Andrews, Gentzkow, Shapiro on Divergence and Transparency of Structural Estimates

Presented by: Angie Acquatella & Sarah Armitage

April 7, 2020

Outline

1. Introduction

2. AGS 2019: Informativeness
   - Mechanics of the Informativeness Ratio
   - Example 1: RCT
   - Example 2: Long Term Care Insurance

3. AGS 2020: Transparency

Overview of AGS Research Agenda

1. Series of papers are part of a broader research agenda on interpreting structural work
2. Papers covered today:
    - **How descriptive statistics and reduced form estimates can inform structural parameters?**
    - **How to increase the transparency of assumptions made in structural work?**
3. Earlier papers part of this agenda address:
    - How sensitive a structural parameters to specific moments?

Outline

AGS 2019: Informativeness

Mechanics of the Informativeness Ratio

Big idea: **connecting** "descriptive analysis" to structural estimation

- All of our statistical analysis informed by an economic model
- AGS leverage the underlying economic model to connect these two forms of analysis

	Reduced From	Structural
Empirics	tax experiment dif-in-difs, bunching at the kinks, etc	data moments (GMM) to map model primitives to observables
Objectives	various	counterfactual simulation
Robustness of estimates	laboriously verified	potentially very sensitive

AGS 2019: Informativeness

Mechanics of the Informativeness Ratio

- Ingredients of AGS (2019):
    1. c: model primitives
    2. $\hat{\gamma}$: reduced form estimate
    3. $\hat{c}$: structural estimate
- We will first illustrate each of these with a PF example
- Then show how informativeness Δ connects the two

AGS 2019: Informativeness

Mechanics of the Informativeness Ratio

An example we're familiar with: **labor supply elasticity**

- Model of the world c:

$$z \in \arg\max_{z} U(\underbrace{(1-\tau)z + R}_{c\text{ , post-tax}}, \underbrace{z}_{\text{income}})$$

- Object of interest

$$\varepsilon_z = \frac{(1-\tau)}{Z} \frac{dZ}{d(1-\tau)}$$

- in PF we care about this for income tax policy, arguably a counterfactual
- this is what AGS (2019) mean by c

AGS 2019: Informativeness

Mechanics of the Informativeness Ratio

Non-structural approach $\hat{\gamma}$:

- How do we typically go about estimating ε_z?
  - Tax experiments: Negative Income Tax experiment 1960/70s (Ashenfelter and Plant JHR' 90 +)
  - Natural experiments: Lottery winners (Imbens, Rubin, Sacerdote AER '01)
  - Bunching at the kink: elasticity = excess mass at kink / change in NTR (Saez '10)
  - ...or Swedish administrative data and incredible regulatory cooperation in research
- This is $\hat{\gamma}$ in the AGS framework

AGS 2019: Informativeness

Mechanics of the Informativeness Ratio

Structural approach $\hat{c}$:

- How would an IO person go about estimating ε_z?
  - FOC of our agent: $0 = (1-\tau)u_c + u_z$
  1. GMM: parametrize $U(c, z)$, assume the agent is on their FOC, solve for model parameters that rationalize observed data, then calculate elasticity
  2. Moment inequalities: parametrize $U(c, z)$, assume the agent's decisions are better than any other decisions they could have made, then solve for the range of model parameters that are consistent with this optimality condition
  - ... Prescott '04 calibration of GE model has a similar flavor: back out ε_z that rationalizes data in the world
- This is $\hat{c}$ in the AGS framework

AGS 2019: Informativeness

Mechanics of the Informativeness Ratio

Advantages and limitations of each approach:

- Reduced Form Estimate
  - Gives you the aggregate elasticity, evaluated at a specific point
  - Or may have to assume constant elasticity
  - Since the estimated behavioral response doesn't map to primitives that will stay fixed in a counterfactual world, less clear how to extrapolate
    - Short run vs long run? female vs male? low versus high income?
- Structural Estimate
  - Gives you the whole curve of elasticities at different income levels
  - Could simulate the effect of a welfare program at the bottom
  - But may be very sensitive to model misspecification

