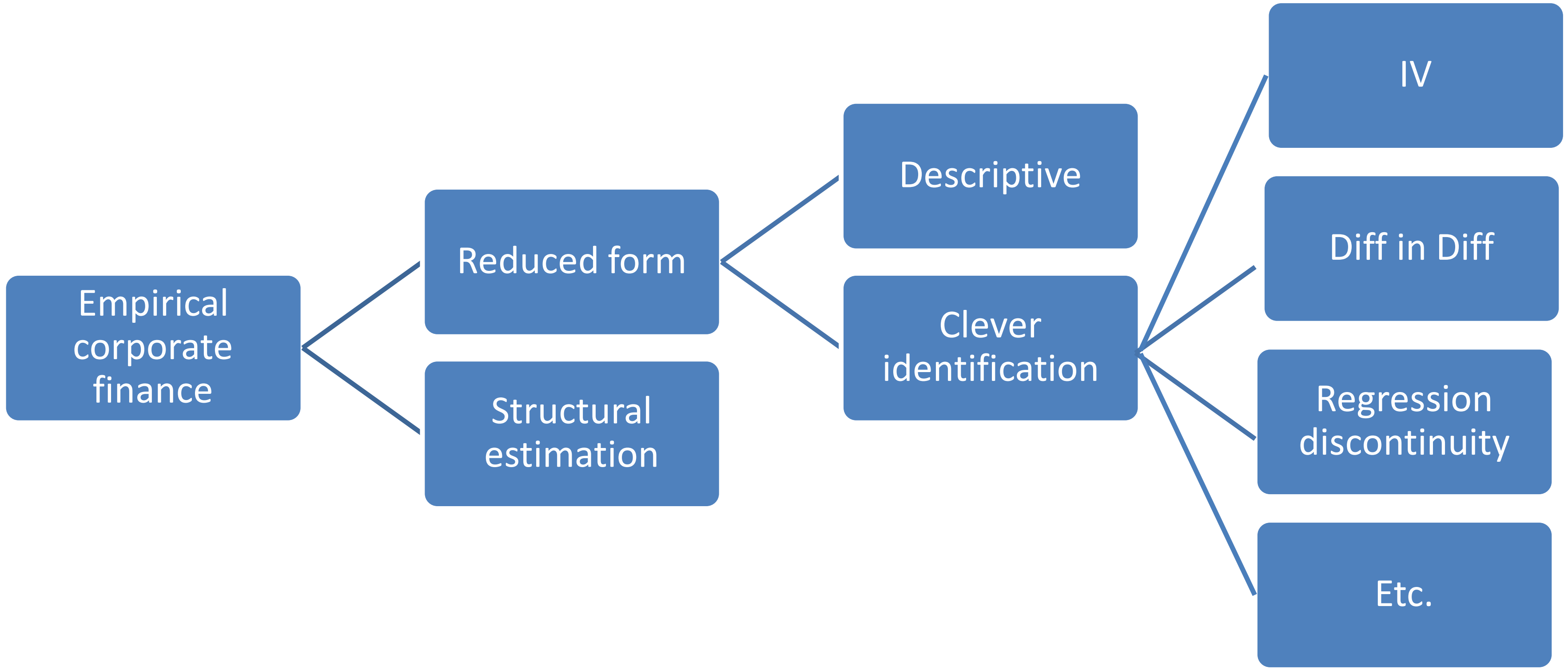
Introduction to Structural Estimation in Corporate Finance

Luke Taylor

April 2017

THE BIG PICTURE



PLAN

- **What is structural estimation?**
  - Terminology
  - An example
  - 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 estimation
  - Is structural estimation good for your career?
- **How to do it**
  - How to generate research ideas
  - How to do SMM
  - Which estimator should you use?
  - Identification
  - Tricks of the trade / warnings
- **How to referee a structural estimation paper**
- **Where to learn more**

FIRST, SOME TERMINOLOGY

- It makes no sense to say “structural model”
- All economic models are “structural”
- Usually when people say “structural model,” they really mean “dynamic model”
- It makes a lot of sense to talk about “structural estimation” versus “reduced-form estimation”

STATISTICAL AND ECONOMIC MODELS

- A **statistical model** describes the relation between two or more random variables:

$$Y=X'b+e$$

- An **economic model** starts with assumptions about
  - Agents' preferences
  - Constraints
  - Information environment
  - Firms' production functions
  - Some notion of equilibrium, etc.

