A Direct Estimate of Rule-of-Thumb Consumption using the Method of Simulated Quantiles and Cross Validation

Nathan M. Palmer

Federal Reserve Board of Governors

September 10, 2019

¹Views expressed in this presentation are those of the speaker and not necessarily of the Federal Reserve Board or System.
Section 1

Overview
Immediate Motivation

- A feature of household wealth may be consistent with learning-to-optimize over the life cycle
  - want to select between optimizing & learning
  - spendthrift behavior as rough first pass
Immediate Motivation

- A feature of household wealth may be consistent with learning-to-optimize over the life cycle
  - want to select between optimizing & learning
  - spendthrift behavior as rough first pass

- Take a “textbook” model of household life cycle choice, and:
  1. Extend estimation to account for data feature (Method of Simulated Quantiles)
  2. Extend model to allow for generalization of Cambell and Mankiw’s (1989, 1990) “rule of thumb” consumers (consume all income)
  3. Select between models of “rule of thumb” consumption using k-fold cross validation

Results are preliminary; feedback extremely welcome
Immediate Motivation

- A feature of household wealth may be consistent with learning-to-optimize over the life cycle
  - want to select between optimizing & learning
  - spendthrift behavior as rough first pass

- Take a “textbook” model of household life cycle choice, and:
  1. Extend estimation to account for data feature (Method of Simulated Quantiles)
  2. Extend model to allow for generalization of Cambell and Mankiw’s (1989, 1990) “rule of thumb” consumers (consume all income)
  3. Select between models of “rule of thumb” consumption using k-fold cross validation

- Results are preliminary; feedback extremely welcome
Long-term Motivation

- Want to build large agent-based / HA models
  - Geanakoplos et al. (2012) [Details]
  - ECB and BOE models
  - FRB-US*
Long-term Motivation

- Want to build large agent-based / HA models
  - Geanakoplos et al. (2012)
  - ECB and BOE models
  - FRB-US*

- Rational expectations solution is intractable.

- Two problems:
  1. Lucas critique: “How to model smart, purposeful agents who dynamically model the world in their heads?”
  2. Sims critique: “There are infinitely many ways to be non-optimal, how do you choose the right one?”
Long-term Motivation

- Lucas critique solution: learning to optimize throughout the life cycle (not this paper) [Reinforcement learning details]
Long-term Motivation

- Lucas critique solution: learning to optimize throughout the life cycle (not this paper)

- Sims critique solution: formal selection between models estimated on partial-equilibrium problems (this paper)
Long-term Motivation

- Lucas critique solution: learning to optimize throughout the life cycle (not this paper)

- Sims critique solution: formal selection between models estimated on partial-equilibrium problems (this paper)

- Goal: demonstrate a generally applicable approach to behavior selection for models:
  - with “hard” / no likelihood surface
  - that are non-nested
  - with limited / not ideal data
This Paper

- Estimate a “textbook” model of household lifecycle consumption behavior using the method of simulated quantiles
This Paper

- Estimate a “textbook” model of household lifecycle consumption behavior using the method of simulated quantiles

- Also estimate $N$ models with variations of Campbell and Mankiw’s “rule of thumb:” consume all income
This Paper

- Estimate a “textbook” model of household lifecycle consumption behavior using the method of simulated quantiles
- Also estimate $N$ models with variations of Campbell and Mankiw's “rule of thumb:” consume all income
- Select between the above models using k-fold cross validation
This Paper

- Estimate a “textbook” model of household lifecycle consumption behavior using the method of simulated quantiles

- Also estimate $N$ models with variations of Campbell and Mankiw’s “rule of thumb:” consume all income

- Select between the above models using k-fold cross validation

- Spoiler: in line with Campbell, Mankiw’s (1989, 1990) 50% result
  - Preliminary results: 26% of consumers follow spendthrift-like rule
Related Literature

1. Rules of thumb for dynamic macro problems:
   - Campbell & Mankiw (1989, 1990) and cottage literature
   - Ganong & Noel HARK, AER (2019)
Related Literature