AGS 2019: Informativeness

Mechanics of the Informativeness Ratio

Hypothetical Empirical Exercise:

- Get $\underbrace{\varepsilon_z}_{c}$ for the top income earners
- RF way
  - Data: have an exogenous change in the top income tax rate
  - Computation: can get the elasticity off the dif-in-dif
- Structural way
  - Data: have labor/leisure ratios for individuals across income groups + survey data on marginal utility of consumption
  - Computation: estimate random coefficients discrete choice model to get the MRS between labor/leisure + MU of consumption

- **Idea: in principle, the elasticity calculated from the MRS of labor and leisure + MU of consumption of top income earners should coincide with the dif-in-dif**

AGS 2019: Informativeness

Mechanics of the Informativeness Ratio

Characterizing bias

- Why might the estimates of ε_z differ under these two approaches?
  - $\hat{c}$: MRS and MU of consumption are the $\hat{c}$ ($\implies$ $\hat{\varepsilon_z}(z)$)
  - $\hat{\gamma}$: dif-in-dif is the $\hat{\gamma}$ (which in this case $= \hat{\varepsilon_z}(z_{top})$)
- Bias captures notion of error in estimators from either approach
- b^N: largest possible bias in $\hat{c}$
- b^R: largest possible bias in $\hat{c}$ if additionally restricted to $\hat{c}$ to be consistent with $\hat{\gamma}$
- Object of interest: b^R/b^N

AGS 2019: Informativeness

Mechanics of the Informativeness Ratio

Understanding Δ (what AGS call *informativeness*)

$$b^R/b^N = \sqrt{1 - \underbrace{\Delta}_{\text{informativeness}}}$$

- Main result: Δ is "the R^2 from a regression of the structural estimate on the descriptive statistics when both are drawn from their joint (asymptotic) distribution"
- AGS characterize the bias ratio $b^R/b^N = \sqrt{1-\Delta}$
  - in the normal linear model
  - and in the asymptotic analogue under local misspecification and non-local misspecification

AGS 2019: Informativeness

Interpreting values of Δ

- If $\Delta \to 1$, a reader interested in worst-case bias can focus on evaluating the assumptions that govern $\hat{\gamma}$ and relate it to $\hat{c}$
- If $\Delta \to 0$, the reader would want to focus on evaluating assumptions that govern features orthogonal to $\hat{\gamma}$
- In our example:
  - $\Delta = 1$ iff the implied elasticity from the $\hat{c}$ MRS and MU of consumption for top income earners is exactly equal to the observed dif-in-dif from the tax experiment
  - $\Delta = 0$ if these two are orthogonal

AGS 2019: Informativeness

Remarks:

- Computational ease of Δ makes their results extremely powerful
- Values of Δ close to 1 or 0 have a practical implication for the reader interpreting structural estimates, but less clear what to do if $\Delta = .5$
- Aside: if we were to add the exogenous tax rate change as a moment in the estimation of $\hat{c}$, $b^N = b^R$ and $\Delta = 1$. The 2017 paper is about which moments we want to take more seriously for estimation when $\Delta = 1$.
  - The Sensitivity section of this paper generalizes this analysis for when $\hat{\gamma}$ is not the vector of moments, so that possibly $\Delta < 1$.

Outline

AGS 2019 Example: PROGRESA

- **Paper setting:** mothers in poorest households in targeted Mexican villages were given grants to keep their children in school.
  - *"A tightly parameterized model... could identify the effect of the program even before its implementation, using variation in the opportunity cost of schooling (i.e. the wage) across communities where the program is not available."*
  - *"The use of non-experimental data to carry out ex ante evaluation, with no variation in school grants, requires additional assumptions: one needs to assume that conditional on the activity of the child (education or work), the income of the child and other household income have the same effect on utility."*

AGS 2019 Example: PROGRESA

- **Paper setting:** mothers in poorest households in targeted Mexican villages were given grants to keep their children in school.
  - *"A tightly parameterized model... could identify the effect of the program even before its implementation, using variation in the opportunity cost of schooling (i.e. the wage) across communities where the program is not available."*
  - *"The use of non-experimental data to carry out ex ante evaluation, with no variation in school grants, requires additional assumptions: one needs to assume that conditional on the activity of the child (education or work), the income of the child and other household income have the same effect on utility."*
- Paper estimates structural model of education choices using data from PROGRESA experiment and uses model to simulate effect of changing program parameters.