- Then it makes predictions about the relation between observable, often endogenous variables

WHAT IS STRUCTURAL ESTIMATION?

- **Structural estimation** is an attempt to estimate an **economic model's** parameters and assess model fit.

- Parameters to estimate often include

  - Preference parameters (e.g., risk aversion coefficient)

  - Technology parameters (e.g. production function's curvature)

  - Other time-invariant institutional features
    (e.g. agents' bargaining power, financing frictions)

- Structural estimation ascertains whether optimal decisions implied by a model resemble actual decisions made by firms

EXAMPLE: “DYNAMIC DEBT RUNS...” from 2014 JFE

Economic model:

- **Setting**: Continuous time, 1 borrowing firm, continuum of lenders
- **Production function**:
  Asset value follows geometric Brownian motion
- **Financing**: Firm buys an asset by issuing equity & short-term debt
- **Preferences**: Risk-neutral lenders optimally choose whether to roll over debt or “run”
- **Information**: a lender’s decision depends on beliefs about other lenders’ decisions (strategic complementarity)
- **Equilibrium**: debt is priced in competitive market

EXAMPLE: “DYNAMIC DEBT RUNS…”

Parameters to estimate:

1. Volatility for asset’s Brownian motion
2. Drift “ “ “ “ *
3. Average debt maturity
4. Average asset maturity
5. Perceived weakness of firm’s backup credit guarantee
6. Asset’s liquidity = recovery rate in default
7. Cap on yield spreads
8. Investors’ discount rate

* Drift is not identified. We assume a value, use alternate values in robustness section.

EXAMPLE: "DYNAMIC DEBT RUNS…"

Data:

- Panel data on firms issuing ABCP (short-term debt) in 2007
- Variables:
    - Weekly spreads (i.e. prices) on ABCP
    - Indicator for whether firm is experiencing a run

EXAMPLE: “DYNAMIC DEBT RUNS…”

Assessing model fit: How well does model fit

- Frequency and timing of “recoveries” from runs
- Average debt yields in event time leading up to runs
- Yield volatility and its relation to yield levels
- Probability of future run, given current yield level (forecasting regression)

WHAT KIND OF MODEL I LIKE TO USE

- The model has to be an economic rather than statistical model
- Should produce realistic magnitudes and distributions
  - No two-state, “profits-are-either-high-or-low” models
  - Structural estimation may or may not require a dynamic– as opposed to static– model
    - Hennessy and Whited (2005, JF) → Dynamic
    - Albuquerque and Schroth (2010, JFE) → Static
  - But usually no two- or three-period models

CALIBRATION VS. STRUCTURAL ESTIMATION

Calibration	Structural estimation
• Take parameter values from other papers	• Infer parameter values from the data
• Usually have more parameters than moments → model isn't identified, can't put standard errors on parameters	• Get standard errors for parameters
• <u>Mainly a theoretical exercise</u>	• <u>An empirical exercise</u>

Both:

- Can assess how well model fits the data– but no statistical tests with calibration
- Can use model to ask counterfactual questions:
  - What would happen if we shocked this variable?
  - How would world look if we changed this parameter's value?

STRUCTURAL VS. REDUCED-FORM ESTIMATION

	Reduced-form	**Structural estimation**
Questions	What is the (causal) effect of X on Y?	• Why does X causes Y? • What are the parameters' magnitudes? ["Parameters" = economic primitives "Parameters" ≠ slopes, correlations] • How well does theory fit the data? • How would the world look if one of the parameters (counterfactually) changed? • What would happen if you (counterfactually) shocked the system

STRUCTURAL VS. REDUCED-FORM ESTIMATION

	Reduced-form	Structural
Tools	Estimators: • OLS • IV • Diff-in-diff • Regression discontinuity Software: Stata	Solving the model: • Value function iteration • ODE/PDE solvers • Simulation Estimators: • GMM • Simulated method of moments (SMM) • Maximum likelihood (ML) • Simulated maximum likelihood (SML) • [See Streb. and Whited (2012) for more] Software: Matlab, C++, Julia, Fortran, etc.