1. Rules of thumb for dynamic macro problems:
   - Campbell & Mankiw (1989, 1990) and cottage literature
   - Ganong & Noel HARK, AER (2019)

2. Reinforcement Learning / Approximate DP for macro
   - Allen & Carroll (2001) and cottage literature
   - Hull JEDC (2015), Sinitskaya and Tesfatsion JEDC (2014)
   - Calvano et al. (2018), “Q-Learning to Cooperate” NBER Econ of AI
   - Fernández-Villaverde et al. (2019), “Fin Frictions & Wealth Distrib”
   - RL as solution method: Duarte, Winant, Maliars, Scheidegger
Related Literature

1. Rules of thumb for dynamic macro problems:
   - Campbell & Mankiw (1989, 1990) and cottage literature
   - Ganong & Noel HARK, AER (2019)

2. Reinforcement Learning / Approximate DP for macro
   - Allen & Carroll (2001) and cottage literature
   - Hull JEDC (2015), Sinitskaya and Tesfatsion JEDC (2014)
   - Calvano et al. (2018), “Q-Learning to Cooperate” NBER Econ of AI
   - Fernández-Villaverde et al. (2019), “Fin Frictions & Wealth Distrib”
   - RL as solution method: Duarte, Winant, Maliars, Scheidegger

3. Estimation and model selection broadly:
   - Non-nested models with a likelihood surface: Vuong (1989)
   - Structural estimation and selection: Li (2009)
Outline

1. Overview
2. Agent Problem
3. Method of Simulated Quantiles
4. K-Fold Cross Validation
5. Preliminary Results, Conclusion
6. Appendix
Section 2

Agent Problem
“Textbook” Household Problem

A household solves the $T$-horizon consumption problem:

$$\max_{c_t} \mathbb{E}_t \left[ \sum_{t=0}^{T} \beta \mathcal{D}_t u(c_t) \right]$$

s.t.

$$m_{t+1} = R_{t+1}(m_t - c_t) + \xi_{t+1}$$

$m_0$ given

where

- $\rho, \beta$ are risk aversion and discount factor
- $m_t$ is total “cash on hand”
- $b_{t+1} \equiv R_{t+1}(m_t - c_t)$ is net worth
- $R_t$ is risk-free return on assets, $\mathcal{D}_t$ is age-dependent survival prob
- $\xi_t$ are mean-1 temporary shocks to income
Solution Method

A household solves the $T$-horizon consumption problem:

$$\max_{c_t} \mathbb{E}_t \left[ \sum_{t=0}^{T} \beta \mathcal{D}_t u(c_t) \right]$$

s.t.

$$m_{t+1} = R_{t+1}(m_t - c_t) + \xi_{t+1}$$

$m_0$ given

Solution method:

- Solution is set of optimal consumption functions $\{c^*_t\}_{t=1}^{T}$
- In the next-to-last period, one-period-ahead $v^*_T$ and $c^*_T$ known
- Numerically solve for $v^*_{T-1}$ and $c^*_{T-1}$ and recurse backwards
- Entire problem has been normalized by permanent income, not shown
Household Solution: Consumption Functions

Consumption Functions: Black Before Retirement, Red After

$\beta = 1.007$ and $\rho = 4.4$

Recall: learning vs optimization
Going from Consumption Functions to Wealth

- Choose $\beta$, $\rho$ and solve problem for consumption functions $\{c^*_t\}_{t=1}^T$
- Simulate large panel of artificial wealth data under $\beta$, $\rho$
Going from Consumption Functions to Wealth

- Choose $\beta$, $\rho$ and solve problem for consumption functions $\{c_t^*\}_{t=1}^T$
- Simulate large panel of artificial wealth data under $\beta$, $\rho$
- Pool simulated data to look like Survey of Consumer Finance (SCF) wealth data, measure distance via quantile functions
  - because we know wealthy tails are fit poorly by this model
- Cross-sectional SCF data are pooled by age in 5-year windows:
Empirical Wealth Distribution: College Educated, Married