AGS 2019 Example: PROGRESA

- **Estimate of interest ($\hat{c}$):** partial equilibrium effect of the counterfactual re-budget on the school enrollment of eligible children, accumulated across age groups.
- **Descriptive statistics available ($\hat{\gamma}$):** Impact on eligible children, impact on ineligible children, both

AGS 2019 Example: PROGRESA

- **Estimate of interest ($\hat{c}$):** partial equilibrium effect of the counterfactual re-budget on the school enrollment of eligible children, accumulated across age groups.
- **Descriptive statistics available ($\hat{\gamma}$):** Impact on eligible children, impact on ineligible children, both
- Recipe for calculating informativeness provided for ML and GMM $\rightarrow$ can be applied here exactly.

Table 1: Estimated informativeness of descriptive statistics for the effect of a counterfactual rebudgeting of PROGRESA (Attanasio et al. 2012a)

Descriptive statistics $\hat{\gamma}$	Estimated informativeness $\hat{\Delta}$
All	0.283
Impact on eligibles	0.227
Impact on ineligibles	0.056

- If $\hat{\gamma}$ is correctly specified by researchers' model: worst-case bias reduced in $\hat{c}$ reduced by $1 - \sqrt{1 - 0.28} \approxeq 0.15$.

Outline

1. Introduction

2. AGS 2019: Informativeness
   - Mechanics of the Informativeness Ratio
   - Example 1: RCT
   - Example 2: Long Term Care Insurance

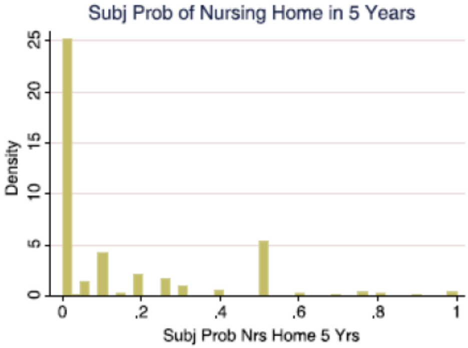
3. AGS 2020: Transparency

AGS 2019 Example: Long-Term Care Insurance

- **Paper Setting:** Hendren looks at phenomenon of insurance rejections, i.e. people who are denied insurance by the insurance company
  - Puzzle in a standard market: why don't prices adjust to offset the sickness of these patients?
  - Leading explanation: private information $\implies$ adversely selected risk mix given any set of pricing characteristics
- **Question**: Test whether residual private information explains rejections
  - private information Z, associated with a potential loss L
  - test whether Z is more predictive of L for rejectees than for non-rejectees
- **Data**: Health and Retirement Study, 1993-2008. He considers three markets: long-term care insurance, private disability insurance, and life insurance
  - LTC Q: What's the % chance you will move to a nursing home?