STRUCTURAL VS. REDUCED-FORM (TERMINOLOGY)

- Economic models often imply a "reduced-form," meaning a statistical model describing the relation between observables generated by the model.

- Example from "Why are CEOs rarely fired?"
  One reduced-form prediction from the structural model:

$$1(CEO\ fired)_{it} = \beta_{0\tau} + \beta_{1\tau}\pi_{it} + \beta_{2\tau}\pi_{it-1} + \ldots + \varepsilon_{it}$$

  The regression slopes β are nonlinear functions of the model's structural parameters.

  The true reduced-form may actually be nonlinear in π.

PLAN

- What is it?
- **Why do it?**
  - What structural estimation buys you
  - How to motivate a structural estimation paper
  - Advantages and disadvantages vs. reduced-form
  - Is structural estimation good for your career?
- How to do it
- How to referee a structural estimation paper
- Where to learn more

STRUCTURAL ESTIMATION BUYS YOU THREE THINGS

From least to most interesting:

1. Estimates of interesting economic primitives

2. "Deep" tests of theory:
   - Formal, joint tests of multiple predictions (e.g., test of overidentifying restrictions in GMM or SMM)
   - Testing quantitative, not just directional predictions
   - Seeing where models fail opens doors to future research (Example: Mehra and Prescott (1985), equity premium puzzle)

3. Can answer interesting counterfactual questions

Caveat: 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, because it's harder to believe "all else equal." Also, it's impossible to shock primitives in reduced-form papers….

EXAMPLE: "WHY ARE CEOS RARELY FIRED?..." from 2010 JF

From least to most interesting:

1. Estimates of interesting economic primitives:
   I estimate a parameter that quantifies CEO entrenchment:
   Directors' disutility from firing a CEO

2. "Deep" tests of theory:
   Model does a good job fitting most moments but struggles to fit
   (1) changes in profitability in the year after CEOs fired, and
   (2) high rate at which CEOs are fired in their first 2 years in office

3. Can answer interesting counterfactual questions:
   How much would firm value change if we eliminated CEO entrenchment?
   - Set the entrenchment parameter to zero→
     Firm value increases by 3%.

EXAMPLE: "DYNAMIC DEBT RUNS…."

1. Estimates of interesting parameters:
   Not so interesting in this paper

2. "Deep" tests of theory:
   Model does a good job fitting most moments but, in one subsample, overpredicts runs when yields are high.

3. Can answer interesting counterfactual questions:
   How does the probability of a run react to a (counterfactual)
   - Equity injection:
     Reducing leverage by 1% lowers Pr{run} by 45%
   - Improvement in asset liquidity
   - Reduction in asset volatility
   - Strengthening of backup credit guarantees
   - Longer debt maturity or shorter asset maturity

MOTIVATING A STRUCTURAL PAPER

- Structural estimation imposes large costs on the reader
- → Any structural paper must put great effort into convincing reader that it's worth going structural
- Next slide: an example

EXAMPLE: "DYNAMIC DEBT RUNS…"

Question: How sensitive are runs to their various potential determinants?

	Reduced-form estimation	**Structural estimation**
Approach	Regress 1(run) on determinants of runs (leverage, liquidity, volatility, guarantee strength…)	• Estimate structural parameters by SMM • Use counterfactual analysis to measure sensitivity of runs to determinants
Data challenges	• Tough to get data on leverage, liquidity, assets' value, assets' volatility, guarantee strength… • Need sufficient heterogeneity in determinants	• Estimate these quantities structurally from data on prices, runs, and recoveries • Do <u>not</u> need heterogeneity in determinants
Identifying assumptions	• Exogenous variation in determinants (i.e., regression does not omit any correlated determinants of runs) • Got the functional form right	• Model is true: - Includes all determinants of runs - Rational investors - Functional forms are correct

The structural approach complements existing reduced-form research by

(1) overcoming certain data challenges

(2) imposing a different type of identifying assumption.

