Structural Estimation in Economics and Finance

Mehran Ebrahimian

Fall 2023

Plan

- Introduction
  - What is structural estimation?
    - terminology
    - model, data, identification
    - structural vs. reduced-form estimation
  - Why do it?
    - what structural estimation buys you?
    - how to motivate a structural estimation paper?
    - advantages and disadvantages vs. reduced-form approach
- Identification in the structural approach
  - Basic concepts, How it works in practice?
  - Examples: Identification in papers doing structural estimation
- "Light-speed" structural estimation
  - The approximate moment technique
  - Applications, Robustness checks in structural research
- Papers doing structural estimation (by you!)

1. Introduction

1a. What is structural estimation?

The big picture

Empirical research:

- Reduced-form
  - descriptive
  - clever identification
    - IV
    - Diff-in-Diff
    - Regression discontinuity
    - ...
- Structural estimation

Some terminology

- It makes no sense to say "structural model"
  - every model imposes structure on the real world
- Usually, by structural model people mean economic model
- It makes sense to say structural estimation versus reduced-form estimation
  - reduced-form estimation: uses statistical model
  - structural estimation: uses economic model
- It is about *what* you estimate, not *how* you estimate!
  wrong to say "I structurally estimate a model of..."
  correct to say "I estimate structural parameters of..."

Statistical and Economic models

- A **statistical model** describes the relationship between two or more random variable. Example:

$$y = xb + \epsilon$$

- An **economic model** makes assumptions about
  - Preferences (demand side?)
  - Production technology (supply side?)
  - Constraints
  - Institutional features and policy framework
  - Some notion of equilibrium
  - ∗ note: a relationship between observable variables x and y is implied, regardless

What is structural estimation?

- Estimate an **economic model**'s parameters, including
  - preference parameters (e.g. risk aversion, time discount factor)
  - technology parameters (e.g. marginal costs, production function curvature)
  - other 'deep' institutional features (e.g. agents' bargaining, financing frictions)

  $\rightarrow$ learn/measure relevant economic forces/"economic" decomposition
- Use the model to run **counterfactual experiments**
  - thought experiments, example: what if there were no financing frictions?
    - deepens our understanding of relevant forces in real world
    - hints to find the optimal policy
  - what if a fundamental parameter changes? example: preferences towards minorities...
  - policy experiments
    - the ones that has been experienced in the past
    - a famous policy proposal, or author's proposed policy

Examples

- Measurement:
  - How large is the consumer surplus in the bank deposit markets?
  - What share of conditional college education gap is explained by borrowing constraints?
  - How changes in hospital ownership after private equity (PE) buyouts impact hospital–insurer price negotiations?
- Understanding economic forces:
  - Why wealth inequality is rising?
  - Why interbank lending arrangements arise?
  - What explain low stock market participation among young households?
- Policy/Counterfactual:
  - What happens to the wealth inequality by imposing capital income tax?
  - Do financially constrained students switch to high-quality colleges by making them tuition-free?
  - What is the patient-surplus gains/losses in response to restricting PE ownership/hospital mergers?

A classic example

- Demand function (from utility maximization):

$$q = -\alpha p + \omega D$$

- Supply function (from production technology):

$$q = \beta p + \eta S$$

- $\uparrow$ "Structural" representation
- Solve the two equations for two unknown
- $\downarrow$ "Reduced-form" representation

$$q = \frac{\beta\omega}{\alpha+\beta} D + \frac{\alpha\eta}{\alpha+\beta} S$$

$$p = \frac{\omega}{\alpha+\beta} D - \frac{\eta}{\alpha+\beta} S$$

A classic example, cont.

- Reduced-form estimation: describes the impact of S, D on q, p
- Structural estimation: describes the underlying economics
  - supply/demand elasticity
  - measure consumer/producer surplus
  - find the impact of taxation
- Note: there is always a link between the reduced-form and the structural representation (coefficients map to each other!)

Structural vs. Reduced-form (Questions)

Reduced form:

- What is the (causal) effect of X on Y?

  ...what would happen if certain things changes but not others

Structural estimation:

- Why does X cause Y?
- What are relevant economic forces—deep parameter magnitudes?
  "Parameters" = economic primitives
  "Parameters" $\neq$ regression slopes/correlations
- If an economic theory is validated with data
- How would the world look like if a (deep) parameter changes?
- What would happen if you change policy or institutional setting?

  ...everything possibly changes in response to a policy!

Structural vs. Reduced-form (Tools)

Reduced form:

Estimators:

- OLS
- IV
- Diff-in-Diff
- RD

Software:

- Stata

Structural estimation:

Estimators:

- GMM
- SMM
- MLE
- SMLE
- Indirect inference

Software:

- C++, Matlab, Julia, Python

Solving the model:

- Value/policy function iteration
- ODE/PDE (continuous time models)
- Simulation

Structural vs. Reduced-form (Terminology)

- All economic models imply a "reduced-form" representation:
  - a statistical (linear) model describing the relation between observables generated by the model
- Regression slopes though are general functions of the model's deep parameters, and policy and institutional settings
  - **Lucas critique**
  - Limited scope of results
- The true relationship between variables (endogenous or exogenous) may be nonlinear as well
  - the noise term captures approximation error in the linear model
- A key note: we do not do structural estimation instead of reduced-form, to not be worried about **identification**!
  - we need a legit source of variation/piece of information anyway
  - identification is the core of a structural estimation project

Calibration vs. Structural Estimation

Both:

- have an economic model
- can ask how model fits data, given the various ingredients
- can use to run counterfactual

But,

Calibration:

- Take parameter values from other papers
- Usually have more parameters than moments of interest
- Cannot perform statistical tests
- Mainly a theoretical exercise

Structural estimation:

- 'Infer' parameter values from the data$\rightarrow$ *identification*
- Get standard errors for parameters/objects of interest
- An empirical exercise

Structural work needs an **economic** model

Requirements for the economic model:

- An economic model, consisting deep parameters, not a 'statistical' model
- Should include relevant and competing economic forces
  - the model should be able to reject your prior!
- Do not write a big complicated model
  - minimalistic, as long as it meets the previous point
- Should produce realistic magnitudes and distributions
  - both in levels and elasticities
  - both in terms of observables and unobservables in your data

Further notes on economic model

- Make sure you understand all ingredients and mechanics of the model
  - do a lot of comparative statistics
  - plot choice/state vars against each other
  - generate a lot of graphs and tables of summary statistics and regression, before even introducing data
- Being simple and small is the objective, having competing forces is the constraint
- Do not name 'residuals': sometimes a block of the model is about what the other key block is *not* about
  - be clear with the audience; it's okay to have residuals, but your interpretation needs to be correct

Estimation

- Inferring model parameters from data...
- Econometricians define an objective function over parameters & data
- The objective function depends on the estimator you use:
  - GMM: the distance between model-generated moments and data moments...
  - MLE: the model-implied likelihood that you observe certain outcomes in the data
- Goal: find a parameter that minimizes the objective function
- A parameter is **identified** if there is a *unique* minimum

* **Identification is the core of a structural estimation project...**
* full discussion $\rightarrow$ next lecture!

Estimation (GMM), a short review

- The model implies a moment condition in the form: $\mathbb{E}[g(x,p)] = 0$
  - x: data. p: parameters. $g(\cdot)$: implied by model
  - example, Euler eq. in Asset pricing: $\mathbb{E}[\beta(c_{t+1}/c_t)^{-\gamma}(r_{t+1} - r_f)] = 0$
  - example (minimum distance/usual smm): $g(\cdot) = m(p) - x \Rightarrow m(p) - \mathbb{E}[x] = 0$
  - note: $g(\cdot)$ can be a *vector* of moments: multiple moment conditions
- The corresponding sample moment restriction: $1/N \sum_{i=1}^{N} g(x_i, p) = 0$
- Find $\hat{p}$ that gets $1/N \sum_{i=1}^{N} g(x_i, p)$ as close as possible to zero
- Do so by minimizing:

$$\hat{p} = \arg\min_{p} \left[1/N \sum_{i=1}^{N} g(x_i, p) \right]^{T} W \left[1/N \sum_{i=1}^{N} g(x_i, p) \right]$$

  - W is a positive definite symmetric weight matrix

Summary — Project stages

1. Develop the theory model
2. Solve/simulate the model
3. Understand how the model works
4. Pick a bunch of moments that you think identify parameters
5. Collect and clean data
6. Estimation
7. Evaluate the fit: show targeted, and untargeted moments
8. Run counterfactual/policy experiment

(And writing/making slides throughout)

Source: Michael Keane, "Practical issues in structural estimation," `https://www.youtube.com/watch?v=0hazaPBAYWE`

1. Introduction

1b. Why doing structural estimation?

What structural estimation does **not** buy you?

- You always need to convince audience, why you do structural estimation
- Remember: It's a lot of pain to do it; so why you do it?
- It imposes large costs on the reader as well

$\Rightarrow$ Any structural paper must put great effort into convincing reader that it's worth going structural

- A wrong but common :(justification:

"...to not deal with identification challenges"

$*$ Structural estimation does not substitute for identification
  - Identification is in the core of estimation anyway...

What structural estimation buys you?

- The correct justification (from the least to the most important one):

1. Estimates of interesting economic primitives
2. Tests of economic theories (quantitative relevance, not just qualitative or direction of changes)
   - no worriers if model fails: you call it a puzzle :) a whole literature then starts

3a. Welfare analysis

3b. Interesting counterfactual and policy analysis[1]

[1] Reduced-form papers can also ask counterfactual questions, by changing a regressor from its actual value to a counterfactual value. But it's usually less convincing, it's hard to believe in/meaningless to say "all else equal." If you change a primitive, everything changes, nothing remains equal. Lucas critique...

What structural estimation buys you?

More **technical** notes:

- If you don't have (plausibly) exogenous variations in data; not a bold event in a Diff-in-Diff context, say

$\Rightarrow$ use an economic model to account for the *endogenous* relationship between variables in data

- If you don't have data on everything$\rightarrow$ the issue of omitted variable...

$\Rightarrow$ Economic model would infer what's missing, by imposing the *economic* relationship: "marginal cost=marginal benefit"

  + you need to know parameters/date governing marginal cost *or* marginal benefit, not both

Advantages/Disadvantages I

Reduced form:

Approach:

- regress variables of interest y on a bunch of determinants X

Data challenges:

- Needs sufficient heterogeneity in X

Identifying assumptions:

- Exogenous variation in X
- Got the functional form right

Structural estimation:

Approach:

- Estimate economic parameters governing y wrt X

Data challenges:

- Do *not* need heterogeneity in X; can infer parameters from mean of X and y

Identifying assumptions:

- Model is true
  - Includes all determinants of y
  - Functional forms are correct
  - Rational agents...
  - ...

