Journal of Institutional Economics (2016), 12: 2, 277–303
 doi:10.1017/S1744137415000429
First published online 30 October 2015

Not by technique alone. A methodological comparison of development analysis with Esther Duflo and Elinor Ostrom

AGNÈS LABROUSSE*

Centre de Recherche sur les Institutions, l'Industrie et les Systèmes Economiques d'Amiens (CRIISEA), University of Picardie – Jules Verne, Amiens, France

Abstract. This contribution aims at an original comparison of development analysis with Elinor Ostrom and Esther Duflo from a methodological standpoint, scrutinising their relationship to theory and their operative research strategies. Both perspectives are investigated as case studies for a broader discussion about significant trends in economics and social sciences. Duflo and the J-PAL's approach illustrates – in its own way – new trends and some blind alleys in contemporary forms of mainstream economics, whereas Ostrom and the Bloomington school point towards the marked theoretical and methodological reflexivity of institutionalism, its sensitivity to historical diversity and openness towards social sciences. Distinct social philosophies and episteme are at stake displaying a great divide between two brands of realism and pragmatism, two relationships to development, expertise and knowledge. The paper also contrasts Duflo's methodological monism and mechanistic piecemeal analysis with Ostrom's methodological pluralism and adaptive complex systems analysis.

1. Ostrom and Duflo: two distinguished women, some commonalities and obvious differences

Exceptionally enough, two prestigious awards were successively awarded to two women, the 2009 Nobel Prize in Economics for Elinor Ostrom (1933–2012) and the 2010 Clark medal for Esther Duflo (1972-), which is commonly considered as the antechamber of the Nobel Prize. Beside this outstanding symbolic capital, it is noteworthy that both authors analyse economic development in the broad sense – including socio-economics dimensions – and tend to focus on the local dimension of development processes, either to govern common pool resources or

*Email: agnes.labrousse@u-picardie.fr
The author wishes to thank without implicating Bernard Chavance, Stéphane Longuet, participants in the Vienna EAEPE and the Paris joint AFEP-AHE-IIPPE conferences, the CRIISEA Sage-Workshop, as well as the two anonymous referees for their highly helpful comments. She is also indebted to Alain Desrosières for inspiring discussions and his encouragement at the early stage of this work.

to fight poverty. Both mobilise experiments in their investigations (laboratory and field experiments, as far as Ostrom is concerned, field randomised experiments as regards Duflo). Both practice forms of 'working together', their researches involving a strong collective and international dimension, be it with the Bloomington *Workshop in Political Theory and Policy Analysis*, which Elinor Ostrom led together with her PhD supervisor, university colleague and husband Vincent Ostrom or with the *Jameel Poverty Action Lab* (J-PAL) that Esther Duflo co-founded with her PhD Supervisor, MIT colleague and life partner, Abhijit Banerjee. Moreover, both approaches led to startling results measured up to standard economics: following the Bloomington school, the so-called 'tragedy of the commons' is a misleading metaphor that can be overcome in numerous situations; free access to medical devices does not reduce their use, on the contrary, according to the J-PAL (Cohen and Dupas, 2007; Duflo, 2009). Fieldwork and the attention to concrete devices play an important role in their respective works, although to varying degrees and according to different modalities.

Beyond these common denominators, differences are palpable. To mention one among many, the J-PAL approach rely almost exclusively on the method of randomised controlled trials (RCT), claimed to be the 'gold standard' to produce 'hard evidence'; the Bloomington school combine a variety of qualitative and quantitative methods; in their understanding, there is no such thing as a gold standard. A systematic comparison of these prominent approaches has not been conducted so far. The fact that both lay emphasis on their methodology invites to a methodology-centred appraisal.

By contrasting these approaches, we hope to raise broader issues, going beyond a mere comparison between two high-profile scientists and research teams, considering them as exemplars of significant trends in economics and social sciences. The J-PAL approach illustrates – in its own way – new trends as well as some blind alleys in contemporary mainstream economics (with a great deal of statistical empirics, a social engineering posture and links to behavioural psychology), whereas the Bloomington school points towards the marked theoretical and methodological reflexivity of the 'old' institutionalist tradition, its sensitivity to historical diversity and its openness towards social sciences. This leads to delineating two social philosophies and episteme: two brands of realism and pragmatism and two relationships to development, expertise and knowledge, opposing also methodological monism and static mechanistic piecemeal analysis (Duflo) with methodological pluralism and adaptive complex systems analysis (Ostrom).

Duflo and Ostrom are not seen here as isolated individualities but as leading figures of two epistemic communities with strong intellectual and personal ties. Hence, this comparative investigation is based on the analysis of books[1]

1 Included are notably three books in French by Duflo. All translations from French by the author.

https://doi.org/10.1017/S1744137415000429 Published online by Cambridge University Press

and articles by the two authors and their close collaborators, associated with interviews given to scientific reviews and to the press. In Ostrom's case, decades of work led to rich accounts of her research praxis and a well-defined epistemological and conceptual framework, even though her social philosophy is less discernible at first sight (Aligica and Boettke, 2011). By contrast, the overall approach by Duflo and the J-PAL is less developed: while the statistical technique is well documented, her fieldwork experience, her relationship to theory and her social philosophy are often implicit, constraining to draw on hints and practices transpiring from publications and interviews to delineate her epistemology, at the risk of misinterpretations. Hence, the effort undertaken here to supplement this bibliographical material with an emergent body of literature examining how RCTs are conducted in practice (Bernard *et al.*, 2012, Bureau *et al.*, 2013, Devaux-Spatarakis, 2014; Gomel and Serverin, 2013; Jatteau, 2013; Morvant-Roux *et al.*, 2014; Quentin and Guérin, 2013). Since this paper focuses on methodology in operation, this literature provides valuable insights in the research praxis of the 'randomistas'. To grasp the transfer from medicine to development studies, medical and historical investigations on RCTs are also included.

The remainder of the paper proceeds as follows. Section 2 scrutinises Duflo's and Ostrom's relationships to theory, historical specificity and emergence. Section 3 sheds some light on their operative research strategies (construction of observation and experimentation, role of the observer, relationship to other disciplines). Section 4 concludes that distinct episteme and social philosophies are at stake.