STRUCTURAL VS. REDUCED-FORM ESTIMATION

	Reduced-form	Structural estimation
Advantages	• "Fewer" assumptions? Results more convincing? • Easier to do • Easier to understand→ larger audience	• Often the only feasible option for answering certain important questions • Tough to find good instruments • The connection between theory and tests of theory is extremely tight, which allows more transparent interpretation of any results. Structural estimation "puts the theory first" and makes it explicit. • Results generalize better? • For job market: Makes you look smart

Bottom line:

- 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!!

PLAN

- What is it?
- Why do it?
- **How to do it**
  - How to generate research ideas
  - How to do SMM
  - Which estimator should you use?
  - Identification
  - Tricks of the trade / warnings
- How to referee a structural estimation paper
- Where to learn more

HOW TO DO SMM (1 of 2)

(Modified from Strebulaev and Whited (2012))

First steps:

1. Choose moments to match. Can include means, variances, covariances, regression slopes, etc. Need at least as many moments as parameters. Extra moments provide a test of overidentifying restrictions.
2. Compute moments in actual data, stack them in a vector *M*
3. Estimate the covariance of *M.* Invert it. This is your efficient SMM/GMM weighting matrix, *W.*

Second steps:

1. Pick β_0 = starting guess for the parameter vector
2. Using $\beta_{0,}$ solve model, create simulated data using policy function, calculate same moments that were calculated with real data. Stack them in vector $m(\beta_0)$. Note simulated moments *m* depend on parameter values β_0.
3. Compute the SMM objective function as

$$Q(\beta_0) = (M - m(\beta_0))' \, W \, (M - m(\beta_0))$$

4. Find the parameter vector $\hat{\beta}$ that minimizes *Q.* This is your parameter estimate.

HOW TO DO SMM (2 of 2)

5. Compute parameters' standard errors while adjusting for simulation error.

   N = number of observations in actual data

   J = number of simulated observations / *N*

   Covariance matrix of $\sqrt{N}(\hat{\beta} - \beta)$ is

$$\left(1 + \frac{1}{J}\right)\left(\frac{\partial m(\beta)'}{\partial \beta} W \frac{\partial m(\beta)}{\partial \beta}\right)^{-1}$$

6. Compute the test of overidentifying restrictions, which tests whether the model jointly matches all moments. The test statistic is

$$\frac{NJ}{1+J} Q(\beta) \rightarrow \chi(\#moments - \#parameters)$$

7. Optional but interesting: counterfactual exercises. Change one or more parameters from their estimated values to counterfactual values, then examine how model's predictions change. Shock one or more exogenous variables in the model….

WHICH ESTIMATOR SHOULD YOU USE?

Estimator	Pros / cons
GMM	- Need closed-form solution + Fast
SMM	+ Don't need closed-form solutions - Extremely slow (use parallel computing as much as possible) + Can use "complicated" moments, sample the data in strange ways....
GMM & SMM	- Choice of moments is subjective and arbitrary (sometimes a +) + *Semi*parametric: Does not require a complete specification of the probability distribution of the data + Have control over weights put on each moment + Delivers a test of over-identifying restrictions

WHICH ESTIMATOR SHOULD YOU USE?

Estimator	Pros / cons
Maximum likelihood (ML)	+ Fast + Asymptotically efficient: consistent, asymptotically normal, "smallest standard errors" - Need closed-form solutions + Don't need to subjectively choose moments +/- "Uses all the moments" predicted by the model - Fully parametric
Simulated maximum likelihood (SML)	[All the same pros / cons as ML, except slower than ML] + Easy to accommodate heterogeneity in parameter values

WHICH ESTIMATOR SHOULD YOU USE?

Estimator	Pros / cons
Markov chain Monte Carlo	+ Good at estimating non-linear models with many latent variables that require high-dimensional integration to evaluate the likelihood function + Good at handing hierarchical models + Good at handling missing data + Faster than SMM + Good small-sample properties See Arthur Korteweg's webpage for more info

WHICH ESTIMATOR SHOULD YOU USE?