Advantages/Disadvantages II

Reduced form:

- – "Fewer" assumptions? No...
- people forget that reduced-form approach also relies on a model,
- and specifies a relationship...
- \+ easier to do
- \+ easier to understand
- → larger audience

Structural estimation:

- \+ The only feasible option for answering certain important questions
- \+ applicable when it's hard to find good instruments or natural experiments
- \+ tight link with theory
- → transparent interpretation of results with economic stories
- → results generalize better

Complements not substitutes

The structural approach **complements** reduced-form research by

1. overcoming certain data challenges
2. imposing a different type of identifying assumption
3. less problematic extrapolation
4. welfare analysis
5. policy analysis

Concluding Remarks

- Do what lets you answer your research question most convincingly and easily
- If structural and reduced-form will both get the job done, go reduced-form!
- Going structural may be right for you if...

  ... there's not much on your calendar for the next few years :)
- But better signal for the job market!

2. Identification in the structural approach

2a. Basic concepts, How it works in practice

Identification

- **Inferring model parameters from data...**
- Econometrician defines an objective function over parameters and data
- The objective function depends on the estimator you use:
  - GMM
  - MLE (maximum likelihood)
  - SMM (simulated method of moments)
  - SMLE (simulated maximum likelihood)
  - Indirect Inference
  - Score estimators
  - The two-step methods used by structural IO folks...
- Goal: find a parameter that minimizes the objective function
- A parameter is identified if there is a unique minimum

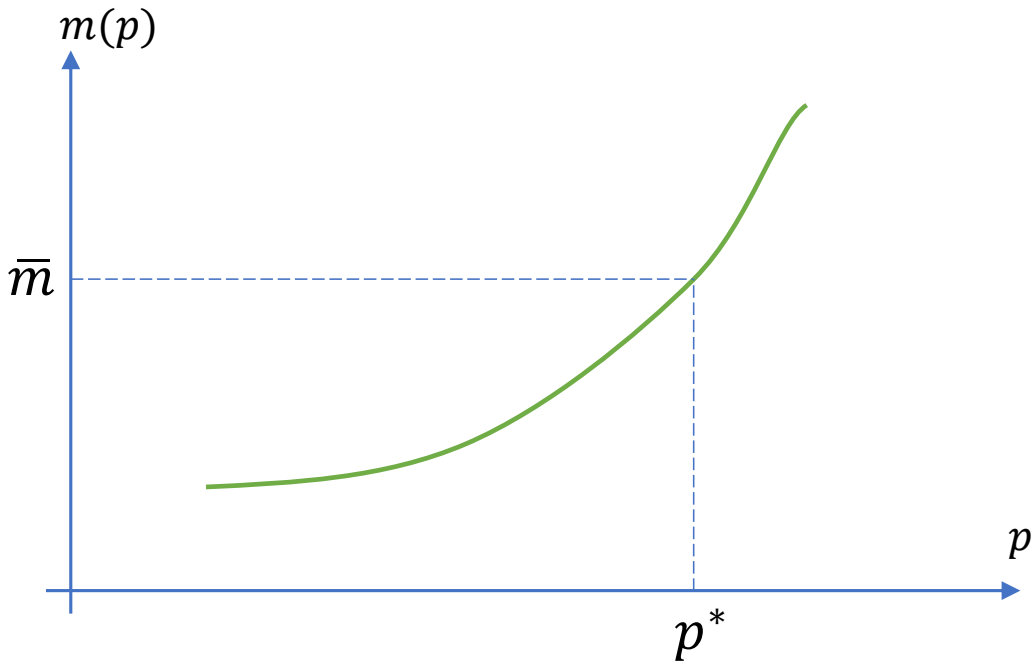
Moments and Likelihoods

- The moment estimators determine whether model-implied moments match with real-data moments
  - $-/+$ choice of moments is subjective and arbitrary
- The likelihood estimators use the economic models to construct the likelihoods for observed patterns in data
  - $+$ asymptotically efficient; delivers the least standard errors
  - $-$ fully parametric
- In both cases:
  - The simulation estimators (SMM and SMLE) are used with models that don't have closed-form estimating equations (**extremely slow**)
  - GMM and MLE are used with models that have closed-form estimating equations (**fast**)
- It really doesn't matter which estimator you choose, as long as the model is *well-identified*...

Identification in SMM

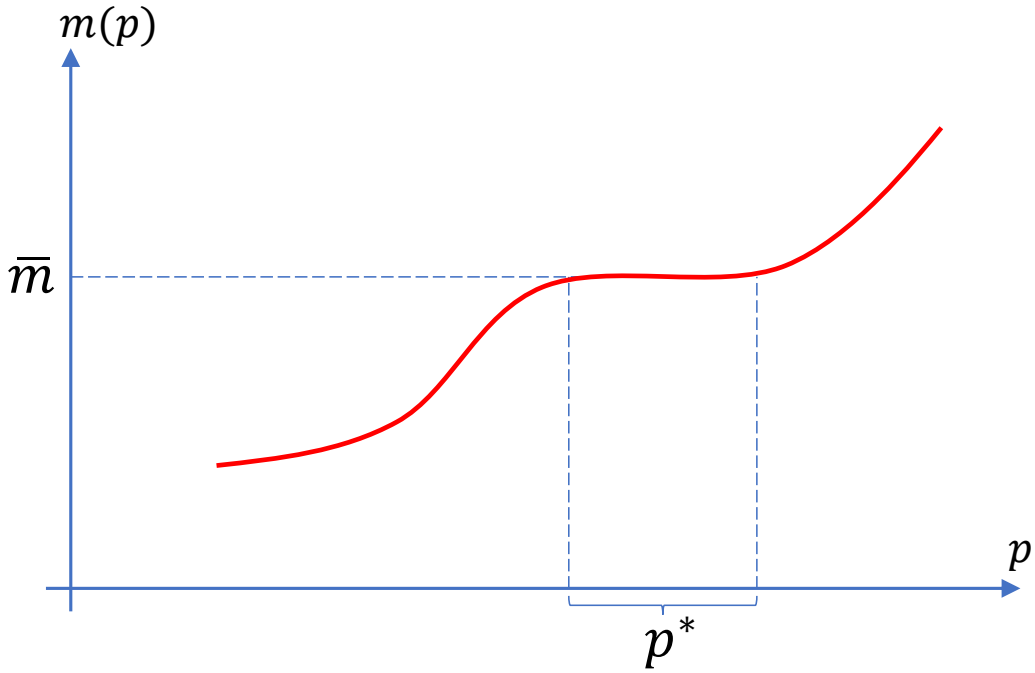
$$\textit{parameter } p \rightarrow \boxed{\textit{Economic Model}} \rightarrow \textit{observable outcomes} \rightarrow \textit{moment } m(p)$$
$$\textit{same variables in data} \rightarrow \textit{moment } \bar{m}$$

- Estimator: tune p to get $m(\cdot)$ as close as possible to the data



Identification failure

- What if m(p) is (almost) flat in 'relevant' regions? multiple solutions for p^*, solving $m(p^*) = \bar{m}$...



$\rightarrow$ The model is not identified!

- The problem is not necessarily solved by changing the estimator: *MLE* may be flat wrt p as well...

Identification diagnostics

- Reduce model ingredients, get more data and target a *relevant* moment:
- Pick a moment that is *sensitive* to economic parameter
- Good news: if the model is not identified, standard errors are arbitrary large[2]

$$Var(p) = (J'\Omega^{-1}J)^{-1}$$

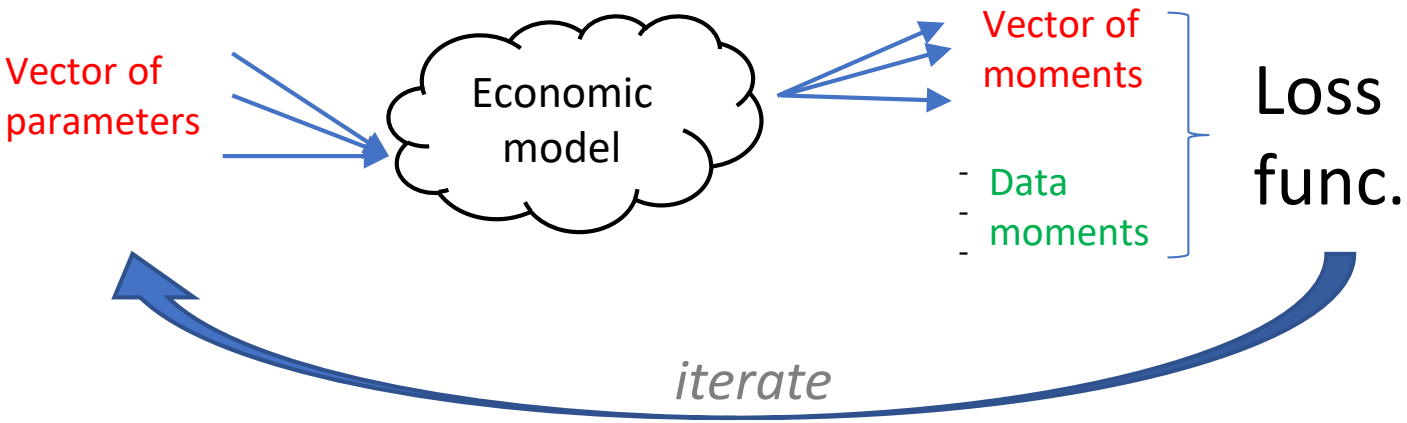
  - where $J = dm/dp$. So $J = 0 \Rightarrow se(p) = \infty$
- Bad news:
  1. you may get finite SEs, just because of numerical/simulation errors in calculating J
  2. It's all based on *local* deviations anyway...
- Better to use economic intuition first; then plot $m(p)$ for a *wide* range, all before getting to data...

[2] This is the formula for SEs when the weight matrix is efficient (see the next slide). but SE is infinity if $J = 0$ regardless.