AGS 2019 Example: Long-Term Care Insurance

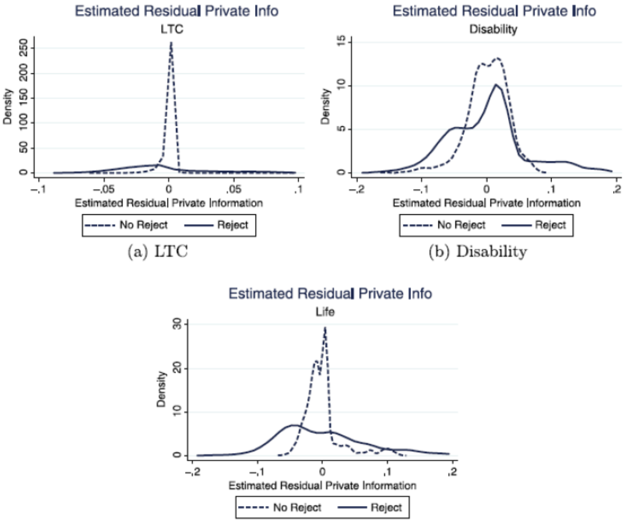
- Survey data (what will be the $\hat{\gamma}$):



(a) Long-Term Care

AGS 2019 Example: Long-Term Care Insurance

- Estimation (what will be the $\hat{c}$)::



(a) LTC

(b) Disability

AGS 2019 Example: Long-Term Care Insurance

- **Estimate of interest ($\hat{c}$):** minimum pooled price ratio among rejectees
  - this quantifies the implicit tax individuals would need to be willing to pay so that a market could exist (meaning st revenue$\geq$ cost)
  - object is the cheapest **cost** of providing (an infinitesimal amount of) insurance
- **Descriptive statistics available ($\hat{\gamma}$):** four vectors
  1. focal point groups: fraction of respondents who report exactly $0, .5, 1$ $\implies \Delta = .01$
  2. non-focal point groups: complement of 1 $\implies \Delta = .02$
  3. fraction of respondents that eventually need LTC $\implies \Delta = .68$
  4. the three vectors appended $\implies \Delta = .7$
- Using (3) fraction of respondents that eventually need LTC reduces worst case bias by $1 - \sqrt{1 - .68} \approx .43$

Outline

AGS 2020: Transparency in Structural Research

Big idea: Consider the experience of reading a structural paper...

AGS 2020: Transparency in Structural Research

Big idea: Consider the experience of reading a structural paper...

- "A reader who accepted the full list of assumptions could walk away having learned a great deal.
- "A reader who questioned even one of the assumptions might learn very little, as they would find it hard or impossible to predict how the conclusions might change under alternative assumptions."

- What are best practices for ensuring that structural work is informative to a range of readers?

AGS 2020: Transparency in Structural Research

Formal definition of transparency:

- First, consider **model of scientific communication**, following Andrews and Shapiro (2018):
  - Data D, relevant to quantity of interest c
  - Researcher reports estimate $\hat{c}(D)$ and auxiliary statistics $\hat{t}(D)$
  - $\hat{c}$ is valid under researcher's assumptions a_0, where $D \sim F(a_0, \eta)$ and $c(a_0, \eta)$
  - Reader r might have different assumptions $a \neq a_0$, so quantity of interest would become $c(a, \eta)$
  - After receiving report $(\hat{c}, \hat{t})$, reader updates prior beliefs, selects his/her own estimate d_r of c, and realizes quadratic loss $(d_r - c)^2$

AGS 2020: Transparency in Structural Research

Formal definition of transparency:

- **Communication risk for reader r:** ex ante expected loss given $(\hat{c}, \hat{t})$

$$\mathrm{E}_r[\min_{d_r} \mathrm{E}_r[(d_r - c)^2 | \hat{c}, \hat{t}]] = \mathrm{E}_r[Var_r(c|\hat{c}, \hat{t})]$$

AGS 2020: Transparency in Structural Research

Formal definition of transparency:

- **Communication risk for reader r:** ex ante expected loss given $(\hat{c}, \hat{t})$

$$\mathrm{E}_r[\min_{d_r} \mathrm{E}_r[(d_r - c)^2 | \hat{c}, \hat{t}]] = \mathrm{E}_r[Var_r(c|\hat{c}, \hat{t})]$$

- **Transparency for reader r:** reduction in expected loss from observing $(\hat{c}, \hat{t})$ relative to observing full data

$$T_r(\hat{c}(\cdot), \hat{t}(\cdot)) = \frac{Var_r(c) - \mathrm{E}_r[Var_r(c|\hat{c}, \hat{t})]}{Var_r(c) - \mathrm{E}_r[Var_r(c|D)]}$$

