Translating Research Results into Practice

Major Elements in the Toolkit for Good Empirical Research

Erich Battistin

Improving Education through Accountability and Evaluation Lessons from Around the World

Roma, October 5 2012

Evaluation Culture

- Researchers over the years have become more weary of claims that are *not* supported by credible **evidence**
- Before intervening on specific school inputs, it is fair to ask where is the **causality** that motivates our approach
- This poses a fundamental distinction between "**hunting causes**" (i.e. using methods to assess what works in the context under investigation) and "**using**" them for policy (i.e. generalize results to contexts other than those in which they were obtained)
- "Hunters" cover the whole spectrum between *practitioners* and *professionals* of the most rigorous methods

Evaluation Culture

- Smart “hunters” understand that they must interact with and inform “policy designers”, and that they both need reliable sources of **data**
- This calls for substantial interaction amongst players, which someone likes to call **evaluation culture**
- **R1**: The evaluation culture is fueled by tilting the **balance** toward the most rigorous hunters, and by learning from the meta-analysis of studies conducted by practitioners
- Make the most out of what we have, and make a plan for moving forward based on a common research agenda tailored around examples of best practice

Hunters and Preys

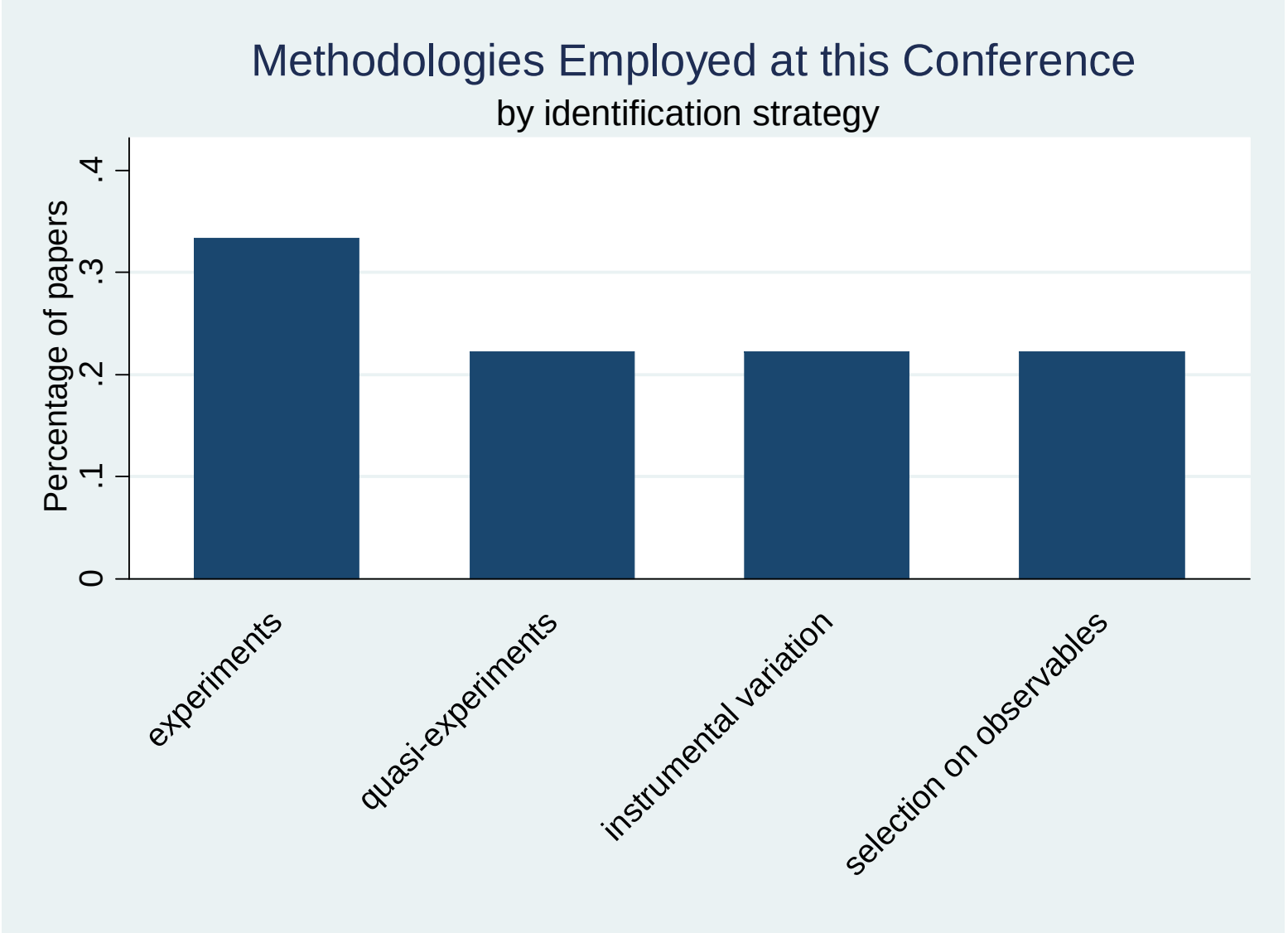
Preys:

- Most papers presented are after the effects of changing (various) **school inputs** on academic achievement
- There is a **ceteris paribus condition** maintained: "target" and "control" groups of schools/students differ only with respect to the input under investigation
- **R2**: Moving many **inputs** at the same time is not the wisest choice to make if we want to pin down neat causal relationships

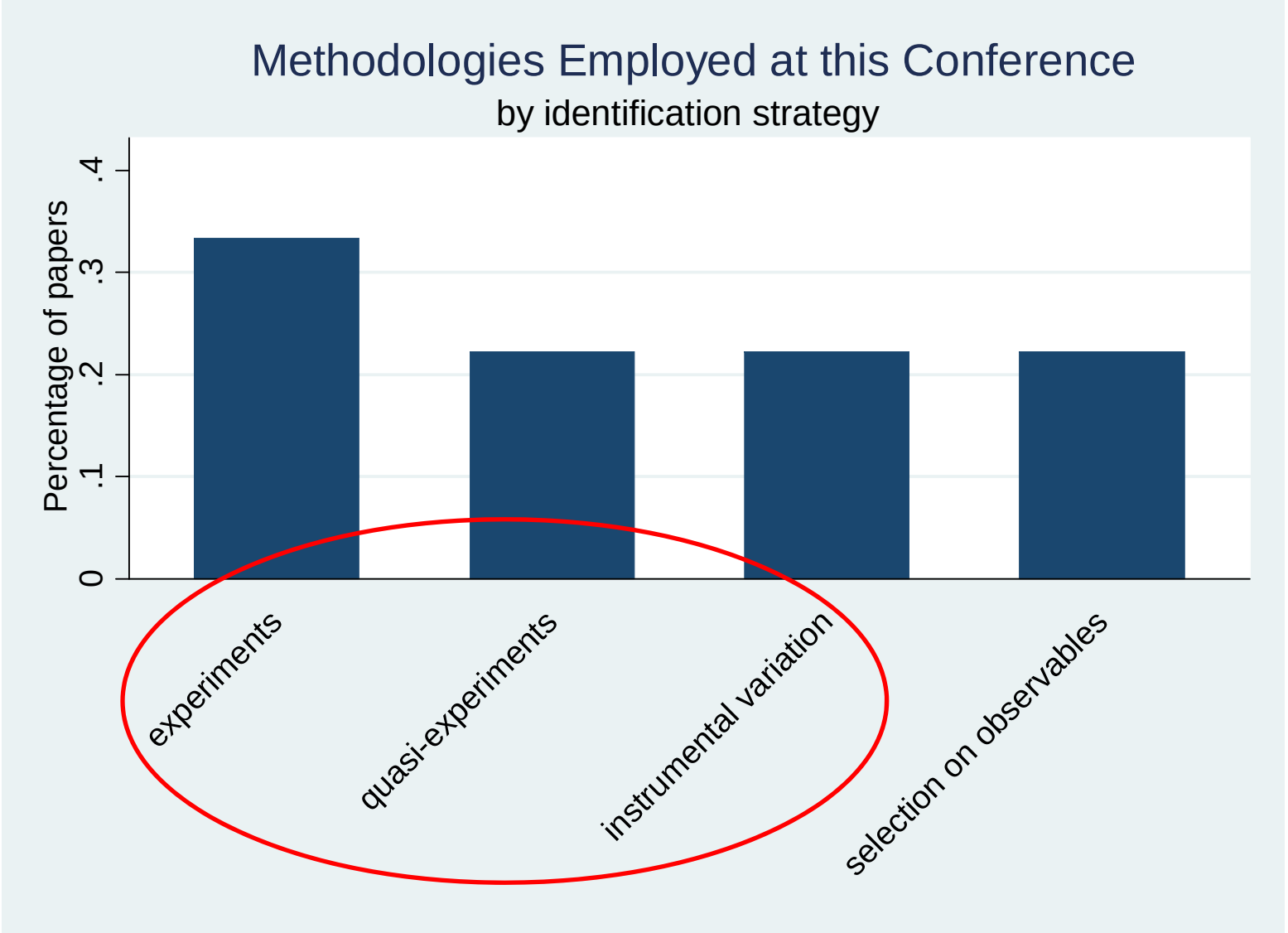
Hunters:

- A variety of alternative approaches (and "assumptions") to learn about "what works", "for whom" and "where"

Hunters and Preys



Hunters and Preys



Relevant Questions

The ultimate aim of the INVALSI should be to play an active role in providing answers to ***two* broad classes of evaluation questions** which are closely interrelated

Q1: **Assessing** the impact of policy interventions that are already (or will be soon) put in place

Q2: **Forecasting** the effects of *new* interventions, never historically experienced, using what it is known of past interventions. This is a problem that policy makers have to solve daily, and calls for an extension of the traditional approach to policy evaluation

Looking Forward: the Role of INVALSI

- **R3**: Most credible (i.e. internally valid) answers obtained through **randomized control trials** (RCTs)
- Typically implemented on a **small scale population** (**external validity**) **with large enough sample size**, depending on the institutional rules in place
- Closely resemble the idea of **pilot/control studies** in the UK, and of interventions in the Tropics
- Call for strict enforcement of assignment rules, as deviations from the mandated status are almost certainly purposeful. This motivates the scale of the experiment
- **R4**: The **involvement of local institutions** is fundamental to induce a shock to the demand for evaluation in Italy (not only in education policies)

Looking Backward: the Role of INVALSI

- **R5**: Understanding that RCTs – under certain conditions - yield the most credible answer **must not be turned into a crusade** against all other methods
- What can we learn from the implementation of **past policies**? Credible and rigorous alternatives to RCTs are available, and exploit for example changes in the eligibility rules over time and/or across subjects/areas
- **R6**: Integrating future (carefully designed) evaluations with the evaluation of past interventions requires **easily accessible data sources**
- This guarantees **transparency** and **replicability** of results for all players

Data Requirements

- We need **integrated sources of information** to answer increasingly more ambitious research questions
- Administrative data produced by the INVALSI through the *Servizio Nazionale di Valutazione* (Italian National Testing Service) are a good example of this
- Determining the effects of early childhood conditions on performance and educational attainment requires **longitudinal data** on the same individuals. **Cohort studies** are useful to this end (e.g. see the UK)
- **R7**: INVALSI should take an active role in promoting the development of a **common repository / data archive** that meets high standards of quality

Learning Ex Ante About New Policies

- In the traditional approach, the variation in policy instruments generated by a policy already experienced (e.g. the availability of scholarships for students scoring values at a cognitive test above a certain cutoff) is used to learn about the effects of *that* particular policy
- Extrapolation to new policy settings requires additional tools, as interest lies in the effects of *new* policy configurations (e.g. a change in the cutoff employed to assign a scholarship) which were *never* historically experienced
- This paves the way for a new era of policy evaluation (there were no papers on this topic)

Wrap Up: Final Recommendations

- The **evaluation culture** is fueled by tilting the balance toward the most rigorous hunters of causal relationships, and by making the most out of studies carried out by practitioners
- Moving many **inputs** at the same time is not the wisest choice to make if we want to learn about causal relationships
- Most credible answers are obtained through **randomized control trials**. Under certain conditions, they represent the "gold standard"
- They should be part of a **structured research agenda**, not just scattered examples

Wrap Up: Final Recommendations

- The **involvement of local institutions** is fundamental to induce a shock to the demand for evaluation and to ensure standards of quality
- Understanding that RCTs – under certain conditions - yield the most credible answer **must not be turned into a crusade** against all other methods
- We need **integrated sources of information** to answer increasingly more ambitious research questions: this guarantees **transparency** and **replicability** of results
- INVALSI should take an active role in promoting the development of a **common repository** with high standards of quality (there are examples of best practice to follow)