2. Far from blackboard economics. Ways to deal with theoretical issues

> 'We are academics, and like most academics we formulate theories and stare at data. But the nature of the work we do has meant that we have also spent months, spread over many years, on the ground'. Banerjee and Duflo (2011)

> 'To discover the diversity of locally designed rules, to understand how the institutional arrangements work given the biophysical conditions of a resource, and the culture of the users, you have no other choice but to go there and do field work. [. . . We] have together spent many years in the field'. E. Ostrom in Aligica (2003: 9)

The attention to empirics, the aspiration to deal with concrete issues in a realistic manner is prominent in both approaches. Duflo emphasises the months spent in fieldwork. She contrasts the unrealisticness of laboratory experiments *vis-à-vis* the realistic conditions of the randomised field experiments she promotes (2009). She is obsessed with day-to-day devices (to increase the use of insecticidal bed nets or the consumption of salt fortified with iron and iodine). Extensive fieldwork was present from the very beginning in Ostrom's work – be it in it her

https://doi.org/10.1017/S1744137415000429 Published online by Cambridge University Press

PhD thesis on the Californian ground water basin (1965) or her post-doctoral work in municipal policing (Boettke *et al.*, 2013; Ostrom *et al.*, 1977), carrying out participant observation, in-depth interviews to 'penetrate social reality' and to develop 'empirically valid theories' (E. Ostrom, in Aligica, 2003: 8). While both are far from 'blackboard economics', they diverge in the way fieldwork is connected with theory and about the place and meaning of theory.

Two relationships to theory, isolation and causal processes

Theory as exploratory and explanatory framework or as propositions ante experimentum?

Ostrom builds an encompassing analytical architecture, based on three theoretical tiers: 'A general *framework* helps to identify the elements (and the relationships among these elements) that one needs to consider for institutional analysis. Frameworks organize diagnostic and prescriptive inquiry. They provide the most general set of variables that should be used to analyze all types of settings relevant for the framework'. [...], *theories* focus on parts of a framework and make specific assumptions that are necessary for an analyst to diagnose a phenomenon, explain its processes and predict outcomes. The development and use of *models* involve making precise assumptions about a limited set of variables and parameters [...] using a particular theory'. (Ostrom, 2011: 8). This architecture enables to explore social reality in a systematic way, diagnosing and explaining the processes behind the recorded regularities and outcomes. A structured system of classification allows to register rules systems and to assess how rules affect each of the variables of an action situation (Ostrom, 2011: 19). The IAD (Institutional Analysis and Development) and SES (Social-Ecological Systems) frameworks are both exploratory and explanatory and are used in repeated interactions between empirical investigation, modelling and theorisation. Moreover, the Bloomington school uses a broad range of references belonging mainly to the family of institutional and evolutionary analysis in political economy and beyond. Crawford and Ostrom (2005) defined a detailed 'institutional grammar'. Additionally, from 'action arena' to 'redundancy', the Bloomington school has forged an extensive lexicon, summarised by McGinnis (2011).

Contrasting with Ostrom's theoretical and conceptual elaborations, the theorisation and the conceptualisation are still in limbo in Duflo's work. Deaton (2009) and Acemoglu (2010) regret the lack of theory in the J-PAL experiments. They are actually few and narrow theoretical references in Duflo's work. These references are mostly recent and oscillate between the two polar extremes represented by Jeffrey Sachs' centralised constructivism and William Easterly's decentralised Hayekism. Conceptual reflection is rather poor: important concepts like autonomy, capabilities or human development are just mentioned *en passant* or in books' subtitles (Duflo, 2009, 2010a, 2010b). Development is implicitly

https://doi.org/10.1017/S1744137415000429 Published online by Cambridge University Press

equated with poverty reduction. In a quest for undisputable 'hard numbers', absolute poverty indicators are preferably used; the 99 cents of the World Bank (former) absolute poverty line are recurrently brought into play as a synecdoche for the lives of the poor. Contrary to experimental technique, the meaning, place and role of theory are not specified. De facto, theory is reduced to propositions *ante experimentum*, i.e. propositions defining the anticipated results of an experiment (if the theory of the poverty trap is true, the poor should maximise their intake of calories to enhance their work productivity and, thus, their income, then one should expect that the food budget should go up proportionally faster than total spending; Banerjee and Duflo, 2011: 24–25). 'Scientists are inspired by theory [but experiments] can contradict theory: The validity of an experiment is not grounded in the validity of the theory', Duflo (2009: 67) points out. Therefore, randomised experiments can contribute significantly to the invalidation of alleged 'universal laws' derived from mainstream economics: for instance, series of experiments show that, contrary to Mead Over's models, free access to medical devices increases their use. Yet, Duflo's approach lacks an articulate theory of agents and cognitive processes, as well as of institutions and structures. Some behavioural biases like procrastination are dispersedly hinted at, but a theory as a coherent set of propositions regarding rationality, agents and agency is still missing (Labrousse, 2010).

Here again, the contrast with Ostrom is striking: the latter develops an articulate framework where individuals are regarded as boundedly rational and communicating 'fallible learners'; being embedded in diverse organisational (farms, firms, households, government units, churches, associations etc.), institutional (nested rules systems), cultural (norms) and biophysical settings. Systematic connections between theoretical considerations, fieldworks, laboratory and field experiments lead to a well-defined anthropology of social actors (see for instance, Janssen *et al*., 2011). In Duflo's case, despite connections with behavioural psychology,[2] there is no systematic behavioural theory (see the only paper addressing directly this issue 'Poor but rational?': Duflo, 2003). Duflo's line may be typical of new trends towards statistical empiricism within mainstream economics, far from the 'high theory' associated with general equilibrium theory. In such works, statistical techniques are viewed as self-sufficient ways to do science. Duflo's relationship to theory – as a coherent set of principles devised to explain a group of phenomena – is weak and partial, especially in the explanation of why experiments yield surprising results compared to theoretical expectations (which is often the case). This leads to the important issue of causal chains.

2 Sendhil Mullainathan, a leading behavioural economist, was one of the J-PAL founders with Banerjee and Duflo.

https://doi.org/10.1017/S1744137415000429 Published online by Cambridge University Press

Evaluating an action or uncovering causal processes?

