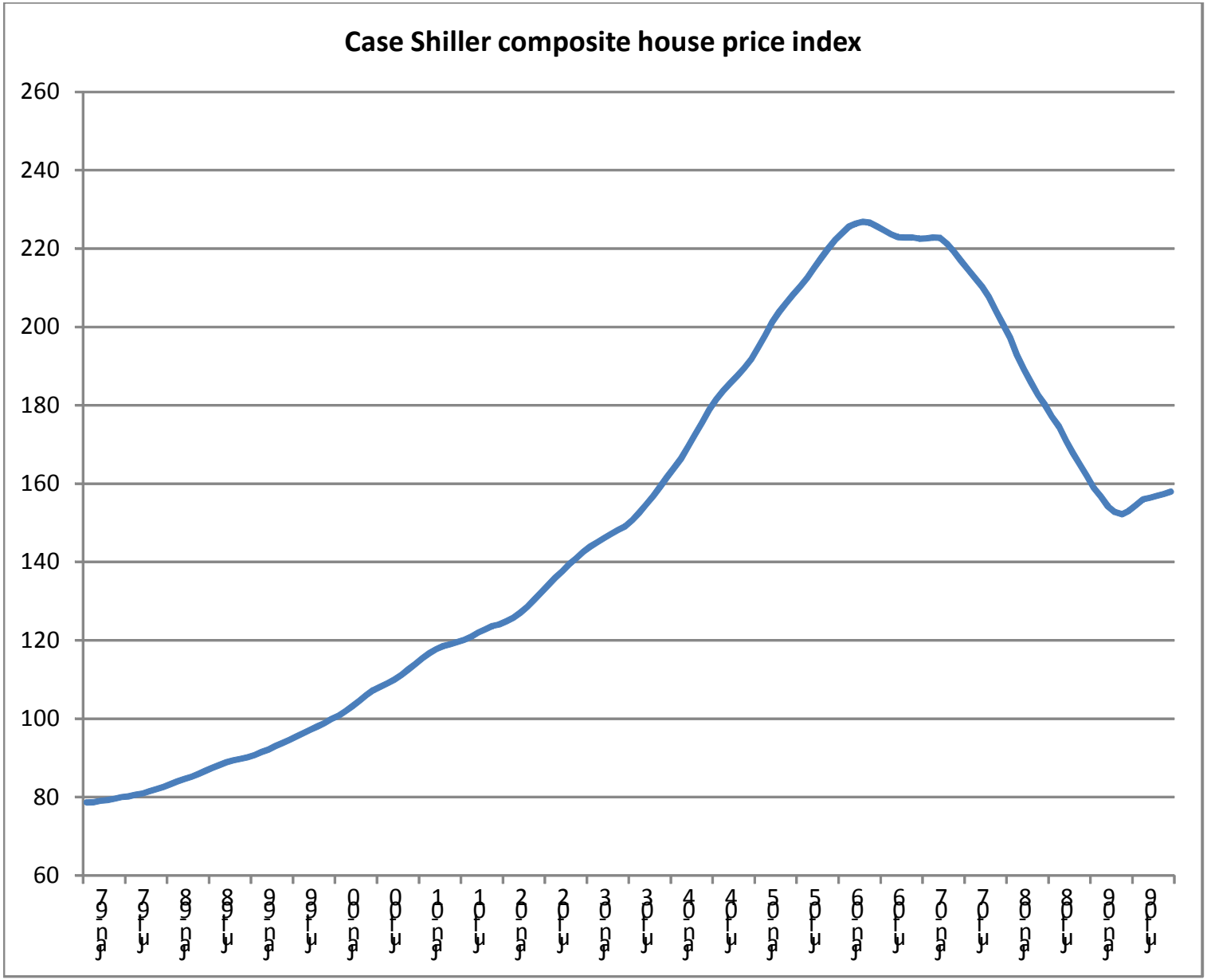
MFM meeting, Chicago, May 2, 2013

Introduction

The housing market in macro modelling

Central to the financial crisis

US house prices



Introduction (cont'd)

Central question: relative importance of interest rates and leverage in the boom and crash

Introduction (cont'd)

Central question: relative importance of interest rates and leverage in the boom and crash

Why an agent-based approach?

- heterogeneity and complexity of decisions
- rich data available
- many aspects of the market to be accounted for
- has been useful on Wall Street

Introduction (cont'd)

Central question: relative importance of interest rates and leverage in the boom and crash

Why an agent-based approach?

- heterogeneity and complexity of decisions
- rich data available
- many aspects of the market to be accounted for
- has been useful on Wall Street

Comparison to meteorological modelling

Methodology

Focus on household behavior

- take banking behavior, income, demographics as given

Methodology

Focus on household behavior

- take banking behavior, income, demographics as given

Calibrate the component modules independently and then put together without fitting to target data

Methodology

Focus on household behavior

- take banking behavior, income, demographics as given

Calibrate the component modules independently and then put together without fitting to target data