(Net Worth)/(Perm Income) by Age Cohort

More education levels
Empirical Wealth Distribution: College Educated, Married

(Net Worth)/(Perm Income) by Age Cohort

More education levels
Section 3

Method of Simulated Quantiles
Method of Simulated Quantiles

- Construct functions of quantiles for all 7 age groups $\tau$:

$$\varphi_{\tau} = \begin{pmatrix}
q_{\tau,50} \\
q_{\tau,75} - q_{\tau,25} \\
(q_{95,\tau} - q_{50,\tau}) - (q_{50,\tau} - q_{5,\tau}) \\
q_{95,\tau} - q_{5,\tau} \\
(q_{95,\tau} - q_{5,\tau})/q_{75,\tau} - q_{25,\tau}
\end{pmatrix}$$

- ...measure location, dispersion, skew, & tail thickness, respectively
Method of Simulated Quantiles

- Construct functions of quantiles for all 7 age groups \( \tau \):

\[
\varphi_\tau = \begin{pmatrix}
q_{\tau,50} \\
q_{\tau,75} - q_{\tau,25} \\
\frac{(q_{95,\tau} - q_{50,\tau}) - (q_{50,\tau} - q_{5,\tau})}{q_{95,\tau} - q_{5,\tau}} \\
\frac{q_{95,\tau} - q_{5,\tau}}{q_{75,\tau} - q_{25,\tau}}
\end{pmatrix}
\]

- ...measure location, dispersion, skew, & tail thickness, respectively

- I use only location and dispersion due to sample size
Define for both the empirical data and simulated data:

- empirical: \( \tilde{\varphi} \)
- simulated: \( \varphi(\beta, \rho) \)

Define the loss function:

\[
\varpi(\beta, \rho) = \left( \tilde{\varphi} - \varphi(\beta, \rho) \right) W \left( \tilde{\varphi} - \varphi(\beta, \rho) \right)
\]

where \( W \) is a positive definite matrix of weights.
Numerically minimize the loss function to estimate \((\beta, \rho)\):

\[
\hat{\beta}, \hat{\rho} = \arg\max_{(\beta, \rho)} \left( \bar{\phi} - \phi(\beta, \rho) \right) W \left( \bar{\phi} - \phi(\beta, \rho) \right)
\]

Quantile choice: as with all SMM, should be motivated by features of model and data.
Estimated Wealth Distribution: Median & IQR

(Net Worth)/(Perm Income) by Age Cohort

- Age: 1, IQR: 0.9
- Age: 2, IQR: 1.1
- Age: 3, IQR: 1.4
- Age: 4, IQR: 1.8
- Age: 5, IQR: 2.1
- Age: 6, IQR: 2.4
- Age: 7, IQR: 2.8
Estimated Wealth Distribution: Median Only

(Net Worth)/(Perm Income) by Age Cohort
Motivation from the Data

Dispersion “miss” remains even when included in estimation objective.

- Learning might explain dispersion
Motivation from the Data

Dispersion “miss” remains even when included in estimation objective.

- Learning might explain dispersion
- Examine simpler case first: “fraction spendthrift”
  - fraction $a$ agents follow a rule of thumb, “consume everything,” as in Campbell and Mankiw
  - $(1 - a)$ are rational optimizers
Motivation from the Data

Dispersion “miss” remains even when included in estimation objective.

- Learning might explain dispersion

- Examine simpler case first: “fraction spendthrift”
  - fraction $a$ agents follow a rule of thumb, “consume everything,” as in Campbell and Mankiw
  - $(1 - a)$ are rational optimizers

- This produces two models:
  - Model I: Estimate $(\beta, \rho)$ with all agents optimizing
  - Model II: Estimate $(\beta, \rho, a)$ with $a$ agents spendthrift, the rest optimizing
Defining Fraction “Rule of Thumb”

Take this a step further:

- Note $\beta \rightarrow 0$ captures “consume everything” in the rational model
  - Data dictates “degree” of spendthrift behavior
  - Jointly re-estimate model with parameters $(\beta_1, \beta_2, a, \rho)$
Defining Fraction “Rule of Thumb”

Take this a step further:

- Note $\beta \rightarrow 0$ captures “consume everything” in the rational model
  - Data dictates “degree” of spendthrift behavior
  - Jointly re-estimate model with parameters $(\beta_1, \beta_2, a, \rho)$

- Straightforward extension to $N$ types:
  - Model I: Estimate $(\beta, \rho)$ with all agents optimizing
  - Model II: Estimate $(\beta_1, \beta_2, \rho, a_1)$ with $a_1$ agents using $\beta_1$
  - Model III: Estimate $(\beta_1, \beta_2, \beta_3, \rho, a_1, a_2)$ with $a_1$ agents using $\beta_1, a_2$
    agents using $\beta_2$
  - ... 
  - Model $N$: Estimate $(\beta_1, \beta_2, ... \beta_N, \rho, a_1, a_2, ... a_{N-1})$ with
    $(1 - \sum_{1}^{N-1} a_n)$ following $\beta_N$
Defining Fraction “Rule of Thumb”

Take this a step further:

- Note $\beta \to 0$ captures “consume everything” in the rational model
  - Data dictates “degree” of spendthrift behavior
  - Jointly re-estimate model with parameters $(\beta_1, \beta_2, a, \rho)$

- Straightforward extension to $N$ types:
  - Model I: Estimate $(\beta, \rho)$ with all agents optimizing
  - Model II: Estimate $(\beta_1, \beta_2, \rho, a_1)$ with $a_1$ agents using $\beta_1$
  - Model III: Estimate $(\beta_1, \beta_2, \beta_3, \rho, a_1, a_2)$ with $a_1$ agents using $\beta_1$, $a_2$ agents using $\beta_2$
  - ... 
  - Model $N$: Estimate $(\beta_1, \beta_2, ... \beta_N, \rho, a_1, a_2, ... a_{N-1})$ with $(1 - \sum_{1}^{N-1} a_n)$ following $\beta_N$

- Each new type adds 2 parameters for a total of $2N$ parameters
How to Avoid Overfitting?

...you should be thinking to yourself, “overfitting!”

Formal model selection: BIC, AIC, k-fold cross validation.
Section 4

K-Fold Cross Validation
Quick Illustration

Consider the following artificial data:²

²Reproduced from Shalizi, “Advanced Data Analysis from an Elementary Point of View”
Polynomial Overfitting
R-squared Looks Great
Loss Function (SSE) Looks Great
However, Very Poor Fit

![Graph showing Mean Squared Error vs Polynomial Degree]

- **In-sample MSE**
- **Re-draw 30,000 times from DGP, MSE**
However, Very Poor Fit
Data Selection, $K = 5$
Data Selection, $K = 5$
Data Selection, $K = 5$
Data Selection, $K = 5$
K-Fold CV Summary

- Essentially a bootstrap of out-of-sample loss
- Many nice model selection properties

Some theory
K-Fold CV Summary

- Essentially a bootstrap of out-of-sample loss
- Many nice model selection properties

- Often recommended as “first model selection tool you should consider” (trivia: CrossValidated on StackExchange)
Section 5

Preliminary Results, Conclusion
Preliminary Results

Two sets of results:

- Early results from ’92-’07 SCF, full education sample
- Updated results from extended through ’16 SCF, education split by college+, HS+associates
- Will present early results, with updated results in parentheses / when indicated
Preliminary Results

- Strong indication of $N \geq 2$ types
Preliminary Results

- Strong indication of $N \geq 2$ types

- and for $N \geq 2$ roughly 46% (updated: 26%) of the population looks like Campbell/Mankiw “rule of thumb”

- Cross-validation plots: evidence for selecting $N \geq 2$ types:
K-Folds CV on N Types

Number of β types

Cross-validation score

+/- 1 sd
“Zoom In:” K-Folds CV: 2-6 Types

![Graph showing cross-validation score vs. number of $\beta$ types]
“Zoom In:” K-Folds CV: 2-6 Types
Preliminary Results

- Original, median-only, $N_{type} = 1$: $\beta = 1.01$, $\rho = 4.4$
- Original, median+IQR, $N_{type} = 1$: $\beta = 1.01$, $\rho = 1.65$