AGS 2020: Transparency in Structural Research

Formal definition of transparency:

- **Communication risk for reader r:** ex ante expected loss given $(\hat{c}, \hat{t})$

$$\mathrm{E}_r[\min_{d_r} \mathrm{E}_r[(d_r - c)^2 | \hat{c}, \hat{t}]] = \mathrm{E}_r[Var_r(c|\hat{c}, \hat{t})]$$

- **Transparency for reader r:** reduction in expected loss from observing $(\hat{c}, \hat{t})$ relative to observing full data

$$T_r(\hat{c}(\cdot), \hat{t}(\cdot)) = \frac{Var_r(c) - \mathrm{E}_r[Var_r(c|\hat{c}, \hat{t})]}{Var_r(c) - \mathrm{E}_r[Var_r(c|D)]}$$

- Note the following relationships:

$$\mathrm{E}_r[Var_r(c|D)] \leq \mathrm{E}_r[Var_r(c|\hat{c}, \hat{t})] \leq Var_r(c)$$

AGS 2020: Transparency in Structural Research

Formal definition of transparency:

- **Communication risk for reader r:** ex ante expected loss given $(\hat{c}, \hat{t})$

$$\mathrm{E}_r[\min_{d_r} \mathrm{E}_r[(d_r - c)^2 | \hat{c}, \hat{t}]] = \mathrm{E}_r[Var_r(c|\hat{c}, \hat{t})]$$

- **Transparency for reader r:** reduction in expected loss from observing $(\hat{c}, \hat{t})$ relative to observing full data

$$T_r(\hat{c}(\cdot), \hat{t}(\cdot)) = \frac{Var_r(c) - \mathrm{E}_r[Var_r(c|\hat{c}, \hat{t})]}{Var_r(c) - \mathrm{E}_r[Var_r(c|D)]} \in [0, 1]$$

- Note the following relationships:

$$\mathrm{E}_r[Var_r(c|D)] \leq \mathrm{E}_r[Var_r(c|\hat{c}, \hat{t})] \leq Var_r(c)$$

AGS 2020: Transparency in Structural Research

Authors identify four high-level ways to improve transparency:

1. through Descriptive Statistics
2. through Identification
3. through Estimation
4. through Sensitivity Analyses

Improving Transparency: through Descriptive Statistics

Researcher may report $\hat{s}$ as part of auxiliary statistics $\hat{t}$: summary statistics, data visualization, correlations illustrating key causal relationships, ...

1. $\hat{s}$ may provide evidence about quantity of interest c that is informative under wider range of assumptions than baseline model assumption a_0

Improving Transparency: through Descriptive Statistics

Researcher may report $\hat{s}$ as part of auxiliary statistics $\hat{t}$: summary statistics, data visualization, correlations illustrating key causal relationships, ...

1. $\hat{s}$ may provide evidence about quantity of interest c that is informative under wider range of assumptions than baseline model assumption a_0

   - Example: In PROGRESA paper, "A reader who does not accept all of the assumptions of the structural model might nevertheless learn a fair amount about the likely effects of the reallocation from comparing the treatment effects on older and younger children."

Improving Transparency: through Descriptive Statistics

Researcher may report $\hat{s}$ as part of auxiliary statistics $\hat{t}$: summary statistics, data visualization, correlations illustrating key causal relationships, ...

1. $\hat{s}$ may provide evidence about quantity of interest c that is informative under wider range of assumptions than baseline model assumption a_0

   - Example: In PROGRESA paper, "A reader who does not accept all of the assumptions of the structural model might nevertheless learn a fair amount about the likely effects of the reallocation from comparing the treatment effects on older and younger children."