Identification in SMM, general form

- Usually models have *multiple* parameters to be estimated: trade-off between two econ. forces, e.g.
- Pick a *set* of moments that are *collectively* informative about the *set* of parameters.
- Form $\mathbf{m}(\mathbf{p})$, where both $\mathbf{m}$ and $\mathbf{p}$ are vectors. Tune $\mathbf{p}$ to get $\mathbf{m}(\cdot)$ as close as possible to the vector of data moments $\bar{\mathbf{m}}$ with the metric W

$$\mathbf{p}^* = \arg\min ||\mathbf{m}(\mathbf{p}) - \bar{\mathbf{m}}||_W = \arg\min \ [(\mathbf{m}(\mathbf{p}) - \bar{\mathbf{m}})^T W (\mathbf{m}(\mathbf{p}) - \bar{\mathbf{m}})]$$

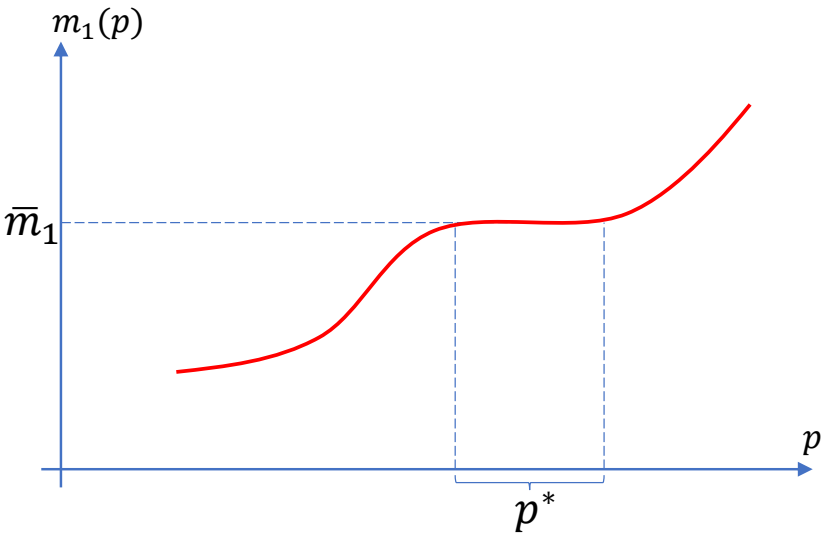


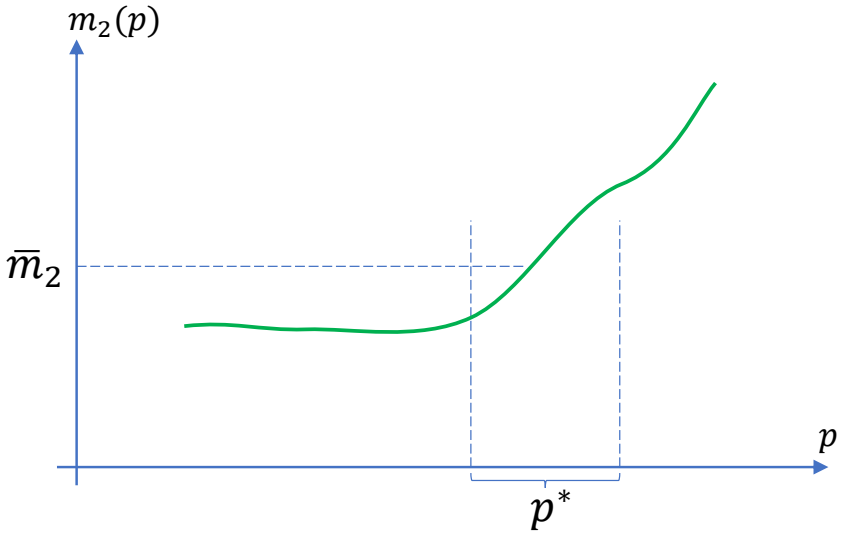
Identification diagnostics, cont.

- A sufficient condition for identification:
  one-to-one mapping; moment i is strictly increasing (or decreasing) in parameter i, independent of other parameters
  - matching moment $m_i(\cdot)$ with data counterpart $\bar{m}_i$ identifies p_i
- Usually, there is no one-to-one mapping between parameters and moments; every moment depend on every parameters
- The mapping needs to be **onto**
- Hard to check in reality, involves an inversion $\rightarrow$ more on this in the next lecture!
- Do a lot of comparative statistics, think a lot on the choice of moments, rely on economic intuition, read papers in the literature to get identification idea...
- Make sure the model is identified, *before* getting to data!

"Over-identified" models

- Target more moments than parameters, to ensure identification





The choice of weight matrix W / Standard errors

- How to reflect prediction error regarding various moments
- Irrelevant, if #moments = #parameters (exactly identified-)
- The efficient choice: inverse var-cov matrix of targeted moments
- The Robust choice: diagonal, appropriately scaled
- Standard errors (general representation):

$$Var(p) = (J'WJ)^{-1}(J'W\Omega WJ)(J'WJ)^{-1}$$

2. Identification in the structural approach

2b. Examples

Example 1: "Household Finance"

Catherine, 2022

A simplified version:

- The individual maximize

$$\mathbb{E}\sum_{t=t0}^{T}\beta^t\Big(\prod_{k=t0}^{t-1}(1-m_k)\Big)\frac{C_{it}^{1-\gamma}}{1-\gamma}$$

such that

$$W_{it+1} = \{W_{it} + L_{it} - C_{it} - f_{it}\}[\pi_{it} r_{t+1} + (1-\pi_{it}) r_f]$$

- C is consumption and π is risky share in portfolio (choice variables), W is wealth at time t
- m is age-dependent mortality rate, L_{it} is labor income (plus social security/retirement benefit minus taxes), r_t is the risky stock return and r_f is the risk free rate $\rightarrow$ all calibrated from data
- $f_{it} = \Phi L_{it}$ is the participation fee (in the stock market), only if $\pi_{it} > 0$

Example 1: "Household Finance"

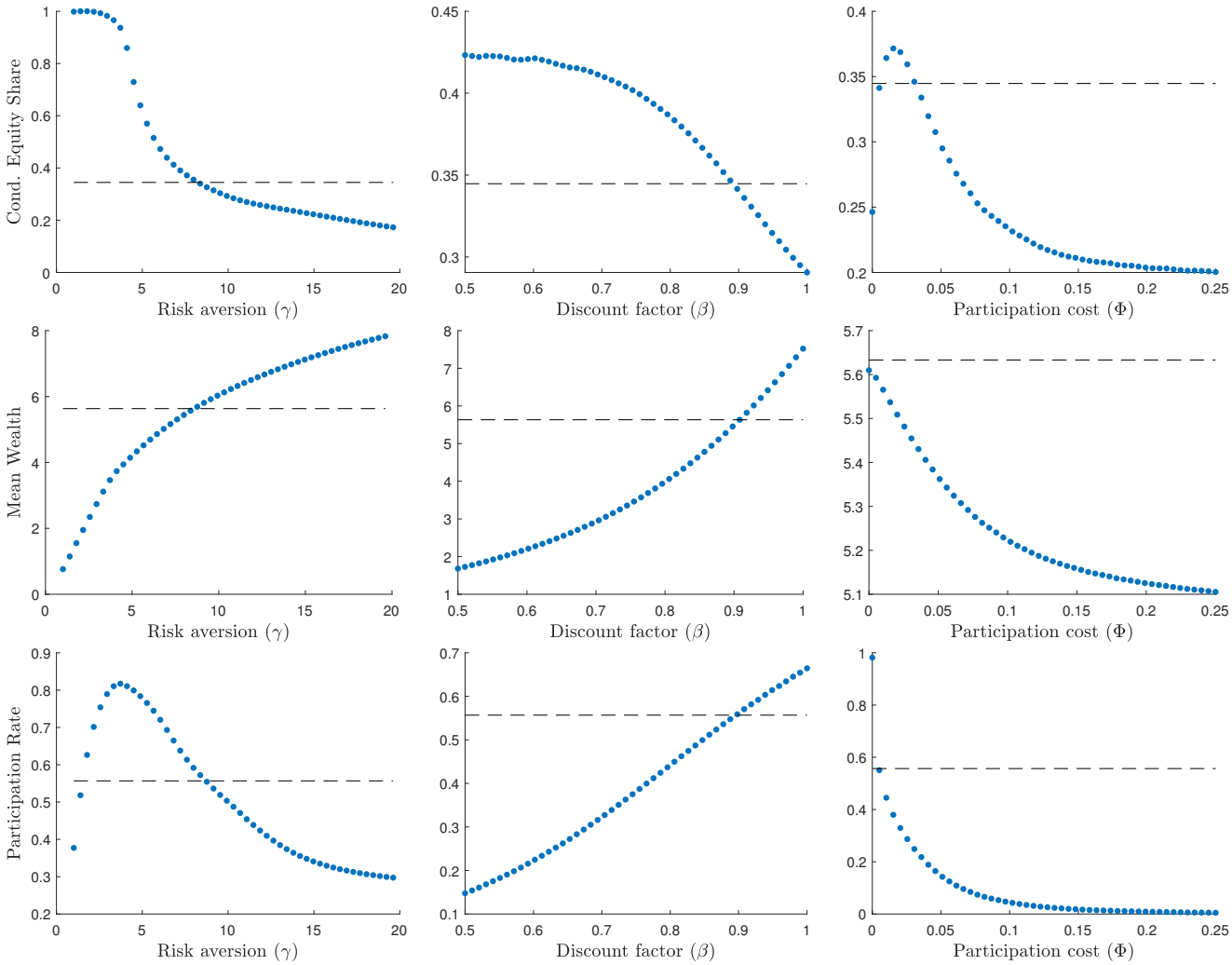
Catherine, 2022

- Deep parameters to be estimated:
  - β: patience
  - γ: risk aversion
  - Φ: participation cost
- Identifying moments:
  1. Average wealth (normalized by wage): $\mathbb{E}[W/L]$
  2. Average risky share in portfolio (conditional): $\mathbb{E}[\pi|\pi > 0]$
  3. Fraction of population who hold stocks: $\mathbb{E}[\pi > 0]$

Example 1: "Household Finance"

Catherine, 2022

Moments are informative about parameters + onto mapping!



Example 2: "Corporate Finance"

Hennessy and Whited, 2007

A simplified version:

- The firm maximize

$$V(k, A) = \max_{i} \quad Ak^{\alpha} - i + \frac{1}{1+r} \cdot \mathbb{E}_{A'}[V(k', A')|A]$$

such that

$$k' = (1 - \delta)k + i$$

- V is the market cap
- k is the stock of capital
- i is investment, builds tomorrow's capital k'
- A is realized TFP
- A' is TFP, tomorrow; usually we consider AR(1) spec. $a' = \rho a + \sigma\epsilon$; little a is log A

Example 2: "Corporate Finance"

Hennessy and Whited, 2007

- Deep parameters to be estimated:
  - α: return to scale
  - δ: capital depreciation
  - ρ, σ: persistence and volatility of TFP
  - $*$ we usually calibrate r
- Straightforward to show (analytically!):

$$k' = \left\{ \frac{\alpha \mathbf{E}[A'|A]}{r+\delta} \right\}^{1/(1-\alpha)}, y' = A' \left\{ \frac{\alpha \mathbf{E}[A'|A]}{r+\delta} \right\}^{\alpha/(1-\alpha)} \Rightarrow y'/k' = \frac{A'(r+\delta)}{\alpha \mathbf{E}[A'|A]}$$

$$\mathbf{E}[k'] = \mathbf{E}[k] \Rightarrow \mathbf{E}[i] = \delta \mathbf{E}[k]$$

- Identifying moments:
  1. profitability: $mean(y/k)$
  2. investment ratio: $mean(i/k)$
  3. volatility/autcorr of operating profit $Std(y), Corr(y, y_{-1})$

Example 2: "Corporate Finance"

Hennessy and Whited, 2007

- The original model includes capital adjustment cost $C(i, k)$...
- ...and equity issuance cost $-\xi D\mathbf{1}\{D < 0\}$, for $D = Ak^{\alpha} - i - C$
  - ▷ financing frictions!