Except in the rare cases of linear monocausality, randomised experiments evaluate the efficiency of a special intervention ('whether it works') but they do not uncover the underlying causal agencies ('why and how it works'). It is a well-known concern in the field of clinical studies: there is clinical evidence that acupuncture can prevent postoperative nausea (Lee and Done, 1999) but the causal mechanisms behind these outcomes remain unidentified. Different types of evidence are delineated in 'evidence-based medicine' (Berriet-Solliec *et al.*, 2014): evidence of existence (description and verification of facts to build an agreement among different actors on the state of the world), evidence of causality (explaining the generating mechanism underlying the observed event), evidence of effectiveness (a given action yields the desired result), evidence of harmlessness (obtained when negative effects have been looked for and not found). The J-PAL produces mainly evidence of effectiveness. RCT is not well-suited to capture actors' motives and perceptions, the working of social interactions and norms, ambient contingent conditions – i.e. crucial factors to understand causal mechanisms in social settings – contrary to qualitative approaches (ethnography, in-depth interviews etc.).

RCTs are less armed to uncover complex causal processes in comparison to Ostrom's multiple methods (see part 3.2) and theoretical trilogy (framework, theories and models). This nested triptych enables to elucidate causal chains explaining both why collective action occurs and sometimes collapses, to unravel 'core theoretical [and empirical] puzzles', for instance, 'how resource users in the field develop rules to increase performance [considering that for example] farmers in Nepal, who lack academic or formal training, can on average outperform highly educated engineers in the design and operation of irrigation systems. What is the process that produces these outcomes'? (Ostrom and Basurto, 2011: 320). The SES framework enlightens processes to craft rules, processes behind the evolution of rules and norms and how different rules systems affect outcomes (performance, equity etc.), whereas RCT as a technique is incomplete: causal, theoretical links are missing.

Isolating a pure effect or analysing complex bundles of rules and actions?

RCT is basically a powerful isolation method to measure the mean effect of one action: it doesn't assess the joint impact of a bundle of actions. The interactions between multiple actions (sometimes called 'cocktail effects') are out of the picture: it is common knowledge in pharmacology that clinical studies do not take into account the synergy effects (correlated action or cooperation by two or more drugs so that their combined effect is greater than the sum of their individual effects),[3] the potentiation effects (the action of one adjuvant b intensifies the action of the drug a: a + b = A) and the antagonistic effects between drugs (the

3 Definitions excerpted from: http://medical-dictionary.thefreedictionary.com/.

https://doi.org/10.1017/S1744137415000429 Published online by Cambridge University Press

drug a counteracts the effect of the drug b for example), except in the rare cases of the joint test of two or three active principles. Because of statistical and costs constraints, the branches of a given experiment cannot be multiplied infinitely in practice. To test the impact of a measure (or sometimes different treatment modalities) and not bundles of measures is a prominent artificial element in the randomisation process. This modus experimendi – characterised by the exclusivity and the invariability of the treatment during the trial – contrasts and often conflicts with the modus governandi: in 'real life', governments and local actors usually tinker with appropriate packages of measures to adjust them pragmatically to (evolving) local conditions (Bureau *et al*., 2013), a tinkering process underlined by Ostrom.

In Ostrom's approach, isolation seems to occur at the level of laboratory experiment and simulations, whereas bundles of rules and actions are examined *in vivo* in specified action arenas. Ostrom's group designed diverse models (from game theory to agent-based models) and experiments to isolate factors explaining the regularities observed in field studies (for instance, the impact of communication on patterns of cooperation, Ostrom, 1999). Relevant in the field is rather a 'combinatory logic' (Ostrom and Basurto, 2011). It follows that the object of empirical analysis is not to isolate the intrinsic effect of a single element (is the effect of the rule to hire a guard from inside the user community good or bad *per se*?) but to understand the contextualised interactions and interrelationships between the joint elements in actual configurations (systems of rules) and their emergent properties. Complex structures are at stake: 'Many differences in surface reality can result from the way these variables combine with or interact with one another. Thus, it is the configurational nature of the set of elements present in each of these layers that helps to explain the great complexity observed in the world'. (Ostrom and Basurto, 2011: 322). Whereas, Duflo is in quest of optimal measures that could be generalised to fight global poverty, dealing empirically and theoretically with diversity is a leitmotiv in Ostrom's work.

Theory and diversity in time and space: the issue of historical specificity

> 'Static and mechanistic analysis is not adequate to understand the changing world in which we live. [. . .] we need to develop analytical tools for analyzing dynamic situations – particularly institutional change [. . . it] is one important step we can take to reduce emphasis on institutional monocropping that currently dominates much of social science thinking as well as that of development agencies'. (Ostrom and Basurto, 2011: 337)

Duflo's and Ostrom's pathways to deal with the specificity over time and space of the local arrangements they scrutinise diverge highly. Whereas, the first tends to shun this issue, the second aims at understanding architectural principles beyond the sheer variety of socio-ecological systems.

https://doi.org/10.1017/S1744137415000429 Published online by Cambridge University Press

Universal or contextual theory? Replication or comparison? Which generalisation?