   - Example: If $|Corr_r(c, \hat{s})|$ is large and $\hat{s}$ is scalar, then researcher can directly bound average posterior variance:

$$\mathrm{E}_r[Var_r(c|\hat{s})] \leq Var_r(c)(1 - Corr_r(c, \hat{s})^2)$$

Improving Transparency: through Descriptive Statistics

Researcher may report $\hat{s}$ as part of auxiliary statistics $\hat{t}$: summary statistics, data visualization, correlations illustrating key causal relationships, ...

2. $\hat{s}$ may help reader to evaluate model assumptions a_0

Improving Transparency: through Descriptive Statistics

Researcher may report $\hat{s}$ as part of auxiliary statistics $\hat{t}$: summary statistics, data visualization, correlations illustrating key causal relationships, ...

2. $\hat{s}$ may help reader to evaluate model assumptions a_0

   - Example: Structural model of grocery demand with prices from other stores in chain used as instruments (Hausman instruments); descriptive statistics show this price variation is $\perp$ to key demographics.

Improving Transparency: through Descriptive Statistics

Researcher may report $\hat{s}$ as part of auxiliary statistics $\hat{t}$: summary statistics, data visualization, correlations illustrating key causal relationships, ...

2. $\hat{s}$ may help reader to evaluate model assumptions a_0

   - Example: Structural model of grocery demand with prices from other stores in chain used as instruments (Hausman instruments); descriptive statistics show this price variation is $\perp$ to key demographics.
   - Example: Structural model of bank lending that exploits credit score thresholds; descriptive statistics show observed borrower characteristics are smooth around discontinuities.

Improving Transparency: through Identification

Formal discussion of identification can improve transparency:

1. Discussions of identification should be precise.
   - Econometric identification of a point estimate is a binary property – should not say "primarily identifies" or "intuitively identifies"
   - Formally, if quantity of interest c "is identified by" $\hat{s}$, then the distribution of $\hat{s}$ is sufficient to infer the value of c under the model.
   - If identification is conditional on knowledge of some other parameters, then this relationship must be explicit – otherwise, not identified!

Improving Transparency: through Identification

Formal discussion of identification can improve transparency:

2. Discussion of model identification should be clearly distinguished from discussion of estimation.
    - If "c is identified by $\hat{s}_j$," that does not necessarily imply that $\hat{s}_j$ is an important determinant of $\hat{c}$.
    - Formal definition of identification with $\hat{s}$: Quantity c is identified by a specific vector of statistics $\hat{s}$ if $c(a_0, \eta) \neq c(a_0, \eta')$ implies distinct distributions of $\hat{s}$ under $F(a_0, \eta)$ and $F(a_0, \eta')$.

Improving Transparency: through Estimation

Understanding how estimator $\hat{c}$ depends on statistics is important to transparency:

1. Target descriptive statistics $\hat{s}$ directly in estimation (that is, $\hat{c} = h(\hat{s})$):
   - In practice, often implemented by minimum distance estimator.

Improving Transparency: through Estimation

Understanding how estimator $\hat{c}$ depends on statistics is important to transparency:

1. Target descriptive statistics $\hat{s}$ directly in estimation (that is, $\hat{c} = h(\hat{s})$):
    - In practice, often implemented by minimum distance estimator.
    - Simply providing formal definition of estimator may not make clear how $h(\cdot)$ depends on assumptions – some functions $h(\cdot)$ may be convincing given many different assumptions a, while others may only be convincing to readers who accept a_0.

Improving Transparency: through Estimation

Understanding how estimator $\hat{c}$ depends on statistics is important to transparency:

1. Target descriptive statistics $\hat{s}$ directly in estimation (that is, $\hat{c} = h(\hat{s})$):
    - In practice, often implemented by minimum distance estimator.
    - Simply providing formal definition of estimator may not make clear how $h(\cdot)$ depends on assumptions – some functions $h(\cdot)$ may be convincing given many different assumptions a, while others may only be convincing to readers who accept a_0.
    - Andrews et al. (2017) propose focusing on local sensitivity of estimator to targeted statistics. Sensitivity is defined as derivative of $h(\cdot)$.