No analytical solution, but the intuition holds...

- Target benchmark moments, plus...
  1. auto-correlation of investment: $Corr(i, i_{-1})$
  2. mean equity issuance: $mean(-D\mathbf{1}\{D < 0\}/k)$

Example 3: "Discrete-choice models"

Berry, Levinsohn, Pakes, 1995, 2004

- Consumer s has the utility v_{su} for product u:

$$v_{su} = -\alpha p_u + \beta \cdot X_u + \xi_u + \epsilon_{su}$$

  - p_u: price. α: price-elasticity
  - X_u: vector of product characteristics
  - ξ_u: common residual taste shock
  - ϵ_{su}: iid random shock across consumers/products
  - tag the outside option "u=0" with $p_u = X_u = \xi_u = 0$
- Discrete choice model: consumer s chooses product u iff

$$\forall u' : v_{su} > v_{su'}$$

- Analytical solution for $Prob[v_{su} > v_{su'}]$ if $\epsilon_{su} \sim$ type 1 extreme value distribution (logit shock)

$$prob_{su} = \frac{\exp(-\alpha p_u + \beta \cdot X_u + \xi_u)}{1 + \sum_u \exp(-\alpha p_u + \beta \cdot X_u + \xi_u)}$$

Example 3: "Discrete-choice models"

Berry, Levinsohn, Pakes, 1995, 2004

Define $\delta_u = -\alpha p_u + \beta \cdot X_u + \xi_u$

- $prob_{su}$ is independent of s
- market share of product u is

$$mktsh_u = \exp(\delta_u)/(1 + \sum_u \exp(\delta_u))$$

- Identify $\{\delta_u\}$ via the observed market shares in data (a standard fixed-point algorithm converges)
- How to identify coefficients β, α?
  1. $\xi_u \perp X_u$ (like in OLS regr of δ_u on X_u)
  2. Find an IV to estimate α...
     - $Cov(C_u, \xi_u) = 0 \rightarrow \alpha = \frac{-Cov(C_u, \delta_u)}{Cov(C_u, p_u)}$, for some cost-shifter C_u

Example 3: "Discrete-choice models"

Berry, Levinsohn, Pakes, 1995, 2004

Can be generalized for α, β depending on s (demographics, D_s):

- $\alpha_s = \alpha + \Gamma D_s$, $\beta_s = \beta + \Lambda D_s$
- $prob_{su}$ follows the same eq., but it depends on s
- $mktsh_u$ is derived numerically (by taking the sum over the population with heterogeneous D_s)
- Same approach to identify mean-tastes $\{\delta_u\}$ and mean coefficient α, β
- Plus, target as the identifying moment the covariance between demographics and choice characteristics: $cov(D_s, X_{u(s)})$

Example 4: "Student Loans and Social Mobility"

Ebrahimian, 2022

A simplified version:

- Students of different backgrounds with cash-in-pocket m and college-related ability A pursue higher education

$$\max Ah^{\alpha} - R\overbrace{(h-m)}$$

h : investment in human capital, proxy with college tuition
$A = \exp(a_0 + a_1 * SAT + \sigma * \nu)$: proxy ability with observed SAT score, and an observable shock
$L := h - m$ is student loans

- $R = R_0$: 10yr treasury (risk-neutral agents) if $L \leq \bar{L}$ and $R_0 + \eta$ if $L > \bar{L}$: friction wedge on taking private loans

$$\text{off the loan limit:} \quad h^* = (\alpha A/R)^{1/(1-\alpha)}$$

Example 4: "Student Loans and Social Mobility"

Ebrahimian, 2022

- Suppose we knew η, so $R(L)$

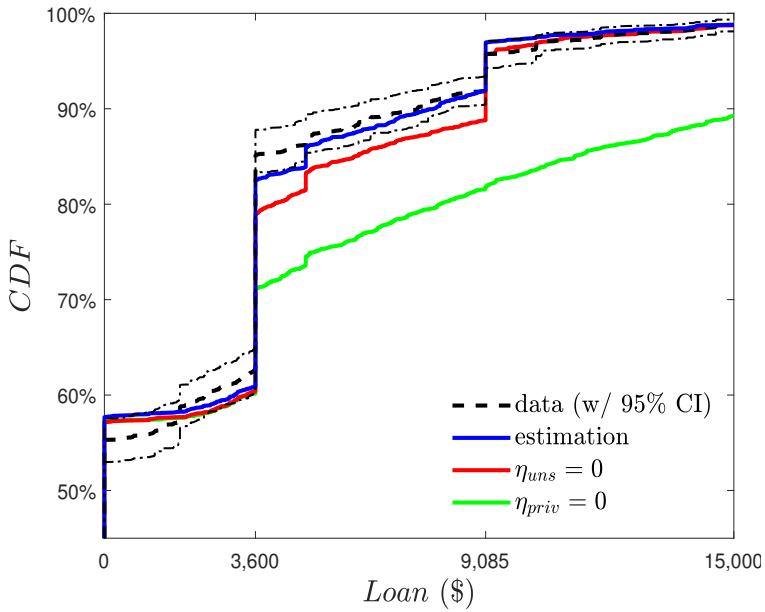
$$\log(h^*) = \frac{1}{1-\alpha}[\log\alpha - \log R_0 - \mathbf{1}\{L > \bar{L}\}\log(1+\frac{\eta}{R_0}) + a_0 + a_1 SAT + \sigma\nu]$$

- Simulate a set of random shocks ν, m, model the tuition
- Run the regression of *log* of tuition on SAT, and a loan level indicator $\mathbf{1}\{L > \bar{L}\}$ in the model, for those with $L \neq \bar{L}$
- Run the same regression for *log* of tuition in the data
- Identify a_1, α, σ: regression slope of SAT and $\mathbf{1}\{L > \bar{L}\}$, the variance of regression residual
  - identifying moments: the average of *log* tuition, for subsamples $L > \bar{L}$ and $L < \bar{L}$, covariance of log tuition and SAT, and std of log tuition
- Note: L and so $\mathbf{1}\{L > \bar{L}\}$ is endogenous: it covaries with ν! and m... no worries :) same endogeneity is in the simulated regression...

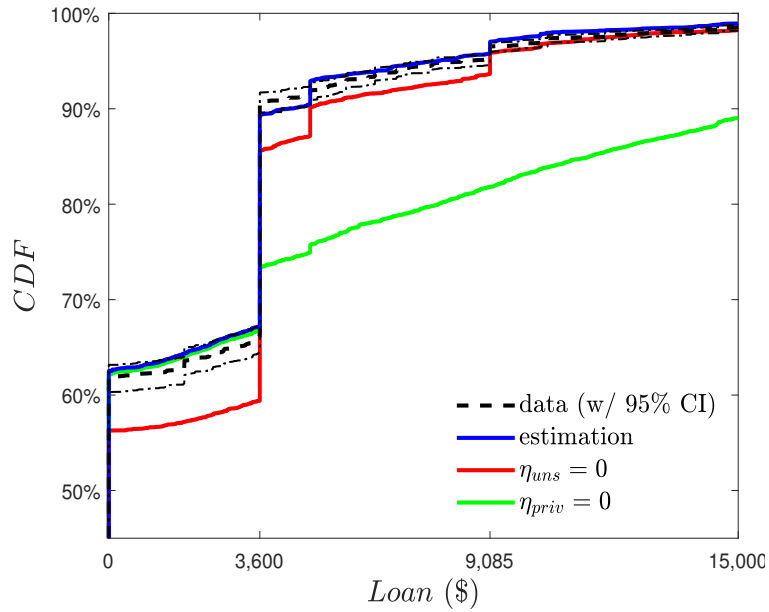
Example 4: "Student Loans and Social Mobility"

Ebrahimian, 2022

- How to estimate η? Bunching: mass in population with $L = \bar{L}$
- Intuition: costly private loans hinders students from increasing leverage beyond federal cap
- Math: $L = \bar{L}$ if $A \in [A_d, A_u]$, where $m + \bar{L} = (\alpha A_d / R_0)^{1/(1-\alpha)}$ and $m + \bar{L} = (\alpha A_u / (R_0 + \eta))^{1/(1-\alpha)}$, $A_u - A_d$ increases with η



(a) low-income students



(b) high-income students

Example 5: "Intermediation in the Interbank Lending Market"

Craig and Ma, 2022

- "A small subset of banks intermediate between all other banks that almost exclusively lend or borrow and that are not directly connected with each other."
- Why the interbank intermediation arrangement arises? Quantify systemic risks in the interbank lending network...
- Model: a set of borrowers and lenders connected through intermediaries $\rightarrow$ delegated monitoring (Diamond, 1984). Lenders perform costly state verification over (partially) diversified intermediaries (Townsend, 1979)
- A borrower forming links to intermediaries: extra monitoring cost (k), but more bargaining power
- An intermediary forming links with borrowers: extra cost for monitoring, but better diversified$\rightarrow$ less premium for CSV (c)

Example 5: "Intermediation in the Interbank Lending Market"

Craig and Ma, 2022

* Take the observed network as the equilibrium outcome and find parameter values that best "rationalize" the observed existing/non-existing links

1. Take a parameter set $\{k, c\}$
2. For any observed/unobserved link, check whether the borrower has incentive to drop/form the link
3. Count the number of "correct" predictions: the score $Q(k, c)$
4. Search over parameter space to obtains the maximum score!

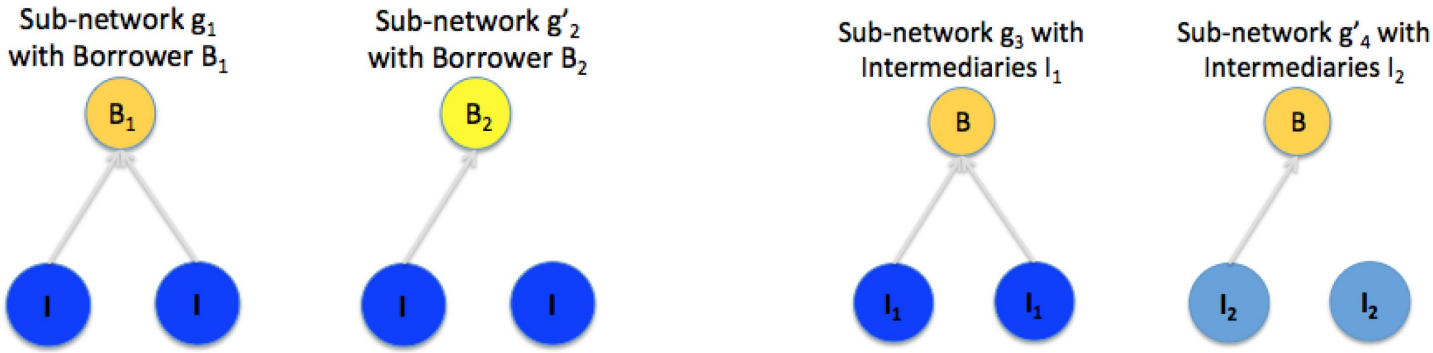
- Consistent estimation of CSV & monitoring cost $\{\hat{k}, \hat{c}\} = \max Q(c, k)$
- Standard errors are obtained by bootstrap method

Example 5: "Intermediation in the Interbank Lending Market"

Craig and Ma, 2022

The sources of variations that identify model parameters:

1. # of connections to intermediaries formed by a borrowing bank identifies monitoring cost relative to $\underbrace{\text{bank's profitability}}_{\text{directly estimated}}$.
2. # of connections to borrowers formed by an intermediary bank identifies CSV relative to monitoring cost.