Duflo is ambiguous regarding the issue of historical specificity (see Hodgson, 2001). One can find in her production some statements like 'context can matter' co-existing with the postulate that incentives produce the same outcomes everywhere or everywhere the poor live, or 'in the environment of developing countries'. It is thus unclear if the J-PAL is aiming at universal laws *à la* neo-classical economics or at regularities specified in time and space as it is the case in historical institutionalism and, particularly, in Ostrom's work, which highlights a 'situational logic', an 'ecological rationality' embedded in specific social and institutional settings (Aligica and Boettke, 2011). Duflo often obliquely assumes that there is a unity of the 'lives of the poor' or of the rationality of the poor, justifying to consider them as a homogenous entity and as a special subfield of economics (i.e. 'poor economics'). But *quid* of the huge diversity of developing countries, where to draw the frontier between the poor and the rest, are the poor in the North following the same logic and constraints as in the South? *Quid* of the internal diversity within the experiment's area? Measuring the mean effect of an action, RCTs often neglect heterogeneity inside groups (Deaton, 2009). A qualitative analysis conducted alongside the J-PAL Al Amana micro-credit experiment demonstrated that the diversity of agro-ecological settings in Morocco, the multiple local meanings associated with credit and debt, the variability of micro-interactions with credit officers and local leaders, explain why micro-credit demand was highly heterogeneous among treatment branches (Morvant-Roux *et al.*, 2014). Not only internal diversity is crucial but also external diversity: the repressed issue of historical specificity resurfaces with the well-known problem of 'external valididity', leaving open the possibility to generalise the result of a given experiment (Deaton, 2009; Rodrik, 2009). Here again, the answer of the J-PAL is equivocal: following Banerjee and Duflo (2009), generalisation means upscaling a development programme, generalising best practices like deworming children in one or several countries. As shown by Gomel and Serverin (2013), there is confusion between a political operation (a government upscales a local programme) and a scientific operation (extending the domain of validity of a local experiment). It is alleged that an accumulation of experiments can lead both to theory and to uncovering best practices: small, modest but 'general' solutions to fight global poverty: 'A single experiment does not provide a final answer on whether a program would universally "work" or not. But we can conduct a series of experiments, differing in either the kind of location in which they are conducted or the exact intervention being tested (or both)'. (Banerjee and Duflo, 2011: 26). But as stressed by Ostrom (2005: 30–1), 'if we cannot link the theoretical results into a coherent overall approach, we cannot cumulate knowledge'. Up to now, a systematic comparison of differing results of identical or has similar experiments not been used as a heuristic tool by

the J-PAL.[4] Academic publications are subject to a bias toward success stories, deplored by J-PAL members themselves, and triggered by journals that are keener to publish positive than negative results, as it is the case for medical journals, where published trials are no more than the 'tip of an information iceberg' (Ramsey and Scoggins, 2008). Besides, success stories can easily be promoted to donors.

Comparing diverse settings is central for the Bloomington school. Understanding 'failures' is as important as understanding successes: the explaining factors at stake being the same (like involving or not the users in the process of crafting rules). Successes and failures are built-in in meta-analysis including hundreds of cases in various types of common-pool resources (fisheries, forestry, grazing lands, irrigations systems, knowledge resources etc.). The joint comparison of differences and common properties (intercase and intracase) is a major tool of theorisation. Built in an iterative process with the empirical materials, the theoretical framework is a 'multi-tier conceptual map' (Ostrom, 2011: 9) that helps to identify key variables in highly diverse contexts. It integrates the historical specificity of institutions, social actors and biophysical environments. Variety, diversity are starting points in Ostrom's work and lead to a theory of evolutionary institutional processes (Ostrom and Basurto, 2011). Considering this emphasis on specific configurations in time and space, what is the meaning of generalisation? Earl and Potts (2011: 18) put it nicely: in Ostrom's approach 'there is no general solution, not master equation'. Akin to Darwin's work, there is a general framework able to deal with historical diversity (Hodgson, 2001). Generalising does not lie in discovering best practices that could be applied everywhere as deemed by Duflo. Generalisation stands for developing a general framework to produce and analyse relevant situated data, in order to find general architectural principles allowing for localised resilient orders.

Blueprint thinking versus institutional diversity, static versus dynamic efficiency

In her early studies, by scrutinising many instances of common pool resources governance systems, Elinor Ostrom (in Aligica, 2003) first sought to reveal a set of optimal and universal rules of governance. Realising that she couldn't find any, a major breakthrough was to go beyond panaceas (Ostrom, 2007a) and to acknowledge the intrinsic diversity of socio-economic systems, reflecting itself the variety of local social and environmental circumstances (Ostrom and Basurto, 2011). It implies to draw on local knowledge to grasp the 'unique combinations of variables present in any one system'. 'Involving users directly in this process increases the likelihood of institutions that are well matched to the local physical, economic and cultural environment' (E. Ostrom, in Aligica, 2003: 11). Therefore,

4 This may evolve: a Bayesian hierarchical analysis of 7 microfinance RCTs is in progress (Meager, 2015).

https://doi.org/10.1017/S1744137415000429 Published online by Cambridge University Press

she became critical of 'institutional monocropping' and 'blueprint thinking' Esther Duflo is bordering on. For Ostrom, 'without the capacity to undertake systematic, comparative institutional assessments, recommendations of reform may be based on naive ideas about which kinds of institutions are "good" or "bad"'. (Ostrom, 2011: 9). This points to a limit in Duflo's analysis. First, it entails no comprehensive institutional analysis, only partial assessments on the one or the other institutional tool (to reduce corruption in driving tests for instance). Second, even if she refuses to assess the helpfulness (or the harmfulness) of 'aid' or 'education' in general – as it is often the case for endogenous growth theory and cross-country growth regressions – the locus of 'good' of 'bad' just shifts at a more concrete level of analysis: for the J-PAL some micro-social devices can be intrinsically good or bad for the poor in each given domain (education, nutrition etc.).

A lack of theorisation of the diversity of social systems, as well of their evolution, is manifest in Duflo's work. It is reflected in the way efficiency is assessed. In experiments conducted by the J-PAL efficiency is implicitly understood in a static way, as it is generally the case in mainstream economics: this notion of efficiency is concerned with the allocation of a given quantity of scarce resources at a moment in time (McCartney, 2004). Yet, experiments are conducted in a dynamic world. Some experiments by the J-PAL reveal the sensitivity of the results to the duration of the experiment. One can find illustrations of this regarding micro-credit (Duflo, 2010b: 43) and the promotion of export-oriented crops in Kenya (Ashraf *et al*., 2008). This time-sensitivity is all the more crucial as actors learn, interact and evolve over time, as well as the institutional matrix (Labrousse, 2010). Interestingly, Ostrom is critical of 'static theories of optimal management' (E. Ostrom in Aligica, 2003): 'no social-biophysical system is a static system, and in order to cope with external shocks one needs robust systems that possess considerable redundancy in their capacities to respond and learn from one another' (see also Low *et al*., 2003). This shift to complex adaptive systems implies to understand efficiency in a dynamic way: it is then associated with the capacity to foster resilience and adaptive learning, differing from the optimisation of the allocation of scarce resources at a given instant.