$^3$Due to: $p_{death}$, $\beta_\tau$. Average $\beta_{total} \approx \beta - 0.05$; average $\beta_{ignore\ death} \approx \beta - 0.01$. 
Preliminary Results

- Original, median-only\(^3\), \(N_{type} = 1\): \(\beta = 1.01, \rho = 4.4\)
- Original, median+IQR, \(N_{type} = 1\): \(\beta = 1.01, \rho = 1.65\)
- (Update, median+IQR, \(N_{type} = 1\): \(\beta = 1.01, \rho = 1.99\))

\(^3\text{Due to:} p_{death}, \beta_T. \text{Average} \beta_{total} \approx \beta - 0.05; \text{average} \beta_{ignore \ death} \approx \beta - 0.01.\)
Preliminary Results

- Original, median-only\(^3\), \(N_{\text{type}} = 1\): \(\beta = 1.01, \rho = 4.4\)
- Original, median+IQR, \(N_{\text{type}} = 1\): \(\beta = 1.01, \rho = 1.65\)
- (Update, median+IQR, \(N_{\text{type}} = 1\): \(\beta = 1.01, \rho = 1.99\))
- With \(N_{\text{type}} = 2 - 6\):
  - \(\rho \approx 4.6\)
  - \(\beta_{lo} \approx 0.25 - 0.45\)
  - \(\beta_{hi} \approx 1.04\)
  - ... with “low” fraction \(\approx 0.46\)

---

\(^3\)Due to: \(p_{\text{death}}, \beta_\tau\). Average \(\beta_{\text{total}} \approx \beta - 0.05\); average \(\beta_{\text{ignore death}} \approx \beta - 0.01\).
Preliminary Results

- Original, median-only\(^3\), \(N_{\text{type}} = 1\): \(\beta = 1.01, \quad \rho = 4.4\)
- Original, median+IQR, \(N_{\text{type}} = 1\): \(\beta = 1.01, \quad \rho = 1.65\)

- (Update, median+IQR, \(N_{\text{type}} = 1\): \(\beta = 1.01, \quad \rho = 1.99\))

- With \(N_{\text{type}} = 2 - 6\):
  - \(\rho \approx 4.6\)
  - \(\beta_{\text{lo}} \approx 0.25 - 0.45\)
  - \(\beta_{\text{hi}} \approx 1.04\)
  - ...with “low” fraction \(\approx 0.46\)

- (Updated, \(N_{\text{type}} = 2\)):
  - \(\rho \approx 3.8\)
  - \(\beta_{\text{lo}} \approx 0.00 - 0.15\)
  - \(\beta_{\text{hi}} \approx 1.00\)
  - ...with “low” fraction \(\approx 0.26\)

\(^3\)Due to: \(p_{\text{death}}, \beta_{\tau}\). Average \(\beta_{\text{total}} \approx \beta - 0.05\); average \(\beta_{\text{ignore death}} \approx \beta - 0.01\).
Original Estimation Results, $N \in (1, 2, 3, 4)$

<table>
<thead>
<tr>
<th>$N_\beta$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho$</td>
<td>1.65</td>
<td>4.65</td>
<td>4.94</td>
<td>4.24</td>
</tr>
<tr>
<td>$\beta$</td>
<td>1.01</td>
<td>0.25, 1.04</td>
<td>0.29, 0.99, 1.05</td>
<td>0.01, 0.41, 0.81, 1.04</td>
</tr>
<tr>
<td>$frac$</td>
<td>n.a.</td>
<td>0.46, 0.54</td>
<td>0.26, 0.35, 0.38</td>
<td>0.17, 0.19, 0.09, 0.54</td>
</tr>
<tr>
<td>$N_\beta$</td>
<td>5</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>--------------------</td>
<td>--------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\rho$</td>
<td>4.70</td>
<td>4.74</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta$</td>
<td>0.00, 0.01, 0.52, 0.79, 1.04</td>
<td>0.11, 0.16, 0.24, 0.28, 0.45, 1.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$frac$</td>
<td>0.09, 0.07, 0.15, 0.17, 0.53</td>
<td>0.03, 0.17, 0.09, 0.08, 0.1, 0.54</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Original, Weighted Combination of all $\beta_{lo} < \beta_{max}$