Improving Transparency: through Estimation

Understanding how estimator $\hat{c}$ depends on statistics is important to transparency:

2. Show extent to which $\hat{c}$ depends on descriptive statistics $\hat{s}$ (that is, for $\hat{c} = h(\hat{s}) + v_h$, show magnitude of v_h and form of $h(\cdot)$).

Improving Transparency: through Estimation

Understanding how estimator $\hat{c}$ depends on statistics is important to transparency:

2. Show extent to which $\hat{c}$ depends on descriptive statistics $\hat{s}$ (that is, for $\hat{c} = h(\hat{s}) + v_h$, show magnitude of v_h and form of $h(\cdot)$).
   - Here local informativeness Δ is useful. When $\Delta = 1$, we return to case on previous slide. When $\Delta = 0$, $\hat{c}$ is asymptotically independent of $\hat{s}$.

Aside: Connection to 2019 paper

- This section of 2020 paper addresses the question: How can the relationship between $\hat{s}$ and $\hat{c}$ tell us about the relationship between $\hat{c}$ and a?
- 2019 paper formalizes relationship between $\hat{s}$ and $\hat{c}$. It may be the case that $\hat{s}$ tells us a lot about $\hat{c}$ (Δ close to 1), or not (Δ close to 0). Knowing Δ might be useful for:
  - Knowing how much $\hat{s}$ reduces bias due to misspecification.
  - Finding ways to increase transparency in scientific communication.

Improving Transparency: through Sensitivity Analyses

Sensitivity analyses can show how results depend on assumptions:

1. Show how $\hat{c}(a_0, \eta)$ changes under specific alternative assumptions (i.e., when number of relevant alternative assumptions is small).

Improving Transparency: through Sensitivity Analyses

Sensitivity analyses can show how results depend on assumptions:

1. Show how $\hat{c}(a_0, \eta)$ changes under specific alternative assumptions (i.e., when number of relevant alternative assumptions is small).
    - Example: BLP report results using logit demand model, random coefficients demand model, alternative utility specifications, etc.
    - This approach increases transparency far more than just reporting bounds of set of estimates!

Improving Transparency: through Sensitivity Analyses

Sensitivity analyses can show how results depend on assumptions:

2. Show how $\hat{c}(a_0, \eta)$ depends on assumptions through explicit function, $u(a) = \hat{c}_a - \hat{c}_{a_0}$ (i.e., when number of relevant alternative assumptions is large).
   - Example: Omitted variables bias formula allows reader to calculate bias in estimator for any possible assumption about correlation of omitted regressor.

Improving Transparency: through Sensitivity Analyses

Sensitivity analyses can show how results depend on assumptions:

2. Show how $\hat{c}(a_0, \eta)$ depends on assumptions through explicit function, $u(a) = \hat{c}_a - \hat{c}_{a_0}$ (i.e., when number of relevant alternative assumptions is large).

    - Example: Omitted variables bias formula allows reader to calculate bias in estimator for any possible assumption about correlation of omitted regressor.
    - Example: Conley et al. (2012) generalize the OVBF to an instrumental variables setting.
    - Example: Andrews et al. (2017) provide analogue of the OVBF for general minimum distance estimators.

Improving Transparency: through Sensitivity Analyses

Sensitivity analyses can show how results depend on assumptions:

3. Show what properties of the data would be required to reverse the (qualitative) conclusion about $\hat{c}(a_0, \eta)$:

Improving Transparency: through Sensitivity Analyses

Sensitivity analyses can show how results depend on assumptions:

3. Show what properties of the data would be required to reverse the (qualitative) conclusion about $\hat{c}(a_0, \eta)$:

    - Example: Structural model of vertical integration between hospitals and insurers; paper finds that vertical integration increases consumer surplus, but this result depends on degree of consumer price sensitivity.

    - Shows that researcher's question is indeed an empirical one!

Discussion, Limitations, and How We Might Apply These Ideas in Our Own Work...

- Questions? Comments?