Theory and level of analysis: micro-reductionism versus multilevel analysis

Ostrom and Duflo strongly focus on the microdimensions of development issues. Beyond this common point lays another divergence that comes to light in the way both authors use a house building metaphor to get the picture of socio-economic structures.

Building blocks and emergent properties: the Meccano versus the house-building metaphor

Banerjee and Duflo (2009) acknowledge that the results of their micro experiments cannot simply be extrapolated to the macro-level because of 'general equilibrium effects', when a programme is generalised (see also Acemoglu, 2010).

https://doi.org/10.1017/S1744137415000429 Published online by Cambridge University Press

But there is a primacy of the micro-economic level: 'macroeconomic models are constructed like a Meccano [erector set for children], based on microeconomic building blocks [. . .]. In each case, basic elements are microeconomic elements' (Duflo, 2009: 73–74). Ostrom uses also a building metaphor, yet it leads to different conclusions: 'A house is constructed out of floor joists, roof beams, lumber, roofing material, nails and so forth. When one wants to talk about the house itself, one usually talks about the number of rooms, the style of the house, the number of stories, rather than the number of nails used in construction [. . .]. When one wants to talk about the street on which the house is located, one uses terms such as the size of the lots, the width of the road, the complementarity or lack of complementarity of the building style, and the like. Descriptions of a neighborhood will use still different concepts, as will a description of an urban or rural political jurisdiction in which a neighborhood is located'. (Ostrom, 2005: 11). Duflo describes a house by the number of bricks and nails, whereas Ostrom emphasises the emergent properties of each 'whole' (house, neighbourhood, jurisdiction) resulting from the relationships (complementarities or lack thereof) between the combined elements and the different qualities corresponding to each tier of analysis.

Micro-reductionism versus multilevel analysis

By focusing on the micro-foundations of the macro-economy, by relying on the idea of a representative agent (see its critique by Kirman, 1992), Duflo's methodology does not grasp emerging properties between micro-, meso- and macro-levels. As previously mentioned, randomised experiments are a tool to isolate the effect of a special action favouring development but they do not grasp the interaction between multiple actions. Yet, development is made of non-linear structural processes involving a complex combination of factors, threshold and irreversibility effects (Boyer *et al.*, 1991), cumulative and circular causation (Myrdal, 1957), agglomeration and junction effects (Perroux, 1961), forward and backward linkages (Hirschman, 1958). Duflo is prone to micro-reductionism: explaining phenomena at one strata in terms of the next lowest level strata, i.e. explaining social phenomena in terms of individual psychological factors like procrastination. Her motto is that multiplying small solutions, 'brick by brick', can lead to a 'quiet revolution' and solve incrementally global poverty (Banerjee and Duflo, 2011). As valuable as these piecemeal solutions are, composition effects have to be put back in the picture. Meso- and macro-issues have also a huge impact on the 'lives of the poor' and do matter in development economics: think of the impact of budgetary and monetary policies implemented during the Washington consensus, the evolution of terms of trade or the insertion in global value chains. Thus, all development issues cannot be solved by punctual treatments and micro-devices, even if they are applied to whole countries. It is a far cry from deworming children to developing an entire nation like China. This country contributed massively to the diminution

of the world poverty in the last decades; its development (and mis-development) cannot be understood without putting industrial and macro-economic policies, polycentric governance structures in the picture (Boyer *et al.*, 2011; Fligstein and Zhang, 2011; Heilmann, 2008).

Like Duflo, Ostrom didn't tackle macro- and meso-economic issues. Yet, no micro-reductionism can be found in her work. Her approach is a branch of complex systems theory, as founded by Simon, Hayek or Koestler to name a few. Bringing into play affiliated notions like nested, polycentric systems, holons, emergent properties and feedback loops, Ostrom integrates the multi-level structure of society: 'Explanations occur at multiple levels and different spatial and temporal scales [. . .]. Building on top of the single individual are structures composed of multiple individuals – families, firms, industries, nations, and many other units – themselves composed of many parts and, in turn, parts of still larger structures. What is a whole system at one level is a part of a system at another level' (Ostrom, 2005: 11–12). It is often underlined that she develops a bottom-up approach (Earl and Potts, 2011). Yet, top-down processes are also crucial for Ostrom (2007b: 244), 'system-level outcomes are generated through a series of linked action situations that [. . .] generate both upward and downward causal processes'. Her favourite scope is local but this grassroots level integrates the top-down impact of other levels (and bottom-up feedbacks), a point underscored by Duflo. Underlying these distinct views, Ostrom and Duflo follow separate ways to do science.

3. What is good science? Comparing research strategies

> 'J-PAL's mission is to reduce poverty by ensuring that policy is based on scientific evidence [. . .] J-PAL and its partners are driven by a shared belief in the power of scientific evidence to understand what really helps the poor, and what does not'. (J-PAL's website, October 2014)

> 'Somebody who takes the perspective of an omniscient observer will assume that he can "see" the "whole picture, "know" what is "good" for people [. . .]. Such a presumption is likely to increase proneness to error. Fallible men require reference to decision-making processes where diverse forms of analysis can be mobilized and where each form of analysis can be subject to critical scrutiny of other analysts and decision-makers'. (Ostrom, 1973)

Observation and experimentation: an objective or a social construction?

Far from the deductive-nomological scheme (deduction-prediction-verification) widely used in mainstream economics, both approaches assign a fundamental role to observation: statistical observation in Duflo's work (with local consultations prior to experimental design) and field observation, databases, remote satellite sensing etc. in Ostrom's work. The fieldwork is undoubtedly important in Duflo's praxis. However, the approach remains more top-down

https://doi.org/10.1017/S1744137415000429 Published online by Cambridge University Press

than bottom-up, especially because of the division of labour between leading J-PAL authors – designing and interpreting the experiment – on the one side, and research assistants and local field workers carrying out the experiment on the other side (Jatteau, 2013). This hierarchical organisation inhibits the feedback from the ground. As field actors are *de facto* limited to NGO workers, local leaders and government officials, the 'co-construction of experiments' is rather an elite-oriented participation: it doesn't involve the very subjects of experiments. In her evaluation report for the Swedish International Development Cooperation Agency, Ostrom recommended that 'all those affected by projects – particularly the beneficiaries – should be involved in the evaluation of projects. This will encourage beneficiaries, donor officials and contractors to learn of each other's concerns and from each other's experiences' (Ostrom *et al.*, 2001: 53). In the Workshop's practice, time-consuming qualitative observations on the ground take place allowing to unearth local knowledge, which is often implicit and informal, and thus, invisible to the researcher scrutinising solely a statistical playing field. As a matter of fact, the ways both authors reflect on the role of the observer and on his interactions with the observed diverge. Whereas, Duflo perceives scientific procedures as essentially objective and technical, endorsing science-based paternalism, Ostrom views knowledge building as a fallible, artefactual, evolving, socially constructed process supporting deliberative self-governance.