<table>
<thead>
<tr>
<th>$N_\beta$</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho$</td>
<td>4.65</td>
<td>4.94</td>
<td>4.24</td>
<td>4.70</td>
<td>4.74</td>
</tr>
<tr>
<td>$\beta_{{lo,hi}}$</td>
<td>0.25, 1.04</td>
<td>0.69, 1.05</td>
<td>0.34, 1.04</td>
<td>0.44, 1.04</td>
<td>0.25, 1.04</td>
</tr>
<tr>
<td>$frac_{{lo,hi}}$</td>
<td>0.46, 0.54</td>
<td>0.62, 0.38</td>
<td>0.46, 0.54</td>
<td>0.47, 0.53</td>
<td>0.46, 0.54</td>
</tr>
</tbody>
</table>
Original, $N_\beta = 2$ Consumption Functions

\[ \beta = 1.04 \text{ and } \rho = 4.7 \]

\[ \beta = 0.25 \text{ and } \rho = 4.7 \]
### Updated Results: Estimates by Education Level

Table 1: Updated Estimates for $N_{type} \in \{1, 2\}$

<table>
<thead>
<tr>
<th>$N_{type}$</th>
<th>All Edu</th>
<th>College</th>
<th>HS, Assc</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$\beta$</td>
<td>1.01</td>
<td>1.02</td>
</tr>
<tr>
<td></td>
<td>$\rho$</td>
<td>1.99</td>
<td>2.30</td>
</tr>
<tr>
<td>2</td>
<td>$\beta_1$</td>
<td>0.00</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td>$\beta_2$</td>
<td>1.00</td>
<td>1.01</td>
</tr>
<tr>
<td></td>
<td>$a_1$</td>
<td>0.26</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td>$\rho$</td>
<td>3.76</td>
<td>4.26</td>
</tr>
</tbody>
</table>
Empirical Wealth Distribution: College Educated, Married

(Net Worth)/(Perm Income) by Age Cohort

- age: 1, IQR: 1.5
- age: 2, IQR: 2.1
- age: 3, IQR: 3.1
- age: 4, IQR: 3.6
- age: 5, IQR: 4.3
- age: 6, IQR: 5.5
- age: 7, IQR: 7.4
Estimated Wealth Distribution: Median & IQR, $N_{type} = 2$:

(Net Worth)/(Perm Income) by Age Cohort
Estimated Wealth Distribution: Median & IQR, $N_{type} = 1$:

(Net Worth)/(Perm Income) by Age Cohort
Summary, Next Steps

Summary:

- I formally select between structurally estimated models of household consumption-savings behavior
- Evidence for $N \geq 2$ types
- ...with ~25% of the population using spendthrift “rule of thumb”

Next steps:

- Data: alternative measures of wealth
- Selection: alternative k-fold formulation
- ...full unit testing; speed via compilation / parallelization
- Selection with learning
Section 6

Appendix
Original SCF Data, All Education Levels

(Net Worth)/(Perm Income) by Age Cohort, All Edu

- age: 1, IQR: 1.7
- age: 2, IQR: 2.9
- age: 3, IQR: 5.1
- age: 4, IQR: 8.6
- age: 5, IQR: 10.8
- age: 6, IQR: 13.7
- age: 7, IQR: 18.3
Updated SCF Data, All Education Levels

(Net Worth)/(Perm Income) by Age Cohort, All Edu
Updated SCF Data, College+

(Net Worth)/(Perm Income) by Age Cohort, All Edu

First data look

Results
Updated SCF Data, High School + Associates

(Net Worth)/(Perm Income) by Age Cohort, All Edu

First data look

Results
Different Learning Implies Different Economic Behavior

- Reinforcement learning ("AI") solves the same DSOP as dynamic programming. We should intellectually arbitrage.
- Different methods imply different behavioral patterns