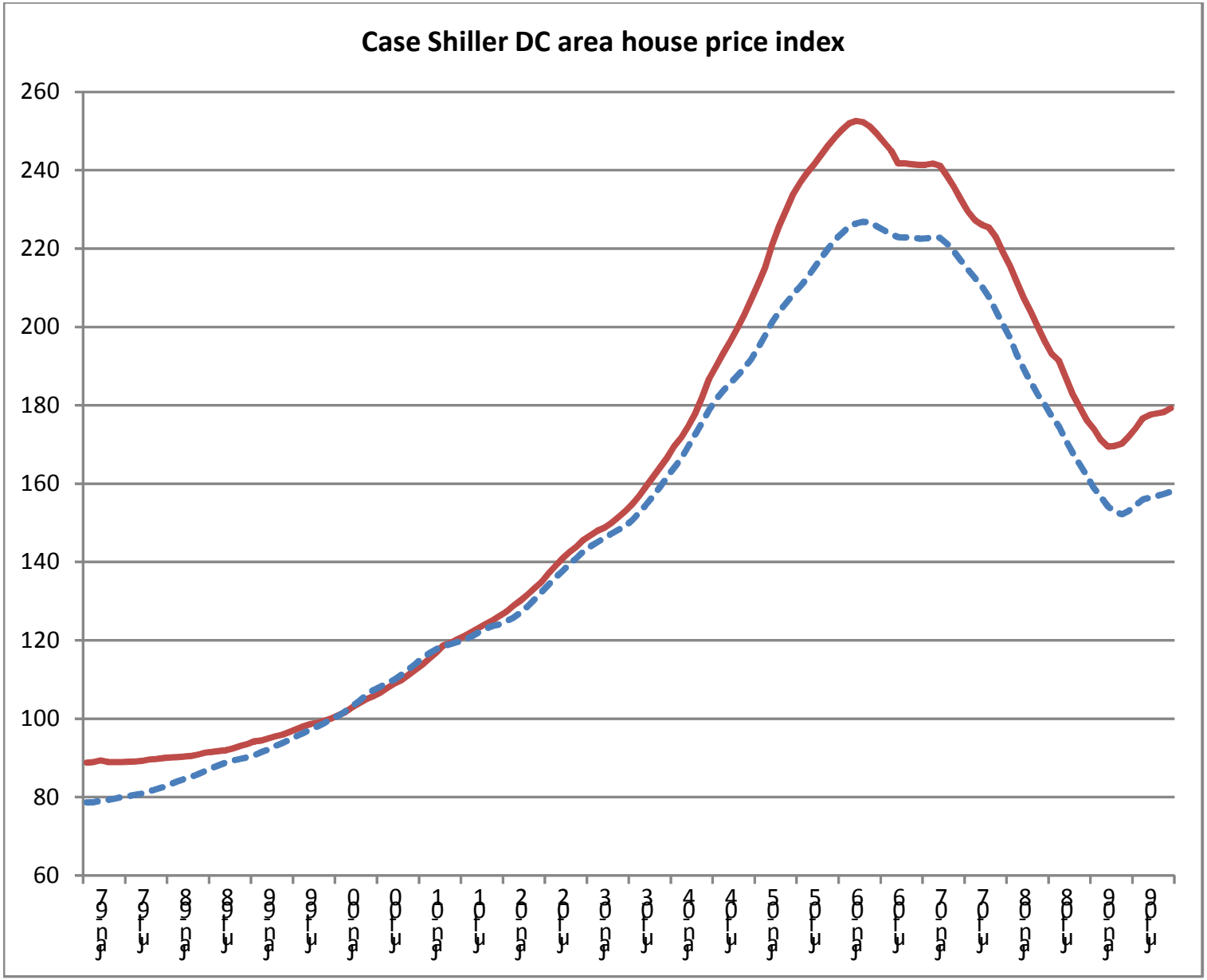
Initialization period of 100 years to get endogenous correlations

Data

Focus on Washington DC area, 1997-2009

Approximately 2.2 million households

DC are house prices



Data (cont'd)

Main Sources:

- Local
  - Core Logic - all public record data, over 3 million mortgages (including "hidden") and other housing variables
  - MLS (listings, price changes, delistings, sales)
  - IRS (income)
  - Loan Performance (more housing variables related to 885,000 of the mortgages)
  - Census Bureau (housing stock)
- National
  - PSID and CEX (national wealth, housing costs)
  - ACS (rental market, housing costs)

Outline of the model

Main objects

- Households (10:1 scale)
  - income, wealth, housing status, mortgage, initialized to data
- Houses of differing quality
  - qualities drawn from distribution of most recent sale prices relative to DC Case-Shiller
  - initial number of houses given by census data
- Mortgages of three kinds
  - interest only, ARM, conventional fixed
- Single bank
  - approves and makes mortgage loans
  - initiates foreclosure on all loans more than 2 months delinquent
  - attempts to sell foreclosed houses

Household actions

Each period (month) each household:

- receives income
- spends on non-housing consumption
- if holding a mortgage, decides whether to strategically default
- if holding a mortgage, decides whether to refinance
- pays housing cost (rental, or maintenance, tax, insurance and mortgage) if wealthy enough

Household actions (cont'd)

- decides whether to attempt buying a house
- if buying, chooses a desired expenditure and leverage
- if living in own house, possibly invests in a rental property
- if living in own house decides whether to list it and what price
- if already listing, decides whether to delist or reduce price
- if owning a vacant rental unit decides on rent

Household income

Carroll (BPEA, 1992) estimated the process from PSID data:

$$\ln Y = \ln Y_p + \ln Y_t$$

where:

$\ln Y_p$ is a random walk with drift (2% pa) and normal increments, and $\ln Y_t$ is the product of an iid normal variable and a (0,1) Bernoulli shock with $1 - p$ proportional to the unemployment rate

Household income

Carroll (BPEA, 1992) estimated the process from PSID data:

$$\ln Y = \ln Y_p + \ln Y_t$$

where:

$\ln Y_p$ is a random walk with drift (2% pa) and normal increments, and $\ln Y_t$ is the product of an iid normal variable and a (0,1) Bernoulli shock with $1 - p$ proportional to the unemployment rate

After this process determines the rank order of each household's income, the distribution is clamped to IRS data. (Each hh gets the actual income of its percentile.