Bottom line:

I don’t care much which estimator you use.

As long as the model is well identified, it should not matter much.

IDENTIFICATION: STRUCTURAL VS. REDUCED-FORM

- Identification is often confused with establishing causation

- Formal statistical definition of identification:
  - Econometrician defines an objective function over parameters and data
  - Goal: Select parameters that minimize this objective function
    - (E.g. Find regression slope that minimizes sum of squared errors)
  - A parameter is **identified** if there is a unique minimum for the objective function at its true value in the population

- A parameter can be identified (in the statistical sense) without being economically interesting
  - Prime example: Regression of endogenous Y on endogenous X

- Our goal: Identify parameters that are economically interesting
  - The parameters may be elasticities defining causal effects
  - But they need not be!
  - **Not all economically interesting parameters are causal elasticities**

Source: "Identification with models and exogenous data variation," by R. Jay Kahn and Toni M. Whited

IDENTIFICATION: STRUCTURAL VS. REDUCED-FORM

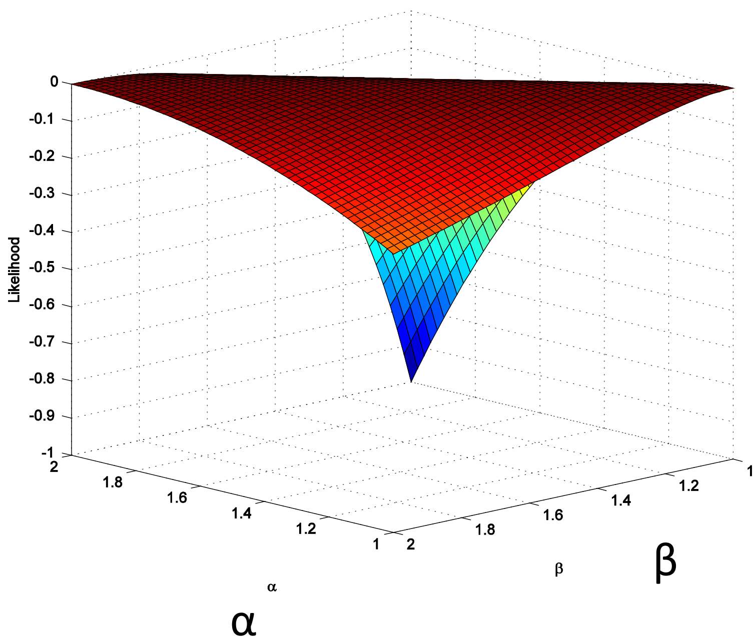
- Exogenous variation is:
  - Always necessary to identify a causal relation
  - Never sufficient for identifying an economically interesting parameter
    - You also need an economic model (either mathematical or verbal)
  - Only sometimes necessary to identify an economically interesting parameter
- Interesting parameters can sometimes be identified without exogenous variation. This is often what's going on in structural corporate finance.
- In what follows, I'll use the formal statistical definition of identification

Source: "Identification with models and exogenous data variation," by R. Jay Kahn and Toni M. Whited

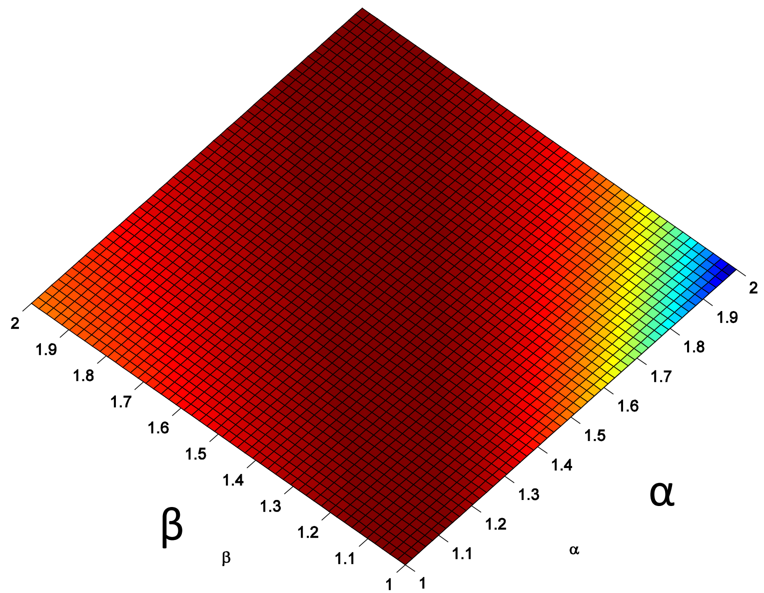
EXAMPLE OF AN UNIDENTIFIED MODEL (MLE)