Further notes: "Identification is not causality, and vice versa"

Kahn and Whited, 2017, Review of Corporate Finance Studies

- "Exogenous" variation is:
  - Always necessary to identify a *causal* relation
  - Never sufficient for identifying an economically interesting parameter
    - You need an economic model (either mathematical or verbal)
  - Only sometimes necessary to identify an economically interesting parameter
- Parameters may be identified *without* exogenous variation!

Further notes: Identification and Endogeneity

- Endogeneity is not necessarily a problem in structural estimation
- Structural estimation accounts for and exploits endogeneity within the model
  - economic model would tell you what the relationship between *endogenous* variables ought to be, given the parameter values
  - tune the economic parameter to get the same relationship between *endog.* vars as in data
  - remember: we can learn from *wrong* regressions
- Note: we rely on the economic model to specify the relationship between endogenous vars
- An important, common criticism: "The economic model omits an important aspect of reality."
  - Do a lot of robustness checks on model specification

Further notes: Identification with Exogeneity

- Exogenous variations can be helpful in structural work
- If an (unimportant) part of the model is specified via regressions
  → A *hybrid* approach.
- Or in identification: compare causal relationship in data with the same simulated regression result in model
  → Both regression in data/model may have omitted variable and economic force and have "endogeneity problems" → apple-to-apple comparison

Identification: Concluding remarks

- No one cares about estimation details/method, as long as you convince readers that the model is *well-identified*...
- Pick some moments/data patterns that are *sensitive* to economic parameter
- If not, you will get large standard errors, so the issue is detected. But not always in practice...
- Do a lot of comparative statistics, think a lot on the choice of moments, rely on economic intuition, read papers in the literature to get identification idea...
- Recall: you can exploit relationship between endogenous variables (identification is not causality and vice versa)
- Finally, make sure the model is identified, *before* getting to data!

3. “Light-speed” structural estimation

Robustness Checks

- **The computational burden**
  Standard practices in reduced-form work:
  - Alternative specifications
  - Sub-sample analyses
  - Alternative mechanisms

  Not much in structural works...
- **Opaque identification**
  - No clear understanding of the mapping between inputs (empirical moments) and findings (parameter estimates/policy recommendations)
  - No clear understanding of how moment selection matters for estimation (i.e. targeted vs. non-targeted moments)
- **Main practical challenge: computational burden**

Why there is a challenge?

- Estimating an economic model requires an *inversion*
  - we can simulate $m(p)$ in a few seconds/minutes,
  - but estimation needs finding $\hat{p}$ that solves $m(\hat{p}) = \hat{m}$
- We wish we knew the function $p = g(m)$, $g := m(\cdot)^{-1}$
- Note: Inversion is needed in reduced-form work as well: $(X'X)^{-1}$, but it's easy and fast there
- Note: There was no issue if evaluating $m(\cdot)$ was fast$\rightarrow$ inversion is done in an outer loop iterating over possible p
- Note: the issue is more severe, if a large number of parameters (unknowns) are to be estimated
  - time to co nverge might go crazy even with $n_p \sim 5$

$\rightarrow$ limited economic forces can be modeled

$\rightarrow$ even a slight modification would keep your PC busy for a week

Local approximation

Measuring the Sensitivity of Parameters to Moments, Andrews, Gentzkow, and Shapiro, 2017, The Quarterly Journal of Economics

- A solution to this challenge: linear (first-order) approximation

$$\delta\hat{p} = -(J'WJ)^{-1}J'W * \delta\hat{m}$$

- Limitations:
  - works locally
  - works only in the *intensive* margin

"Robustness Checks in Structural Analysis"

Catherine, Ebrahimian, Sraer, Thesmar, 2022

- **An alternative Solution:**

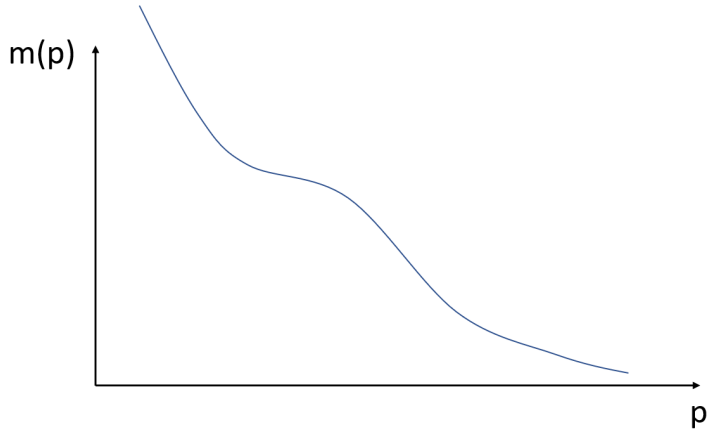
  "Light-speed" structural estimation, allowing for extensive robustness analysis in structural work

- **Key insight:**
  - Do not spend your computational budget on finding $\hat{p}$
  - Spend your budget on finding a good approximation of $m(\cdot) \to \tilde{m}(\cdot)$
  - Then use the approximation of m to run as many SMMs as you wish

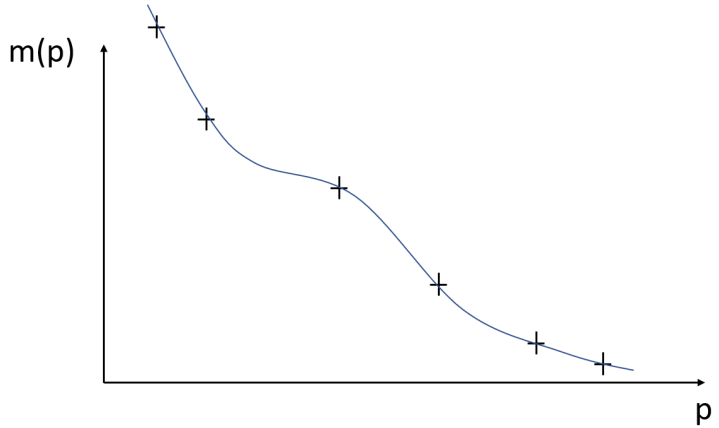
$$\min_p (\tilde{m}(p) - \hat{m})' W (\tilde{m}(p) - \hat{m})$$

How it works in practice

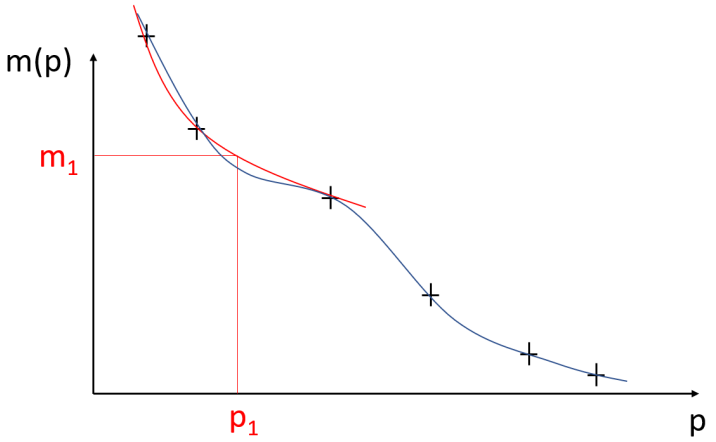
The exact relationship $m(p)$:



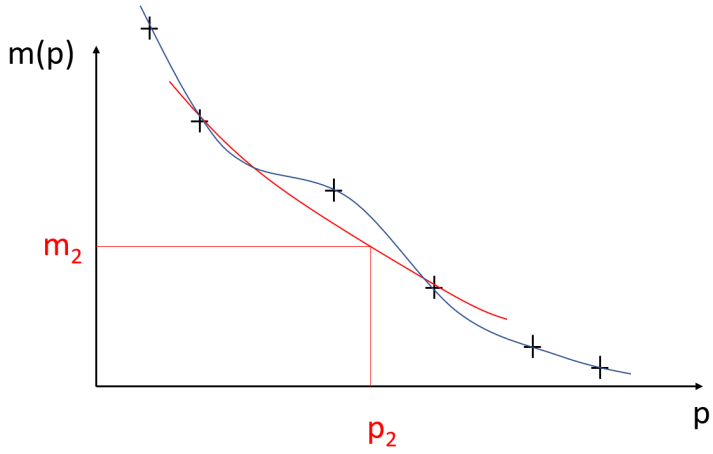
Random parameter draws p:



Approximate SMM via m_1:



Approximate SMM via m_2:



What we are approximating

- **The approximation $\tilde{m}(p)$ is <u>not</u> an approximation of the economic model**

  Relationships between state variables and the agent's decisions are generally not well-behaved because of inaction bands (eg transaction costs), discrete choices (eg renting/owning), corner solutions (eg collateral constraints), institutions (eg policy eligibility threshold) ...