Desired Expenditure and Leverage

- Desired expenditure originally set to make housing cost equal one third of income, with individual shocks

Desired Expenditure and Leverage

- Desired expenditure originally set to make housing cost equal one third of income, with individual shocks
- Now it is a concave function of Y:

$$P^* = \frac{\varepsilon h Y^g}{\tau + c + LTV \cdot i - a \cdot HPA}$$

τ	tax and insurance per dollar house price
c	maintenance per dollar house price
i	prime rate
HPA	last 12 months' avg % house price appreciation
ε	lognormal individual shock

(parameters $h, g, \tau, c, a, \sigma_\varepsilon$ estimated from PSID and ACS)

Desired Expenditure and Leverage

- Desired expenditure originally set to make housing cost equal one third of income, with individual shocks
- Now it is a concave function of Y:

$$P^* = \frac{\varepsilon h Y^g}{\tau + c + LTV \cdot i - a \cdot HPA}$$

τ	tax and insurance per dollar house price
c	maintenance per dollar house price
i	prime rate
HPA	last 12 months' avg % house price appreciation
ε	lognormal individual shock

(parameters $h, g, \tau, c, a, \sigma_\varepsilon$ estimated from PSID and ACS)

- Desired leverage drawn from bin distribution conditional on desired expenditure size using CoreLogic data.

Loan approval process

- Household must satisfy 3 constraints:
  1. LTV
  2. DTI
  3. Wealth
- Bank initially assigns a loan chosen randomly from empirical distribution of type, rate, LTV conditional on expenditure.
- If this doesn't fit, the loan size is adjusted to satisfy constraints.
- If the fit requires a loan greater than our "estimated" max LTV, no loan is approved

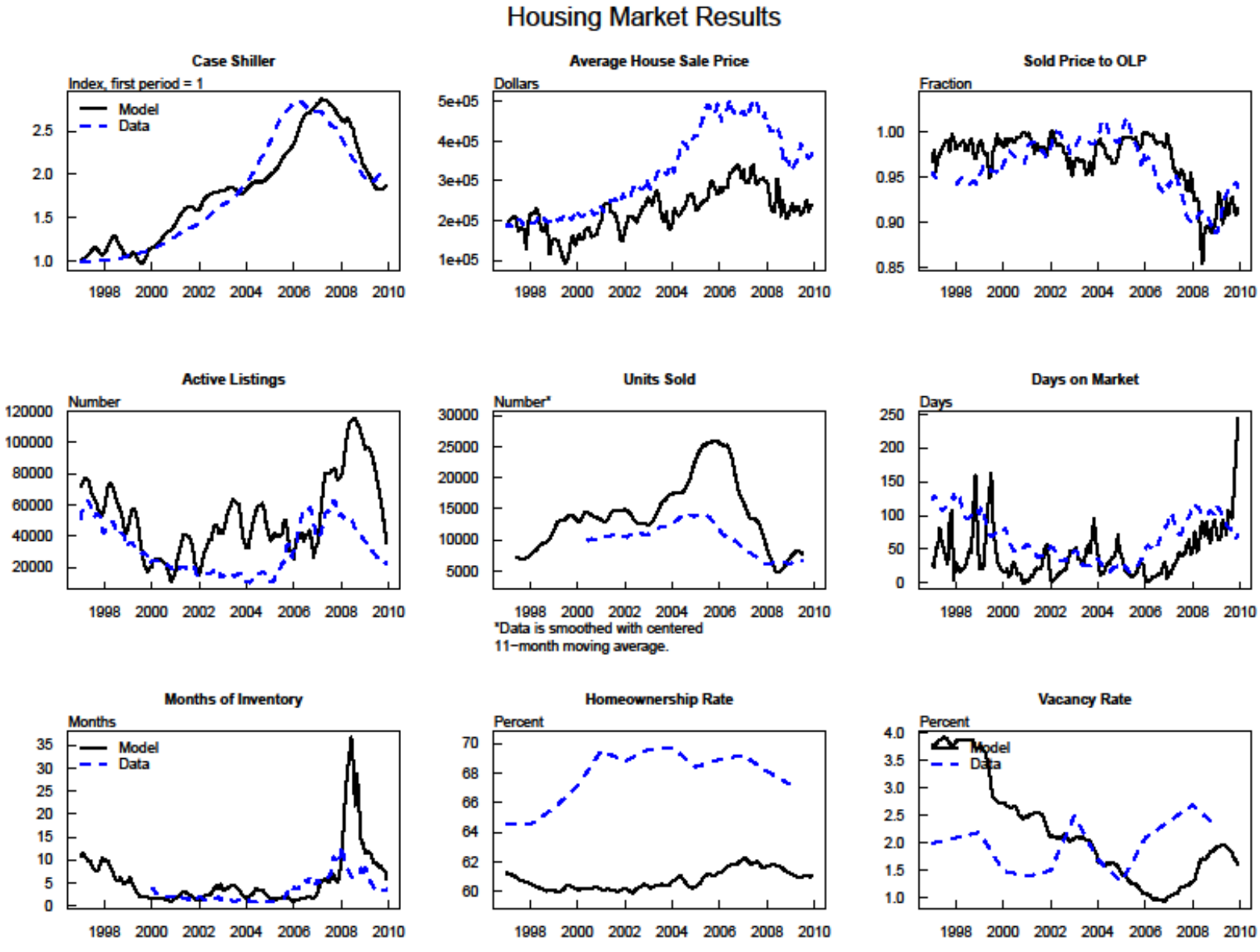
The matching process

- Approved buyers are then selected in random order and matched with the highest quality house on the market at no more than desired expenditure. (List price is the sale price.)
- Before agreeing, the buyer calculates the relative advantage of renting a unit of similar quality:

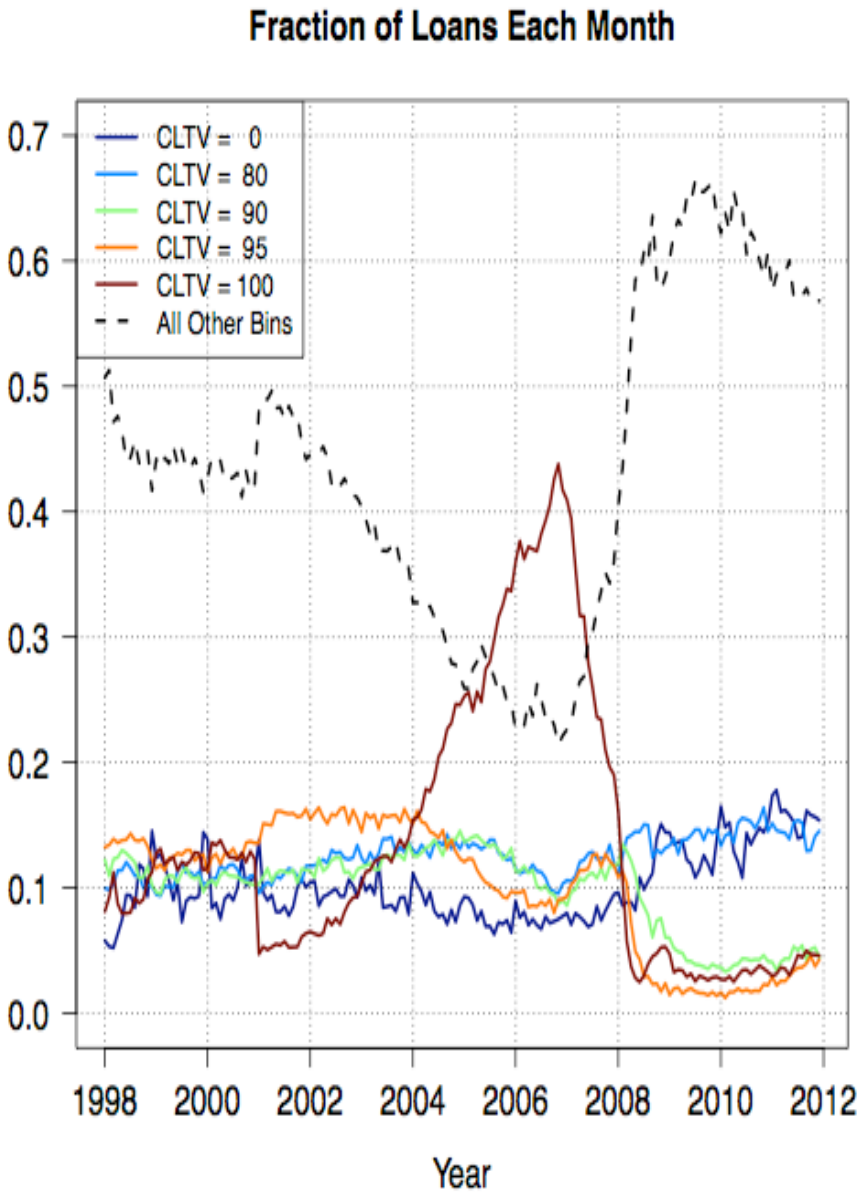
$$RA = P \cdot (\tau + c + LTV \cdot i - a \cdot HPA) - r$$

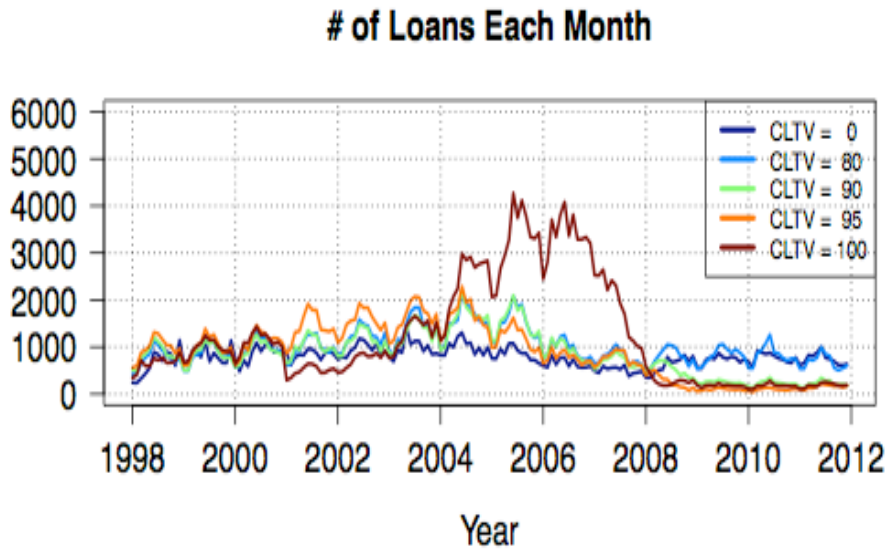
  and rents instead with probability fitted as logistic of RA
- Unsuccessful buyers always choose to rent