- Want to estimate parameters α and β
- Parameters α and β appear in the likelihood function only in the form α/β
- The fraction α/β will be identified, but α and β will not be separately identified
- Likelihood function is flat at its max:

Likelihood function



Likelihood function from above



HOW TO CHOOSE MOMENTS IN SMM/GMM

Must choose moments with extreme care to obtain an identified model

- Best-case scenario: each moment depends on just 1 model parameter:
  "Moment 1 identifies parameter 1, moment 2 identifies parameter 2…"

- More realistic: every moment depends on every parameter

- Do comparative statics to understand how each moment moves with each parameter. Make sure you understand the economics behind each comparative static result.

- Need enough moments, and moments that move in in different directions for different parameters, to obtain identification.

- It's a good idea to use moments describing the policy function
  - Policy function: Mapping from state variables to choice variables
  - See Bazdresch, Kahn, and Whited (2014)

If the author cannot clearly explain which features of the data identify each parameter, the paper is not very convincing. Structural estimation should not be a black box.

IDENTIFICATION AND OMITTED VARIABLES

- "Just as there does not exist any perfectly exogenous source of data variation in observational studies, structural estimation does not magically solve all endogeneity problems." (Strebulaev and Whited, 2012)

- An important, common criticism:
  "The structural model has omitted an important aspect of reality."

- The model may be well identified in a technical, econometric sense, even though it omits this "variable"

- Even if the model is well identified, we may not want to take its results seriously if the omitted "variable" is important.

- Potential solutions:
  - Avoid using moments that are contaminated by omitted forces
  - Use moments that already "sweep out" the omitted variable
    (See Hennessy and Whited, 2005)
  - TONS of robustness exercises. Extend the model to include the omitted variable, re-estimate.

IDENTIFICATION – GENERAL REMARKS

- "Endogeneity" is not necessarily a problem here. Structural estimation accounts for and exploits endogeneity within the model to get parameter estimates.

- It's usually very difficult to prove whether a model is identified

- Two useful checks:

  1. If a parameter's standard error is huge, it's probably not identified

  2. (GMM or SMM) A necessary condition for *local* identification: The Jacobian of moments w.r.t. parameters, $\frac{\partial m(\widehat{\beta})}{\partial \beta}$, must have full rank

- It's okay to "calibrate" nuisance parameters that are hard to identify
  Example: In "Why are CEOs rarely fired?" I set discount factor to 0.9 and try alternative values for robustness.

IDENTIFICATION AND PARAMETER HETEROGENEITY

- Model is usually about a single firm
- Common identifying assumption:
  Parameter values are constant across all firms and years within the sample.
- It's analogous to assuming a regression slope is the same for all observations in a regression
- Sometimes interesting to ask how parameters vary across firms, years, etc.
- How to address / explore parameter heterogeneity:
  - Estimate the model in subsamples. Do results go the way you expected?
  - Use the method in Taylor (2012), "CEO wage dynamics…."
  - Before estimating, purge heterogeneity that's outside the model
    - E.g., remove firm and time fixed effects before measuring moments

PRACTICAL TIPS

General tips:

- Before going structural, convince yourself that a structural approach is absolutely necessary.

- Don't start estimating or gathering data until you're convinced the model is identified and you understand why/how

- A way to check whether model is identified:
  1. Simulate a "fake" dataset off the model
  2. Estimate the model, treating the fake data as if it were real data
  3. Does the estimator recover the true, known parameter values?
  4. Are your standard errors accurate? Repeat exercise several times.