Source: Fig 1.5 in Cao (2007), “Stochastic Learning and Optimization”
Geanakoplos et al. (2012), “Getting at Systemic Risk via ABM of the Housing Market”

- Hypothesis: the joint distribution of HH income, wealth, housing status, mortgage status (type, duration, LTV, DTI), credit score, had 1st-order effects on housing crisis dynamics
- Focus on housing market for Washington DC MSA, 1997-2009
- Markets:
  - GE housing market via matching
  - PE labor and mortgage markets via historical data
- Behavior: policy derived from literature, estimated on data
  - importantly, “nuisance policy parameters” not updated
Housing Market, Experiment

- Early policy experiment: low rates vs credit access
Housing Market, Experiment

Case-Shiller in the baseline simulation

Case-Shiller with interest rates fixed at 1997 levels

Not shown: $\sim 3 \times 7$ other moments
Not shown: $\sim 3 \times 7$ other moments
Housing Market HH Policy Function

One version of desired housing expenditure component:

\[ P = \frac{\epsilon_i \times h \times \text{Income}^g}{(\tau + c) + \text{LTV} \times \text{PrimeRate} - a \times \mathbb{E}[\text{HPA}]} \]

where

- \( g \) is income spent on housing
- \( h \) is scaling factor (eg. from prefs)
- \( \text{LTV} \) is desired LTV
- \( a \) is house price appreciation sensitivity
- \( \mathbb{E}[\text{HPA}] \) expected HP appreciation
- \( \text{PrimeRate} \) is contemporaneous 30-year prime mortgage market rate
- \( (\tau + c) \) are fixed costs, \( \epsilon_i \) is agent-level heterogeneity term
- Expectations easy to update but “nuisance” policy parameters hard
Housing Market Data Sources

- Focus on Washington DC MSA, 1997-2009, ~1.6 million HHs in 1997

- Data sources:
  - Local: Core Logic, MLS, IRS, Loan Performance, Census
  - National: PSID, CEX, ACS

- Data goal: obtain joint distribution of income, wealth, housing status, mortgage status (type, duration, LTV, DTI), credit score for HHs, at individual level. “Structural data combination.”
K-Fold Cross Validation - Some Asymptotic Properties

- For linear models, leave-d-out cross validation\(^4\):
  - is equivalent to BIC for \(d = n \left(1 - \frac{1}{\log(n) - 1}\right)\)
  - is asymptotically efficient if
    \[
    \frac{d}{n} \to 1 \quad \text{and} \quad \frac{n_{\text{var}}}{n - d} \to 0
    \]
  - is consistent if correct model in selection set

- For all likelihood-based models, leave-1-out CV is equivalent to AIC as \(n \to \infty\) (with associated issues as well)


\(^4\)k-fold CV is a special case of leave-d-out CV where \(d = \frac{1}{k} n\).
### Calibration and Sources

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Value</th>
<th>Source and Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma_{xi}$</td>
<td>0.01</td>
<td>Carroll (1992); standard deviation of temporary shock to income.</td>
</tr>
<tr>
<td>$\sigma_{psi}$</td>
<td>0.01</td>
<td>Carroll (1992); standard deviation of permanent shock to income.</td>
</tr>
<tr>
<td>$R_t$</td>
<td>1.037</td>
<td>Jorda et al. (2019); Table 11, average US post-1980 real annual safe return.</td>
</tr>
<tr>
<td>$p_{ymin}$</td>
<td>0.0115</td>
<td>Guvenen et al. (2019); $p_{nu}$ from Model 2. Prob of near-0 income shock.</td>
</tr>
<tr>
<td>$y_{unemp}$</td>
<td>0.00015</td>
<td>Guvenen et al. (2019); $Y_{min}$ from Model 2. Size of near-0 income shock.</td>
</tr>
<tr>
<td>$/D_t$</td>
<td></td>
<td>Cagetti (2003). $1 - \text{prob}(death)$</td>
</tr>
<tr>
<td>$\beta_{\tau,t}$</td>
<td></td>
<td>Cagetti (2003). Lifecycle discount factor</td>
</tr>
</tbody>
</table>