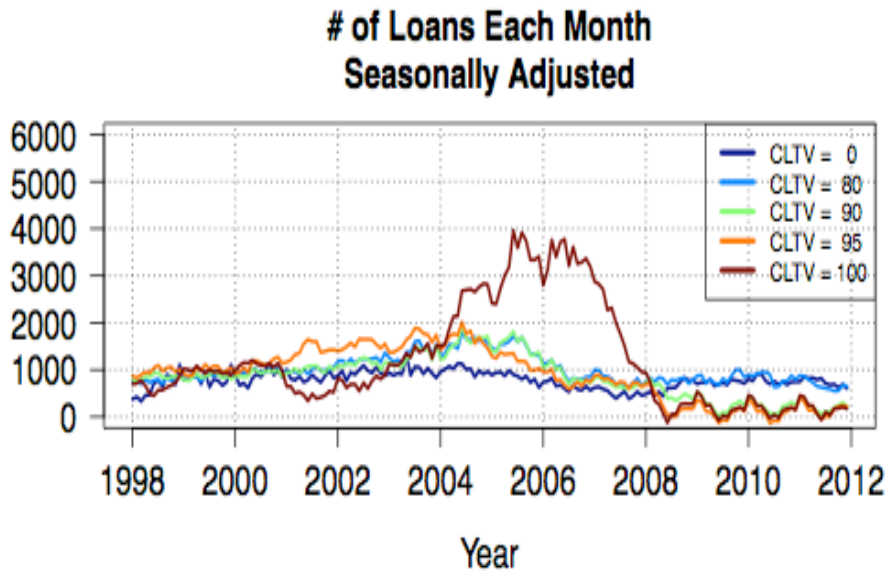
Baseline results



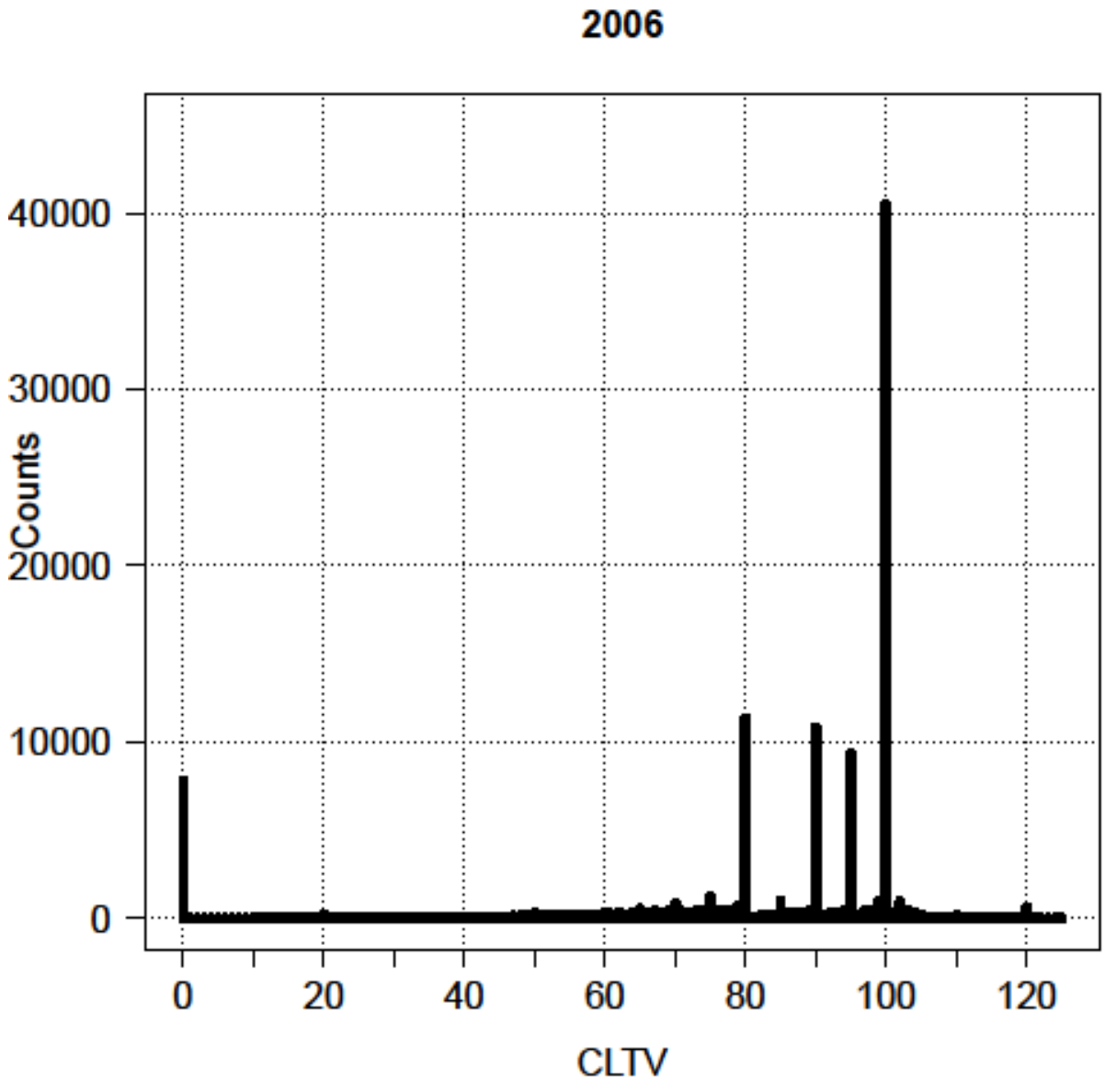
Changing leverage constraints