     Across simulations, should find:

     Stdev(estimates) ≈ Average (standard error)

PRACTICAL TIPS

Tips for SMM/GMM:

- Switching moments is a huge pain. Think carefully about identification before coding/estimating

- On searching the parameter space to minimize the objective function
  - Use the simulated annealing (SA) algorithm to avoid local minima
  - Once SA converges, run a deterministic minimizer like Matlab's fminsearch.
  - Run SA from multiple initial parameter guesses in parallel on Wharton's computing grid.
  - Use the same seed for the random-number generator each time you simulate data off the model.

PRACTICAL TIPS

Tips for SMM/GMM:

- Get the standard errors right.
  - The actual data are usually not i.i.d.
  - When estimating the covariance matrix for empirical moments [*Cov(M)*], must take into account
    - Heteroskedasticity
    - Time-series autocorrelation
    - Cross-sectional correlation
    - Serial correlation, including correlation across moments.

GETTING THE STANDARD ERRORS RIGHT (CONTINUED)

- Example: Empirical moments are slopes β and γ from regressions

$$y_{it} = \beta x_{it} + \varepsilon_{it}$$
$$h_{it} = \gamma z_{it} + \delta_{it}$$

- Need to compute

$$cov(M) = \begin{bmatrix} var(\beta) & cov(\beta, \gamma) \\ cov(\gamma, \beta) & var(\gamma) \end{bmatrix}$$

- Must take into account

$corr(\varepsilon_{it}, \varepsilon_{is}) \neq 0$ [Time-series autocorrelation]

$corr(\varepsilon_{it}, \varepsilon_{jt}) \neq 0$ [Cross-sectional correlation]

$corr(\varepsilon_{it}, \delta_{js}) \neq 0$ [Time-series and cross-sectional correlation across regressions]

- How to do it? Estimate the empirical moments M as a big GMM system with appropriately robust, clustered errors. See Taylor (2010) and Taylor (2012). Or use influence functions, as in Erickson and Whited's papers (see my note, "How to compute the standard error for anything, using influence functions")

PLAN

- What is it?
- Why do it?
- How to do it
- **How to referee a structural estimation paper**
- Where to learn more

QUESTIONS A REFEREE MIGHT ASK

- ☑ Am I convinced that we need structural estimation?
  - Why won’t a reduced-form approach work?

- ☑ Is the economic question important?
  - Or are we using a large hammer to hit a small nail?

- ☑ Is the identification clear, or is it a black box?
  - Which features of the data identify each parameter, and why/how?

- ☑ Is model fitting the data reasonably well?
  - If not, what can we learn from its failure?
  - Usually not a deal-breaker

- ☑ Are moments contaminated by important omitted economic forces?
  - If so, how could the omission bias the estimates?

- ☑ Have authors explored interesting heterogeneity in the parameters?
  - E.g, estimate model in subsamples
  - Enriches paper, provides useful consistency checks

- ☑ Does the paper take full advantage of counterfactual exercises?

WHERE TO LEARN MORE

- These slides owe a large debt to Toni Whited.
- Other resources from Toni:
  - Her 2012 survey article with Strebulaev: “Dynamic models and structural estimation in corporate finance”
  - Her nontechnical lecture slides on structural estimation: http://toni.marginalq.com/fma2011.pdf
  - Her technical slides on structural estimation: http://toni.marginalq.com/FMA.pdf
- Take an empirical Industrial Organization (IO) course
- Read papers that do structural estimation (that’s how I learned!)
- 2017 Mitsui Summer School on Structural Estimation Corporate Finance

WHY GO STRUCTURAL? BECAUSE IT'S FUN!

Going structural may be right for you if…

- … you actually know what "robust" does in Stata
- … you asked Santa Claus for the newest version of Matlab
- … you left your last girlfriend/boyfriend for a Bellman equation
- … you'd rather clean data than clean your laundry
- … there's not much on your calendar for next few years
- … you're emotionally robust

It's easier than it looks. Go for it!