- **The approximation $\tilde{m}(p)$ is a set of equations predicting moments as a function of parameters**

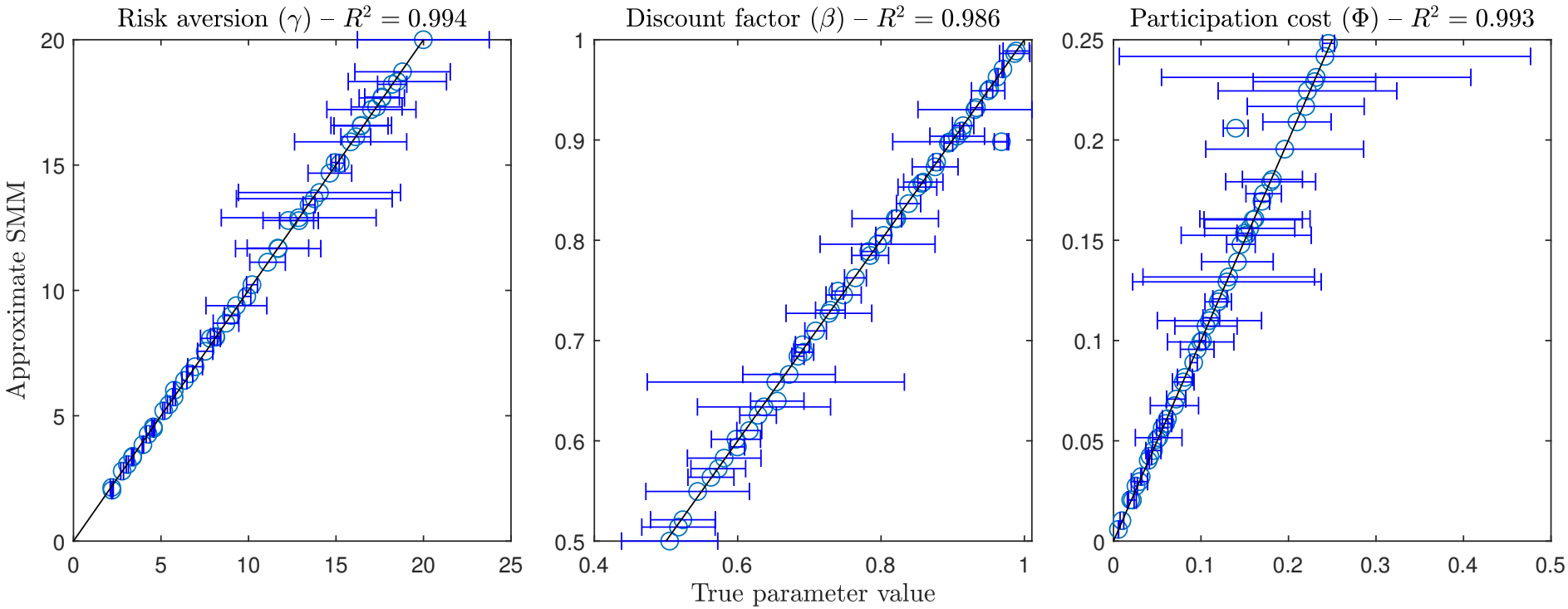
  Each moment is generally a well-behaved function of parameters, a condition for the method of moments to be valid

Summary — "Approximate SMM" stages

- Build a training set:
  - set a reasonably wide range for parameters
  - draw (semi-) random points in this space of parameters (Halton?)
  - use the true economic model to build the set of moments
  - store the set of parameters and moments on disk
- Generate the approximate moment function:
  - make a fit for each of the targeted moments based on input parameters
  - higher order polynomial, (deep) neural net, kernel smoothing, etc.
  - weight observations in the fit based on proximity to *data* target
  - sometime better to make fits over a *transform* of parameters/moments
  - note: evaluate your fit using an out-of-sample set before moving on
- Do SMM with the approximate moment function
  - use generous global optimization routines
  - once you converge, you can further iterate locally with true SMM

Application: Life-cycle portfolio choice model

Out-of-sample performance of approximate SMM:

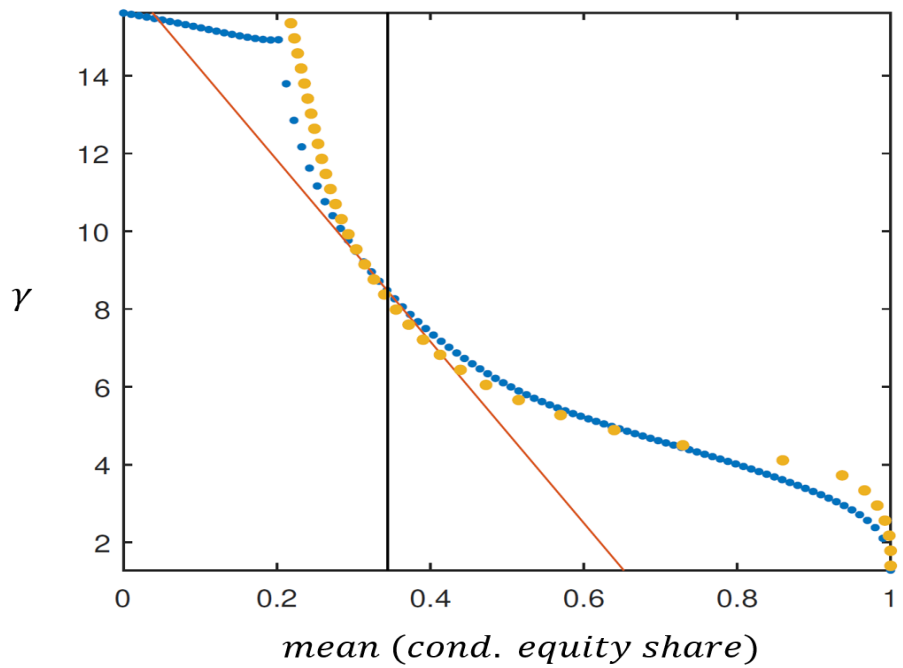


Illustrating identification (1)

- How do moments affect parameter estimation? often opaque in structural estimation
- Best practice: report local comparative statics. How simulated moments vary for local variations of θ around estimation: $p \rightarrow m(p)$ around $\hat{p}$
- With the approximate approach, easy to show how estimated parameters depend on empirical moments: $\hat{m} \rightarrow \hat{p}$
  - (1) fix all moments but one (2) vary moment around empirical value (3) re-estimate parameters
  - computationally, this approach is prohibitively expensive without an approximation

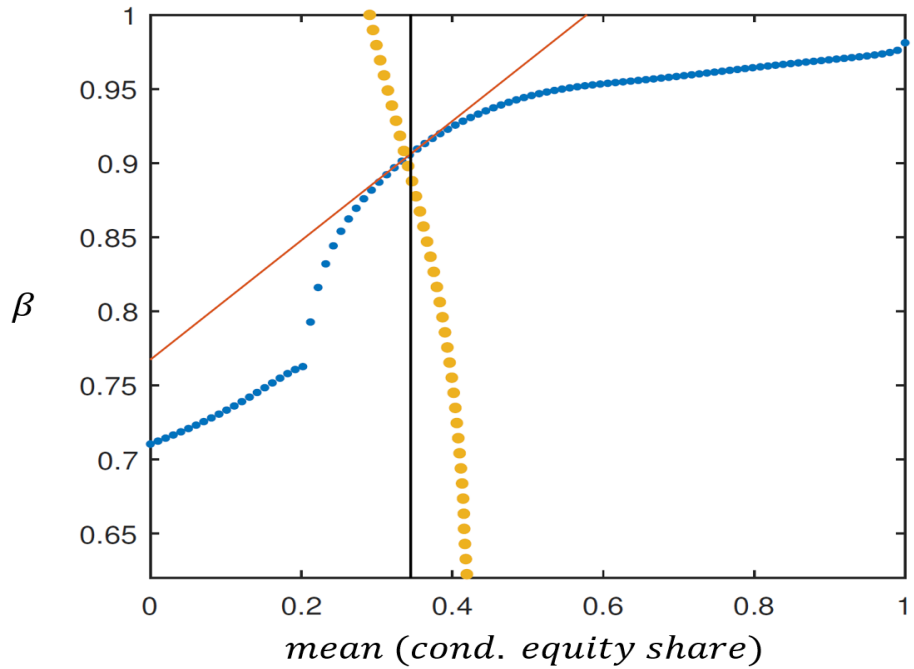
Identification: how does m_2 =mean cond. equity share affect risk aversion γ?

Orange: $p \to m(p)$ – **"standard" identification graph**
Blue: $m \to \hat{p}(m)$ – **"each obs." is a new estimation**



$\Rightarrow$ Similar conclusion from two approaches

Identification: how does m_2 =mean cond. equity share affect time discount factor β?



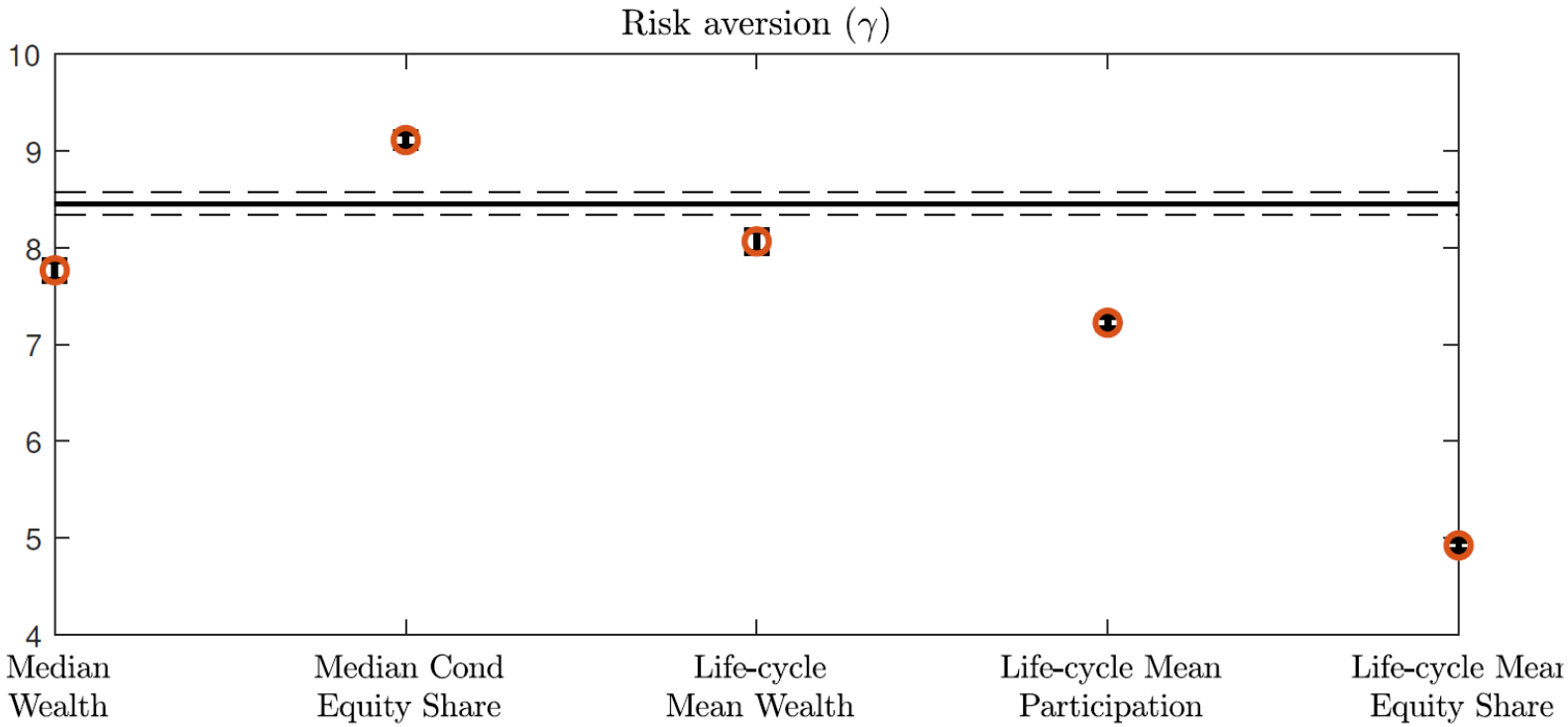
$\Rightarrow$ "Standard" identification graph says m_2 not useful for β:

- most value of β leads to similar m_2 close to empirical value
- **not true**: variations in $m_2 \rightarrow$ changes in $\hat{\gamma} \rightarrow$ changes in other moments $\rightarrow$ changes in estimated β
- **note**: local linear approximation (AGS) not accurate

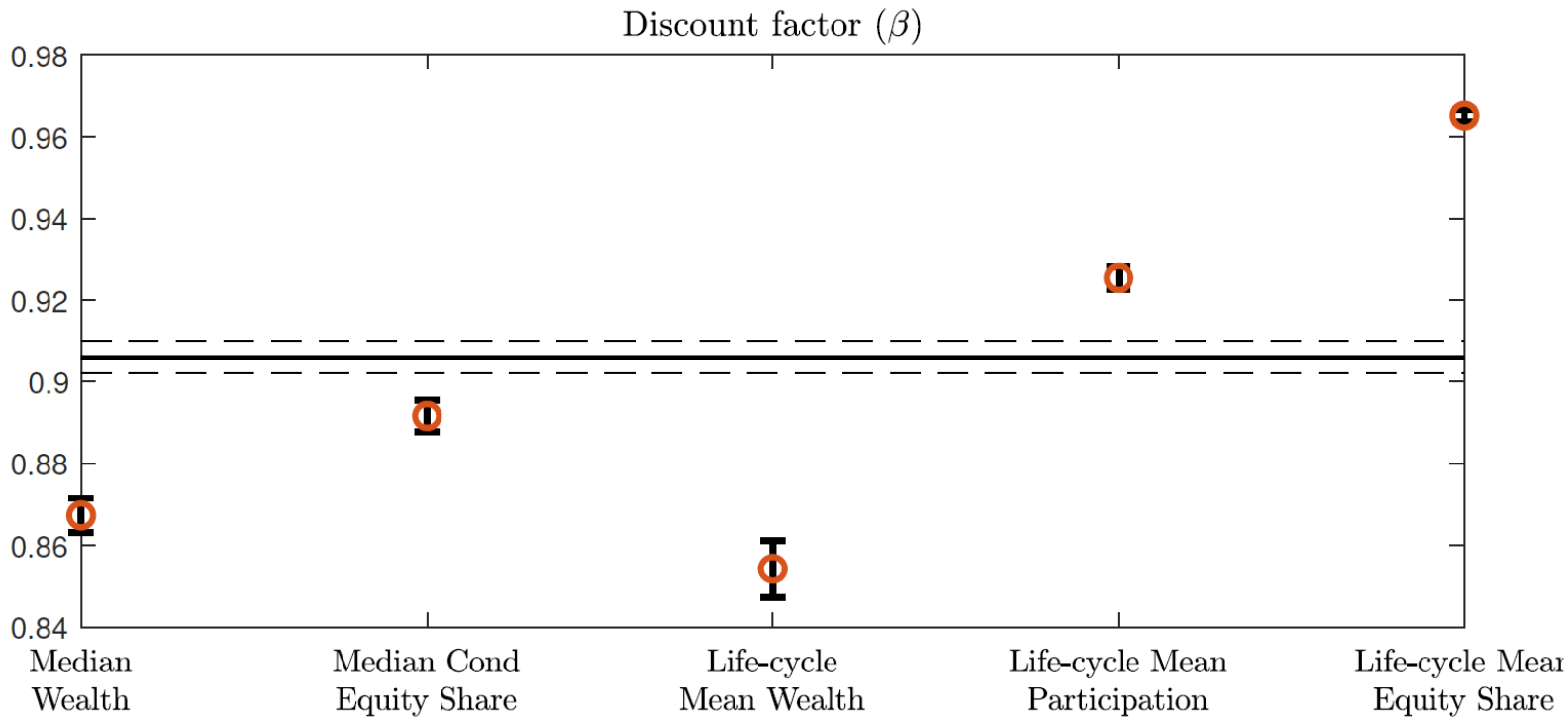
What about moment selection?

- Standard critique of structural approach: difficult to evaluate the importance of moment selection for estimation
- Let's consider the dozen of moments used in this literature. How do estimates and findings depend on moment selection?
- To do this, need to re-estimate the model with different combinations of moments

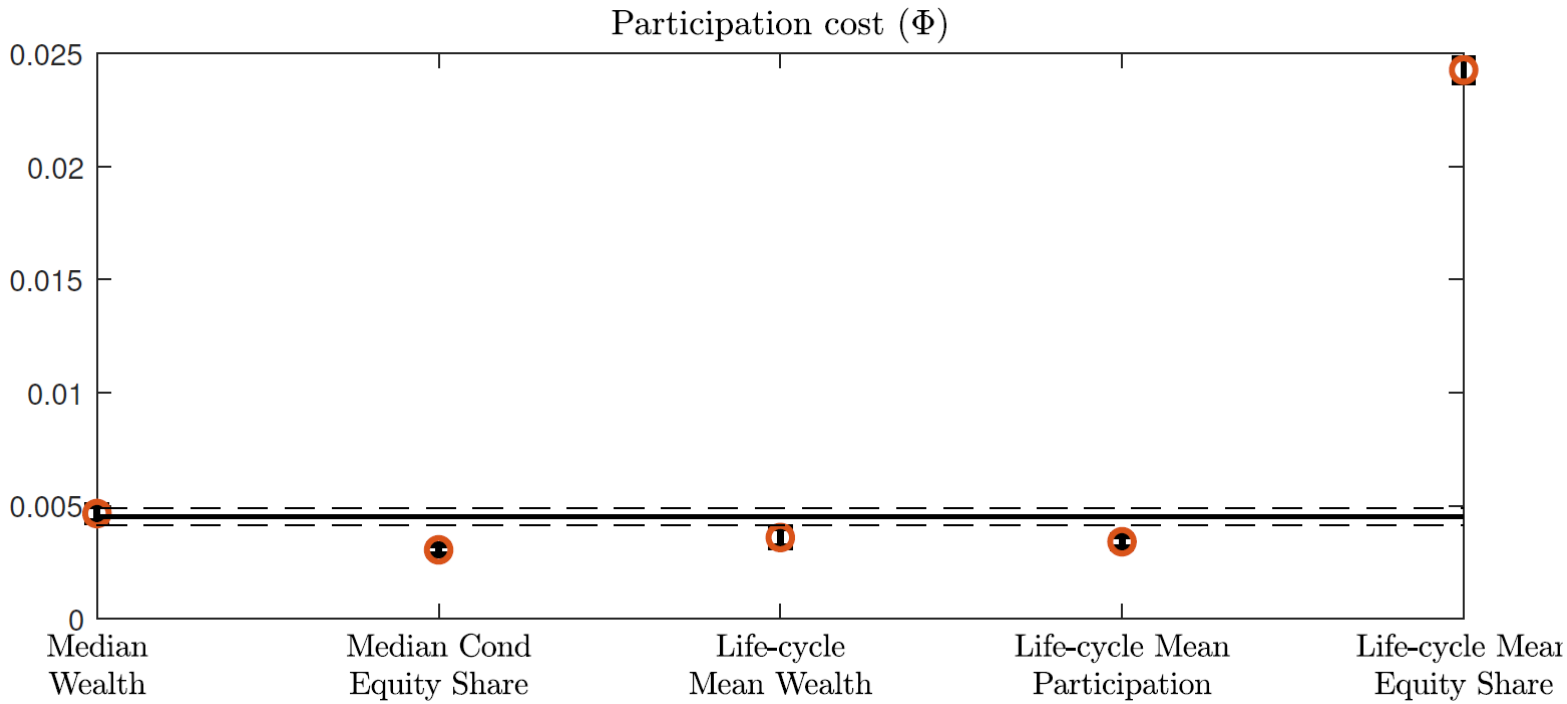
What about moment selection? Life-cycle portfolio choice model