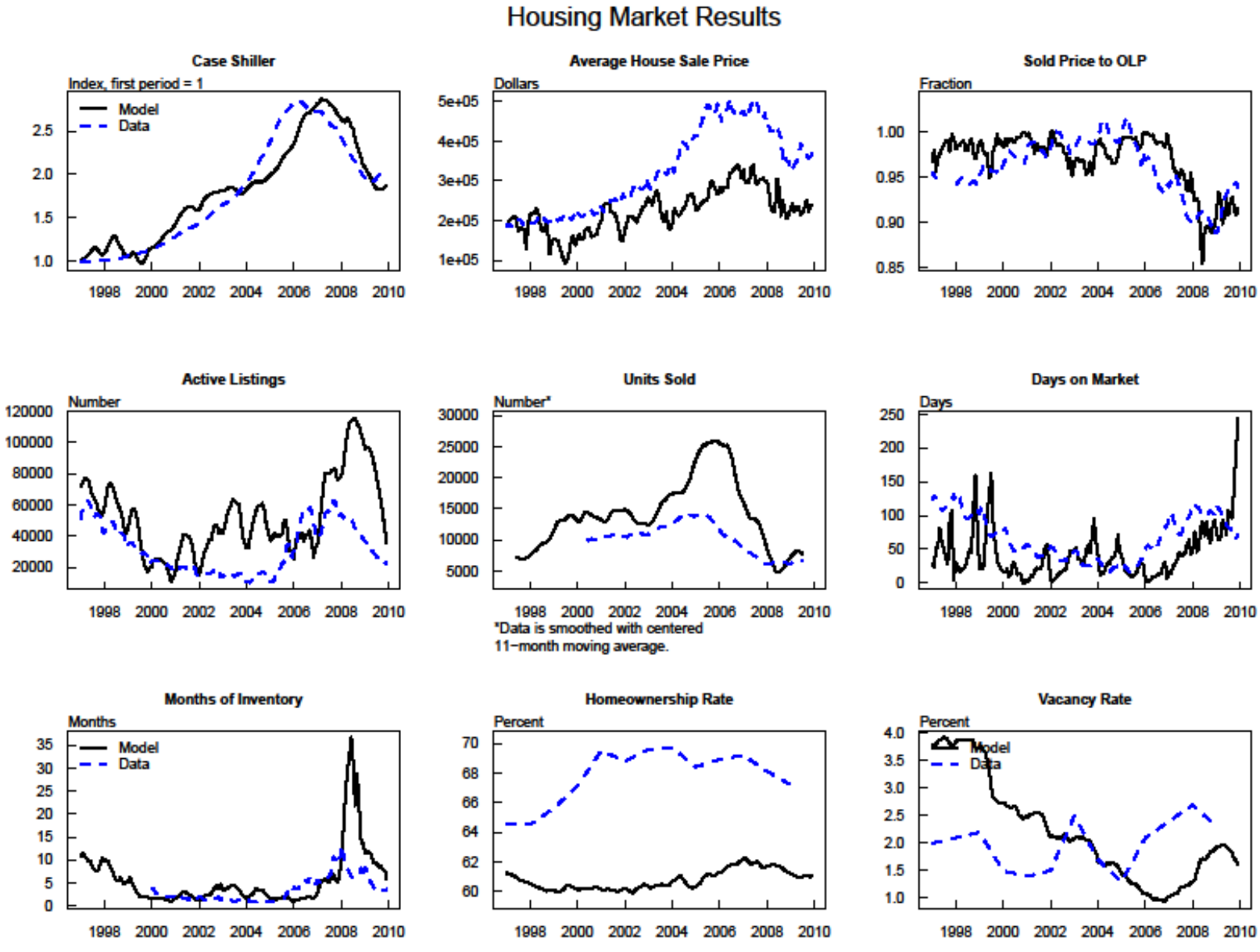




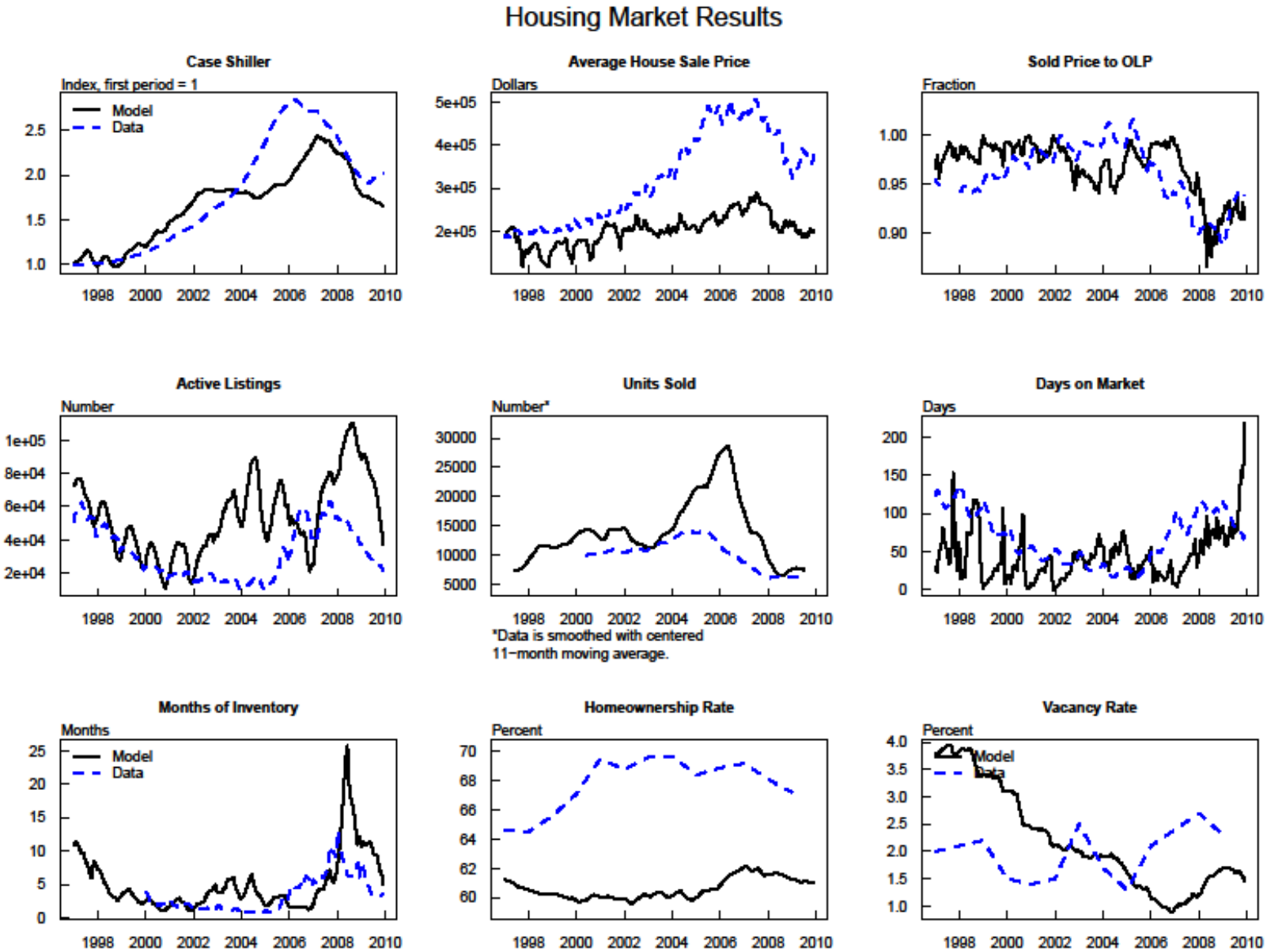
Changing leverage constraints (cont'd)