Against ideology and ignorance: experimentation and the objectivity of the scientist in Duflo's work

Duflo considers the 'randomista' as a neutral and impartial researcher, the spearhead of the 'credibility revolution in empirical economics' (Angrist and Pischke, 2010). In key J-PAL's publications, a positivist temptation is discernible, emphasising RCTs as an objective technique, allowing to settle ideological debates by a kind of 'experimentum crucis' or, more accurately, by series of experiments. Duflo sees RCT's as compelling and denies the interpretative role of the observer: 'Randomized evaluations are rigorous. There is no room for interpretation. Either it works or not. If it doesn't work, one can only try something else', she summarised her method in a 2010 interview.[5] This view is a particular declination of mainstream standards of scientificity, rooted in a technical stance of scientific prowess. For the J-PAL, a true scientist is bias-free, devoid of ideology and ignorance, and thus able to solve the 'three I's problem: ideology, ignorance, inertia'. (Banerjee and Duflo, 2011: 276). Good policies should follow the power of scientific evidence, as was the promise of the 'experimenting society' project from the 1970's (Monnier, 1992: 41). As the socio-historian of statistics Desrosières (2008: 22) wrote: 'statistics was presented as a tool of rationalization in the conduct of human affairs, substituting the rationale of calculus and measure for the arbitrariness of passions and

5 http://www.lejdd.fr/Economie/Actualite/Intellectuelle-de-terrain-166936.

https://doi.org/10.1017/S1744137415000429 Published online by Cambridge University Press

the interplay of power relationships. [...] statistics was involved as a way to promote de-ideologization and objectivation [of social issues]'.

Yet, all in all, occulting the conventions and the tinkering behind statistical techniques, ignoring the active role of the researcher, proves illusory in practice. Far from a simple 'verdict of the data' (Banerjee and Duflo, 2011: 16), interpreting the outcomes involves a delicate, partially open-ended process. It goes beyond the mere issue 'whether a program works or not'. It is indeed tricky to assess what is really put to the test. In the Al Amana micro-credit experiment, the results of the RCT were apparently clear-cut: the micro-finance programme put to the test in rural areas failed (very low take-up rates). Yet, a qualitative study showed that this failure was linked – among other factors – to an ill-timed calendar of repayment, designed to mimic the way Al Amana provided micro-credit in urban areas without taking into consideration the agricultural calendar, a crucial issue in rural areas (Bernard *et al.*, 2012). As follows, an experiment tests the effect of a whole apparatus – corresponding to a bundle of explicit as well as implicit assumptions – so that it is highly delicate to disentangle which one(s) is (are) at work to assess the results. We are facing here a variant of the epistemological problem highlighted by Duhem and Quine: the impossibility to test a hypothesis in isolation because an empirical test of the hypothesis (here: does micro-credit work to alleviate poverty?) requires one or more background (or auxiliary) assumptions (here: the calendar is causally neutral). RCT aims at isolating a pure effect but isolation is tedious and equivocal in practice.

Duflo is right in assessing the superior realisticness of RCT's compared to laboratory experiments. Yet, the very conditions that allow a randomised experiment to isolate a pure effect are also the ones that create artificial conditions *in vivo*. The randomisation protocol remains a scientific artefact, triggering social reactions: possibility of placebo (or Hawthorne effect) and nocebo effects, social rejection of some treatments (Devaux-Spatarakis, 2014; Monnier, 1992). Furthermore, it is not uncommon that field actors refuse the very principle of randomisation, hence the tacit use of alphabetisation in the flagship experiments on worms (Miguel and Kremer, 2004) and on flip charts in schools (Glewwe *et al.*, 2004) 'much cited as evidence in favor of the virtues of randomization': 'Private communication with Michael Kremer has confirmed that, in fact, the local partners would not permit the use of random numbers for assignment [...]. We are then in the world of quasi- or natural experiments, not randomized experiments. [...] alphabetization does not guarantee orthogonality with potential confounders'. (Deaton, 2009: 39). Duflo neglects that a clinical trial is a social construct and not a purely objective technique (Keel, 2011; Marks, 1997). As shown by medical investigations, real-world clinical experiments are repeatedly subjects to deviations from the ideal randomisation process (Schulz, 1995), selective publication strategies, controversial interpretations and 'adverse effects of the industry influence' (De Angelis and Fontanarosa, 2008).

Duflo often lacks reflexivity and comes close to scientism. Good science is seen as a source of certainty and infallible knowledge. Unlike scientific experts, the poor often ignore what is good for them, justifying the 'nudge paternalism' *à la* Thaler and Sunstein she advocates in Poor Economics (Banerjee and Duflo, 2011: 66–70) and in her first Tanner lecture (Duflo 2012): 'A number of outcomes [. . .] should be uncontroversially desirable, and there is a fair amount of scientific evidence for how to achieve them'. For her, 'the paternalism of a government or a foreign aid agency trying to achieve these goals is therefore not equivalent to the paternalism of the French industrial houses'. She compares the economist to a plumber (Duflo 2009: 27–30; 38), understood as a figure of social engineering. 'It is useful to think of the economist not as a pure scientist, but as a skilled technician, an engineer or a plumber'; 'The economist should be a skilled craftsman: [. . .], he can deliver his technical skills' advised Duflo in her inaugural lecture at the Collège de France.[6] This view is in close tune with 'international institutions [that] have promoted a technocratic view of development problems, which sees the economist as a "social engineer" and the scientific apparatus of economics as a device to solve narrowly defined problems of policy advice' (Fourcade, 2006: 177). Beyond her constant plea for modesty, her emphasis on creative trials, Duflo inclines to a technical, mechanistic *Weltanschauung*, as expressed by the above mentioned Meccano metaphor. By means of a good technique, the plumber-economist can mend the social machinery, fix its defects. Interestingly, artisanship is also a notion dear to the Ostroms and a token of modesty, but it is endowed with alternative meanings: the Bloomington school stresses the trial-and-errors, the tinkering processes and the artefacts linked with the work of the artisan in his 'workshop'.