What about moment selection? Life-cycle portfolio choice model

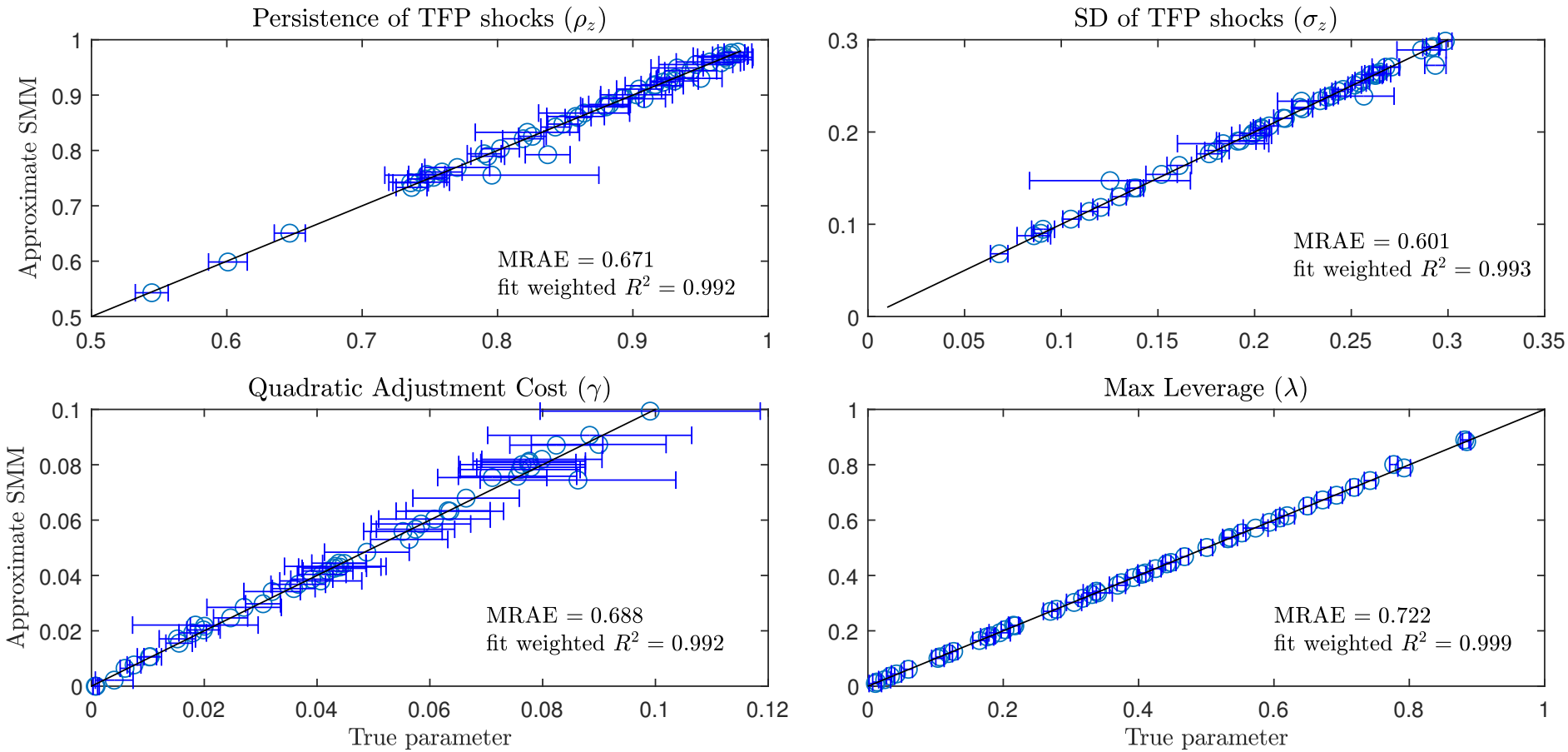


What about moment selection? Life-cycle portfolio choice model



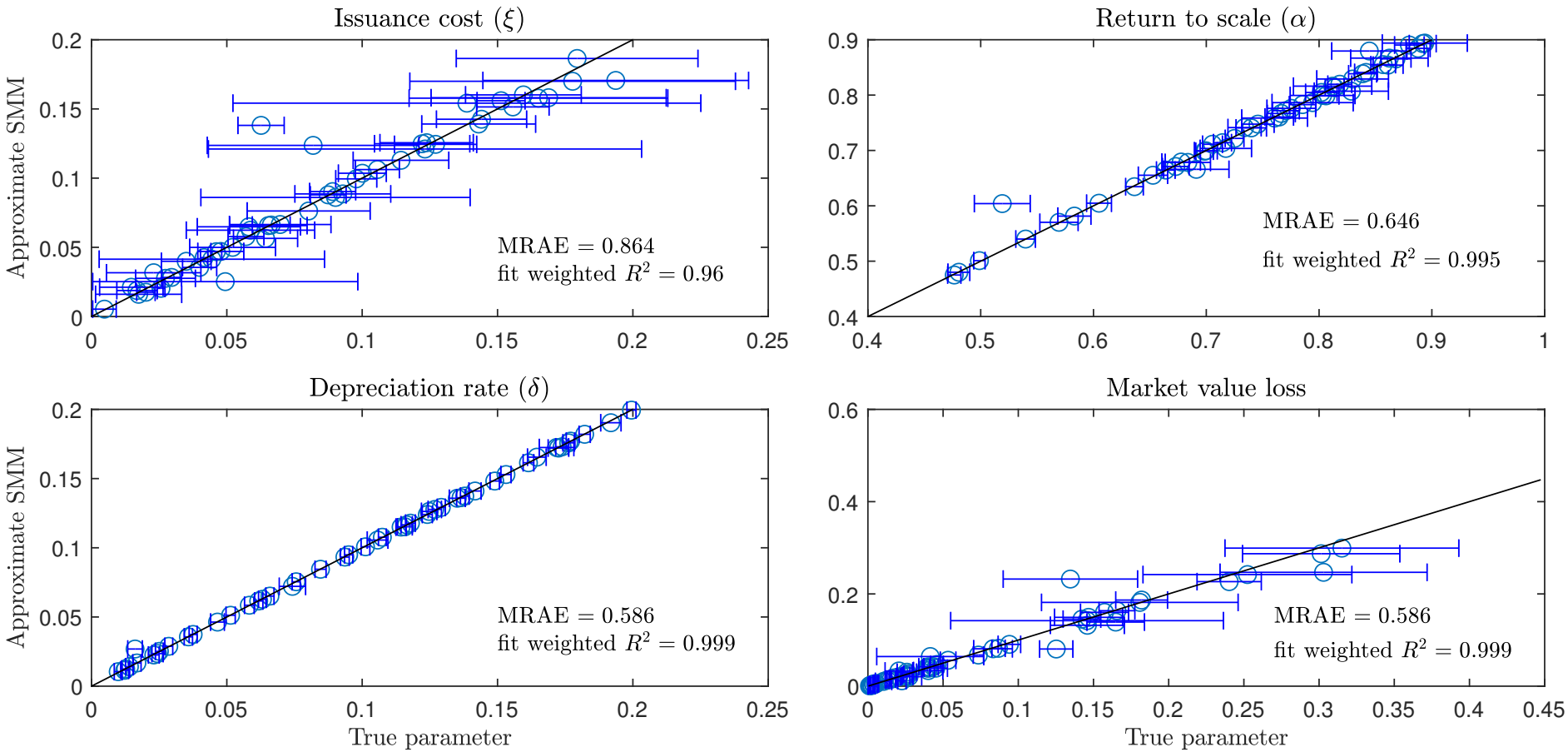
Application: Corporate Finance Model

Out-of-sample performance of approximate SMM (I):



Application: Corporate Finance Model

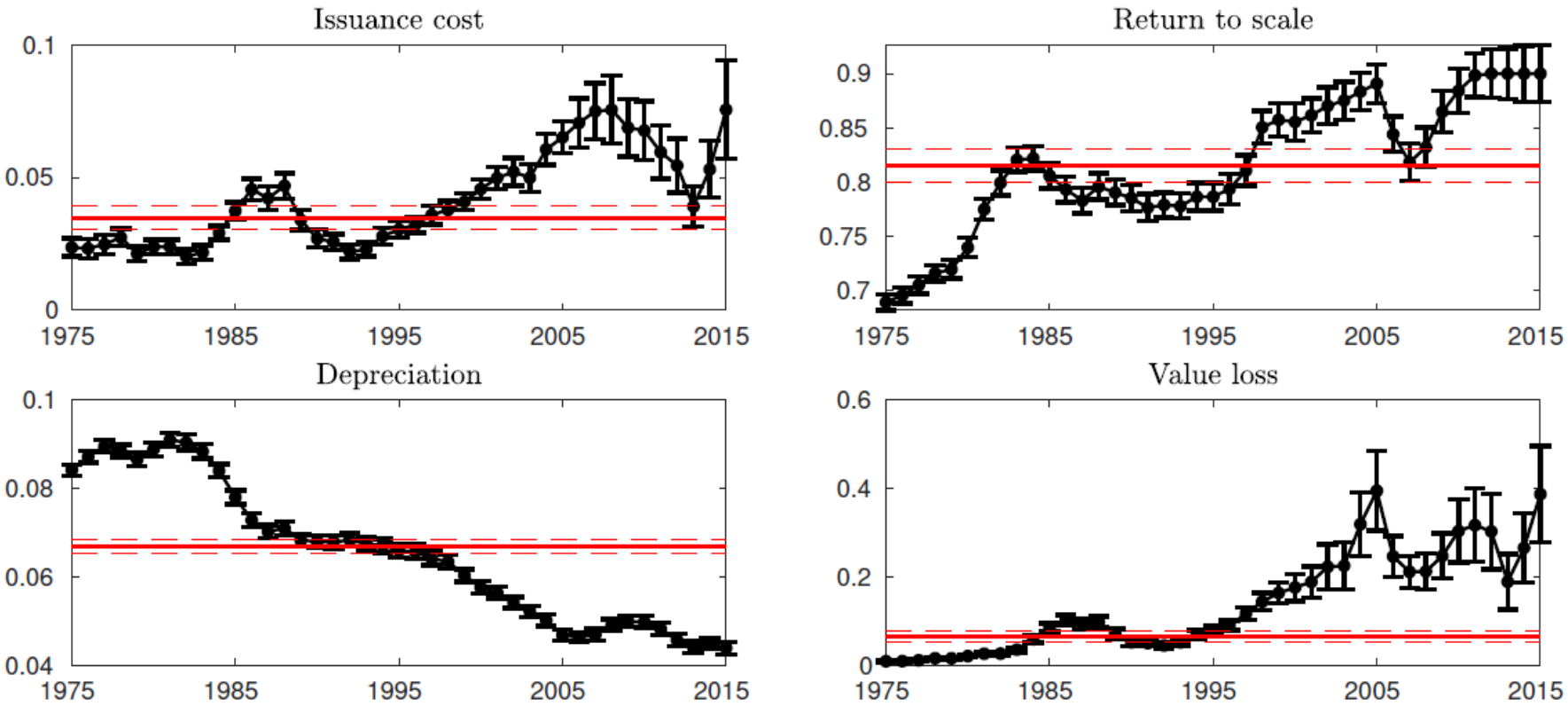
Out-of-sample performance of approximate SMM (II):



Standard robustness check: subsample analyses

Corporate Finance Model (I):

- Estimate the model for 10-year rolling sub-periods.



Standard robustness check: sample splits

Corporate Finance Model (II):

- Estimate the model for various industries:

	ρ_z	σ_z	γ	λ	ξ	α	δ	Value Loss
benchmark (all)	.7130	.2141	.0423	.1085	.0348	.8149	.0670	.0650
- std	.0067	.0018	.0025	.0032	.0022	.0076	.0008	.0062
manufacturing	.6921	.2102	.0403	.0730	.0115	.7512	.0542	.0146
- std	.0089	.0024	.0037	.0039	.0026	.0092	.0007	.0035
retail trade	.8637	.1347	.0846	.1870	.1696	.8295	.0939	.1174
- std	.0140	.0042	.0108	.0096	.0490	.0145	.0022	.0250
services	.7995	.2161	.0277	.0563	.0635	.7488	.0575	.1130
- std	.0166	.0044	.0045	.0096	.0077	.0176	.0020	.0150
transportation	.7374	.1673	.0390	.3102	.0893	.9000	.0955	.2246
- std	.0253	.0075	.0072	.0139	.0122	.0238	.0037	.0634
mining	.6945	.2575	.0194	.3776	.1218	.9000	.1088	.4835
- std	.0362	.0170	.0036	.0181	.0110	.0363	.0059	.1910

Questions?

Have fun doing structural estimation!

References I

Berry, S., J. Levinsohn, and A. Pakes (1995). Automobile prices in market equilibrium. *Econometrica 63*(4), 841–890.

Berry, S., J. Levinsohn, and A. Pakes (2004). Differentiated products demand systems from a combination of micro and macro data: The new car market. *Journal of political Economy 112*(1), 68–105.

Catherine, S. (2022). Countercyclical labor income risk and portfolio choices over the life cycle. *The Review of Financial Studies 35*(9), 4016–4054.

Craig, B. and Y. Ma (2022). Intermediation in the interbank lending market. *Journal of Financial Economics 145*(2), 179–207.

Diamond, D. W. (1984). Financial intermediation and delegated monitoring. *The review of economic studies 51*(3), 393–414.

Ebrahimian, M. (2022). Student loans and social mobility. *Jacobs Levy Equity Management Center for Quantitative Financial Research Paper*.

Hennessy, C. A. and T. M. Whited (2007). How costly is external financing? evidence from a structural estimation. *The Journal of Finance 62*(4), 1705–1745.

References II

Kahn, R. and T. M. Whited (2018). Identification is not causality, and vice versa. *Review of Corporate Finance Studies 7*(1), 1–21.

Keane, M. "practical issues in structural estimation". *https://www.youtube.com/watch?v=0hazaPBAYWE*.

Townsend, R. M. (1979). Optimal contracts and competitive markets with costly state verification. *Journal of Economic theory 21*(2), 265–293.