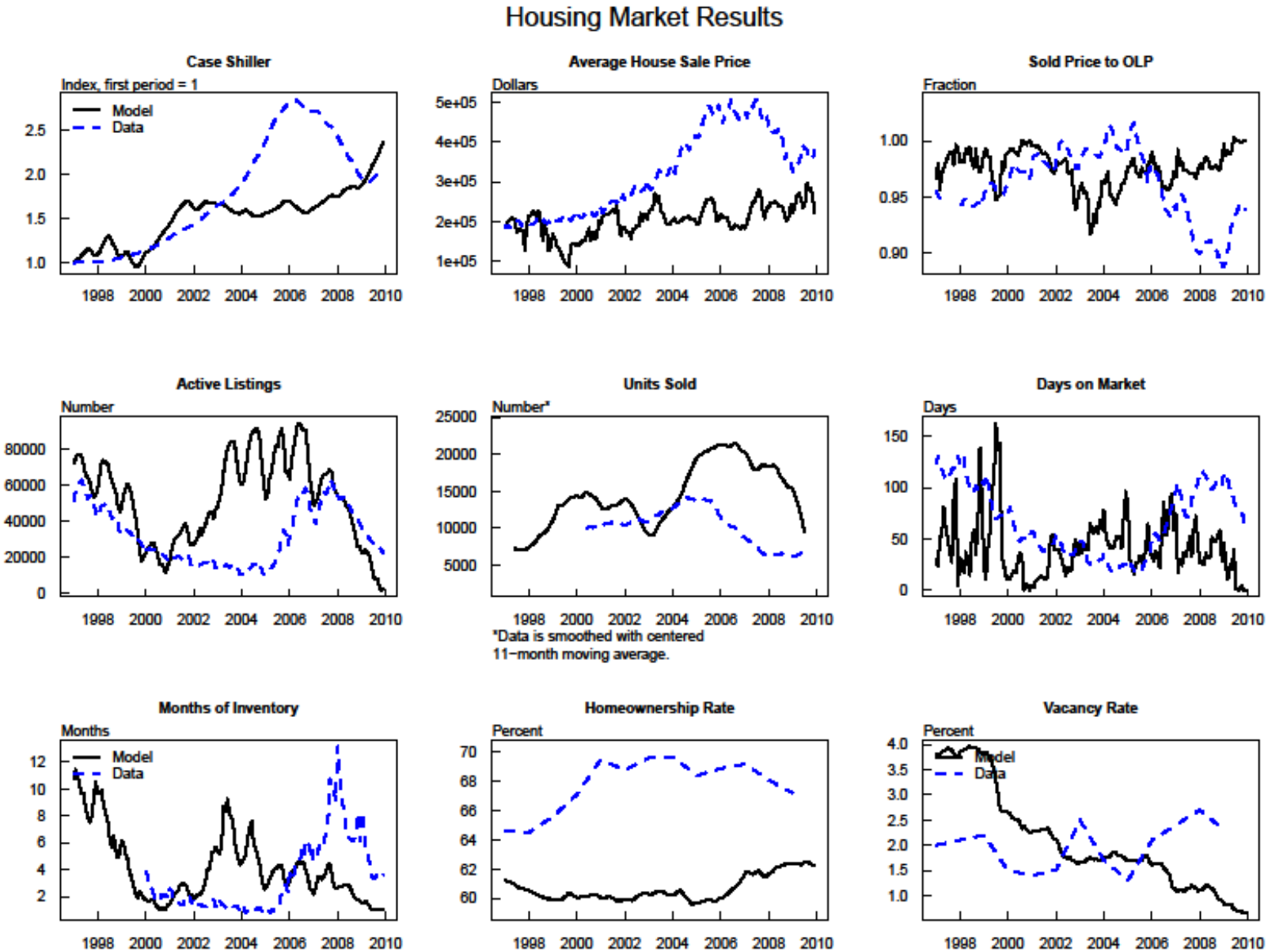
Baseline results



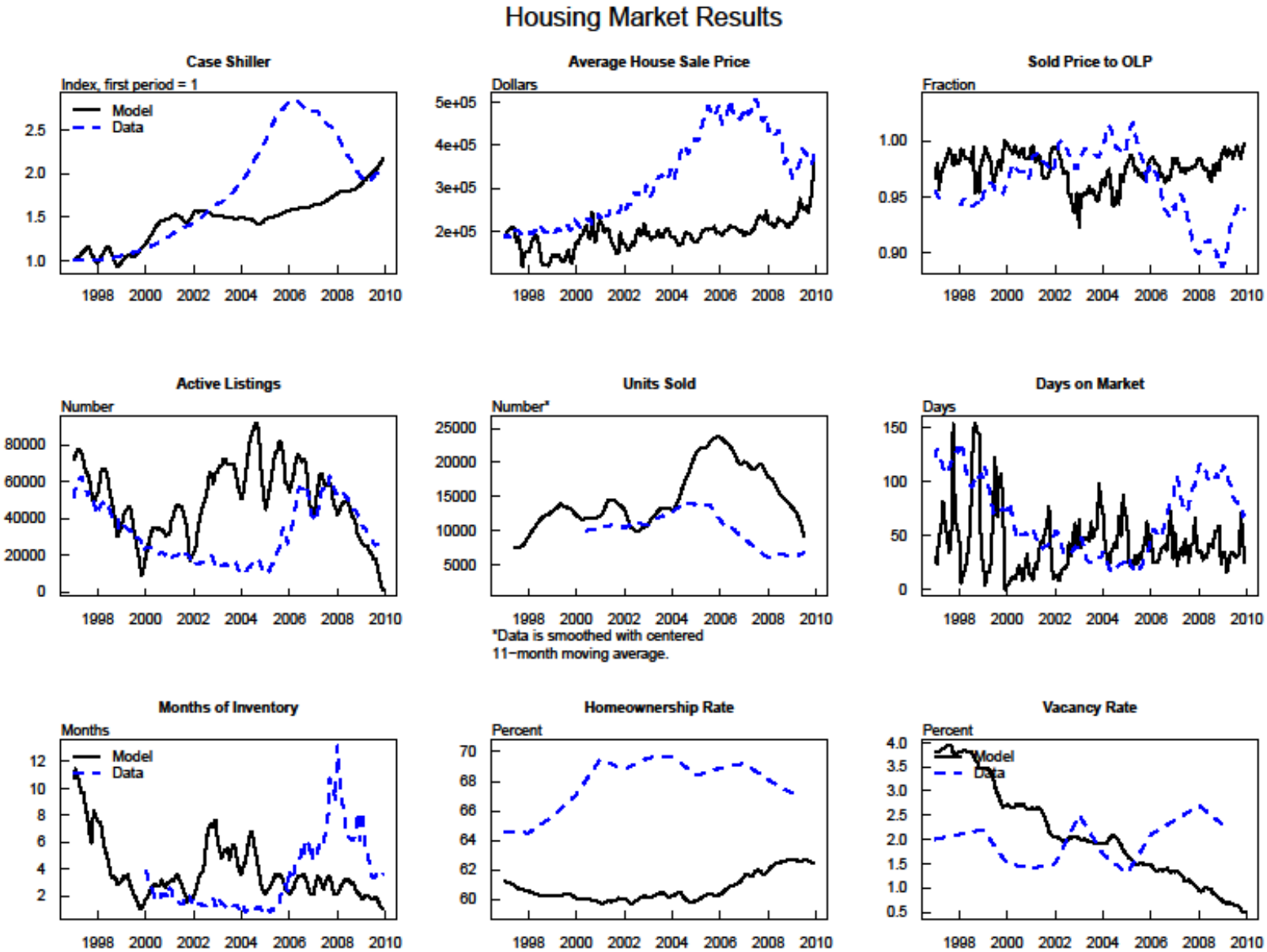
Interest rates fixed



Leverage fixed



Interest rates and leverage fixed



The next steps

- Work on homeownership rates

The next steps

- Work on homeownership rates
- Validate model on other cities

The next steps

- Work on homeownership rates
- Validate model on other cities
- Aggregate across cities

The next steps

- Work on homeownership rates
- Validate model on other cities
- Aggregate across cities
- Incorporate into an agent-based macro model

Other modules

The rest of these slides describe the other modules of the model. They are mainly for the discussant's benefit. I will almost certainly not have time to present them, but will be able to display them if needed during the discussion.

Non-housing consumption

$$\log C = 1.662398 + 0.590391 \cdot \log Y$$

Estimated from CEX data.

Strategic default

The monthly probability of strategic default is

$$P = max\{CCLTV - .5, 0\} \cdot (0.0087 - 0.0235 \cdot HPA)$$

where the parameters are estimated using Current-to-DQ loan probabilities over the period 1997–2010.

Strategic default

The monthly probability of strategic default is

$$P = max\{CCLTV - .5, 0\} \cdot (0.0087 - 0.0235 \cdot HPA)$$

where the parameters are estimated using Current-to-DQ loan probabilities over the period 1997–2010.

Defaulter does not qualify for another loan for 84 months.

Refinancing

1. Rate refinance (same size loan):
   - probability depends on moneyness (prime rate now vs at origination)
   - equals the mean refi rate among actual loans of same age in public records
   - bank offers a rate chosen from empirical distribution in public records
   - hh will accept iff lower than existing rate