The observer and the observed: fallibility and reflexivity in the Bloomington school

As per the Bloomington school, 'observed "facts" do not speak for themselves, they are "artifacts" which may be well analysed if the researcher is fully aware of the shared rules of the members of the society who create the "artifacts"' (Groenewegen, 2011). Scientific and local knowledge interact, frame and permeate social reality. 'The presence of order in the world', Ostrom (1998) writes, 'is largely dependent upon the theories used to understand the world'. The Bloomingtonian knowledge theory is branded by learning, fallibility and uncertainty: 'Any creature that has unique capabilities for learning and generating new knowledge inevitably faces an uncertain future. Learning and the generation of new knowledge are themselves marks of fallibility. Infallible creatures would have no need to learn and generate new knowledge. Fallible creatures need to accommodate their plans to changing levels of information

6 http://www.college-de-france.fr/media/cha_int2008/UPL31862_inaugurallecture_slides__fran_ais.pdf.

https://doi.org/10.1017/S1744137415000429 Published online by Cambridge University Press

and knowledge'. (Ostrom (1980) in Aligica and Boettke, 2011: 39). Knowledge is adaptive, evolutionary and rooted in changing local contexts. Since ambiguity is inherent to communication and language (of the observer and the observed), knowledge is ambiguous.

Mirroring the local actors he observes, the observer is himself a fallible learner with a bounded interpretative rationality, embedded in epistemic communities and following specific professional rules and routines. He interacts with the observed and can learn a great deal from them. According to Vince Ostrom (in Aligica, 2003), 'the researcher or observer needs to take into account the way people think about and experience themselves and their situation. [...] As I said we need to address problems of institutional analysis and development with methods that allow to penetrate social reality rather than distance ourselves from it. A critical dialogue between the observer and those being observed can reduce the potential for observer error'. The J-PAL's trust in scientific knowledge – provided it relies carefully on the best techniques – contrasts with Bloomingtonian reservations. Insisting on the necessity to draw on local knowledge, the later often went against the grain of received expert wisdom: 'Some of the lessons coming out of our institutional analyses [...] show that resource users who have relative autonomy to design their own rules for governing and managing common-pool resources frequently achieve better economic (as well as more equitable) outcomes than when experts do this for them' (Ostrom and Basurto, 2011: 319). There is no such thing as a technocratic and paternalistic view on politics: 'Instead of presuming that optimal institutional solutions can be designed easily and imposed at low cost by external authorities, I argue that "getting the institutions right" is a difficult, time-consuming, conflict-invoking process. It is a process that requires reliable information about time and place variables as well as a broad repertoire of culturally acceptable rules'. (Ostrom, 1990: 14). 'The Ostroms join the company of those who reject the notion that institutional design should be the exclusive object of a profession, a specialization, and expertise; on the contrary they consider it a core element of citizens' competence and as such a basic ingredient of self-governance and democracy' (Aligica, 2013: 165).

Relationships to social and 'hard' sciences

'Just as randomized trials for pharmaceuticals revolutionized medicine in the 20th Century, randomized evaluations have the potential to revolutionize social policy during the 21st' (Duflo and Kremer, 2003: 32).

Mimicking 'hard' sciences? A single gold standard or a controlled combination of methods?

Duflo tries to mimic clinical research (Banerjee and Duflo, 2011: 8) but eschews the rich internal debates about evidence-based medicine. As already mentioned, her posture is analogous to that of a social engineer. Ostrom is critical of

mechanistic views and warns of the danger of an 'engineering bias': 'ignoring social infrastructure [...], property rights or social arrangements' (E. Ostrom in Aligica, 2003). Nonetheless, even if 'there are important specific differences between social sciences and the natural sciences', engineering and even more biology can inspire social scientists: 'The biologists' scientific strategy involves identifying for the simplest possible organism in which the process under investigation occurs in a clarified, or even, exaggerated form. The organism is not chosen because it is representative of all organisms. Rather, the organism is chosen because particular processes can be studied more effectively using this organism than using another. These cases are in no sense a "random" sample of cases. Rather, these are cases that provide clear information about the processes involved' (E. Ostrom in Aligica, 2003: 14). Political economy is seen as an evolutionary science in a way akin to Veblen, Nelson and Winter. Significantly, Ostrom's efforts to practice interdisciplinarity went well beyond social sciences.

'Randomized experiment is the best method' insisted Duflo in her lecture at the Collège de France.[7] She defines RCT as the gold standard, largely ruling out other methods, especially cross-country growth-regressions and case studies, producing supposedly 'wishy-washy evidence' (Banerjee, 2007). In evidence-based medicine and policy, RCTs are viewed as providing the highest level of evidence for impact assessment. Nonetheless, other types and levels of evidence are used and can even prove more relevant depending on the subject matters (Berriet-Solliec *et al.*, 2014; Laurent *et al.*, 2009). In the medical art, RCT is but one of the tools participating in the highly complex medical reasoning processes, involving multiple heuristics of exploration and diagnosis, characterised by uncertainty and singularities (Fagot-Largeault, 2010; Masquelet, 2006). In contrast, there is a kind of methodological monism in the J-PAL approach: RCT amounts to a technical panacea, which is applied to all kinds of micro issues, from the empowerment of women to local corruption, from health and nutrition, to education, credit, entrepreneurship or workforce mobilisation. One could label it a 'one-technique-fits-and-fixes-all-issues'. Still, Duflo overlooks that RCTs are only suited for a limited range of development programmes (Bernard *et al.*, 2012): 'treatment-control methods are best suited to address "tunnel"-like issues characterized by a clearly defined and stable "treatment", a rather short and event-proof causal chains, and a large share of targeted individuals that are effectively affected by the intervention [...] Yet, most development interventions do not satisfy such pre-requisites'. Ostrom (Poteete *et al.* 2010) integrates this issue of the suitability of a given scientific instrument and are critical of the very idea of a single gold standard: 'Often, scholars follow "the rule of the hammer" and apply a single method indiscriminately, regardless of its suitability for a given research project. [...] To rely on a single approach is to be shackled [...] To overcome the limits of any one method, one needs to draw

7 http://www.college-de-france.fr/media/inf_pre/UPL1096_2_Pr_sentation_EvalAleatoire.pdf.

https://doi.org/10.1017/S1744137415000429 Published online by Cambridge University Press

on multiple methods [...]. If social scientists have shared standards, no single method fully addresses all standards'. It seems that this 'rule of the hammer' – a hint at Maslow's famous formula 'if all you have is a hammer, everything looks like a nail' – applies to the J-PAL (Bernard *et al.*, 2012).