Refinancing

1. Rate refinance (same size loan):
   - probability depends on moneyness (prime rate now vs at origination)
   - equals the mean refi rate among actual loans of same age in public records
   - bank offers a rate chosen from empirical distribution in public records
   - hh will accept iff lower than existing rate
2. Cashout refi (new LTV $= 0.85$)
   - probability $= .009 \cdot \left(1 - \frac{CLTV}{0.85}\right)$
   - Khandani-Lo-Merton (2009)

Becoming a buyer

Renters always try to buy when their 12-month lease expires

Successful sellers always try to buy

Becoming an investor

only if income $>$ \$100K, and with probability a logit of estimated yield:

$$\frac{1}{1 + e^{-p \cdot y + q}}$$

where y is yield:

$$y = r + a \cdot HPA - \tau - c - i \cdot LTV$$

and $p = 24, q = 4.5$ chosen arbitrarily to get more than 10% investing each month when $y = 10\%$

Listing

probability of listing clamped to MLS data
list price determined by equation estimated using same data:

$$OLP = 0.99 \cdot \varepsilon \cdot \overline{P} \cdot e^{0.22 + 0.22 \ln s - 0.011 \ln DOM}$$

where:

OLP	original list price
$\overline{P}$	avg price of "comparable" houses recently sold
s	recent avg sold/OLP ratio
DOM	recent avg days on market
ε	non-Gaussian shock drawn from regression residuals

Delisting

probability clamped to bin distribution conditional on days on market (MLS data)

Price reductions

probability and size drawn from binned distribution of markdowns depending on DOM and recent reductions

Rental rate

Equation estimated from ACS data:

$$r_{t+1} = r_t \cdot \left(1 - 0.8355 \cdot \left(\frac{v_t - v^*}{12}\right)\right)$$

where:

r	rent/quality
v	avg apartment vacancy rate (endogenous)
v^*	"target" vacancy rate set at 6% (avg for DC area)