A more qualitative approach is often lacking in the randomistas' research process. As put by Rodrik (2009), 'the "hard evidence" from the randomized evaluation has to be supplemented with lots of soft evidence before it becomes usable'. Qualitative analysis can help to choose treatments relevant for local populations, to identify possible hurdles, misapprehensions and deviations from the protocol in the conduct of the experiment, and to uncover the motives of behaviour of the population and the causal processes behind the experiment outcomes (Labrousse and Zamora, 2013). One can find a few attempts to link qualitative and quantitative material in recent experiments (Morvant-Roux *et al.*, 2014). Yet, in practise, the 'working together' of teams with distinct scientific cultures can be undermined by misunderstandings and tensions, as exemplified by the SKY-RCT in Cambodia investigated by Quentin and Guérin (2013). The shared belief in the power of hard numbers uniting the J-PAL, associated with the neglect of other forms of scientificity and evidence, is a hindrance to such endeavours. By contrast, ethnography and other exploratory techniques are crucial for the Bloomington school. Its menu of qualitative methods of data collection techniques includes ethnography, participant observation, interviews, oral histories and archival research. 'Neither theory nor methodological techniques substitute for a thorough familiarity with the data, gained from diagnostic tests and data exploration. [...]. Data exploration draws attention to potential causal heterogeneity, non-linear relationships, interaction effects and other aspects of the data that are obscured by more sophisticated multivariate techniques. Thus thorough data exploration contributes to theory testing and development by complementing more sophisticated forms of data analysis and drawing attention to empirical patterns that call out for theoretical explanation'. (Poteete *et al.*, 2010: 14). A thorough qualitative exploration is missing in the fieldwork by the J-PAL, even though their familiarity with the field is far more developed than for many economists, who never went 'there'.

Potlach interdisciplinarity or soft economic imperialism?

Following a kind of pragmatist stance, research is, according to Ostrom, a collective and evolutionary process requiring the collaboration of multidisciplinary research teams to overcome some limitations of both individual rationality and disciplinary specialisation, through skill associations and collective knowledge building within a shared framework (Poteete *et al.*, 2010). Ostrom performs a Simonian 'potlatch interdisciplinarity' (Ostrom, 2007b: 239–240): 'Gift-giving between economics and the other social sciences can become a genuine exchange, going in both directions. The metaphor of a potlatch, rather than one of imperialism, best describes what we have achieved

in the relationship between economics and political science'. Like Ostrom, Duflo practices forms of interdisciplinarity but it amounts to a kind of soft economic imperialism. It is indeed not the neo-classical imperialism *à la* Gary Becker relying on 'the combined assumptions of maximizing behaviour, market equilibrium, and stable preferences, [. . . which] form the heart of the economic approach' (Becker, quoted in Hirshleifer, 1985: 52): Duflo and Banerjee's *Poor economics* is often at odds with neo-classical economics, with the idea of self-regulated markets, the maximising behaviour of self-interested individuals with stable preferences.[8] Duflo's 'universal grammar' is the RCT technique or, more accurately, its positivist use embedded in a mainstream habitus (trust in hard numbers and social engineering, discount of historical, ethnographic material and 'soft' methodologies etc.). She applies this technical grammar to multifarious domains attached to other social sciences (sociology, anthropology, education etc.), ignoring frequently their concepts and methods, overlooking decades of accumulated research on social structures and representations governing health, water adduction, politics, farming systems etc.

Ostrom's ambition to re-embed economics in social sciences is explicit: 'My academic career has been devoted to the development of empirically grounded theories to cross the great divide between economics and other social sciences' (Ostrom in Aligica, 2003: 7). Her work is a systematic attempt to transcend the dichotomy of modern political economy (Ostrom, 2007b). While Ostrom underlines the cost of the separation of political economy into two disciplines, Duflo and Banerjee struggle 'against political economy', embracing a peculiar definition: 'Political economy is the view that politics has primacy over economics: institutions define and limit the scope of economic policy' (Banerjee and Duflo, 2011: 237). Targeting Acemoglu and Robinson, this academic struggle inflicts a collateral damage: a variety of longstanding and fruitful traditions related to political economy, to the enabling role of institutions and the social and political embeddedness of the economy is disregarded.

4. The great development divide: two episteme

Following this methodological comparison, it appears that, despites broad commonalities, Duflo and Ostrom views are exemplars of diverging social philosophies and episteme (see Foucault, 1980: 197), that were not obvious at first instance.

While both authors subscribe to realism, they practise two types of realism. The realism supported by Duflo is akin to a naive 'metrological realism' as defined by Desrosières (2008), in which quantification is seen as merely mirroring reality within a margin of error, whereas Ostrom seems closer to critical realism and constructivism: the way we perceive and quantify reality is moulded by our

8 She underlines notably psychological bias, without developing a Simonian view on rationality.

https://doi.org/10.1017/S1744137415000429 Published online by Cambridge University Press

cognitive maps and conventions. The rationales of the social scientist and of the economic actors are also distinctive. Whereas, Duflo underlines the objectivity and rightness of the scientist applying sound techniques – which contrasts with the lack of information and the restrained horizon of local actors – Ostrom emphasises the processual, bounded and interpretative rationality of both the researcher and the observed actors. This leads to diverging views and normative agendas regarding development, politics and economics.

Duflo sees development as the implementation and replication of expert-led fixes to provide basic goods for the poor who are often blinded by their exacting situation. It is a technical quest for certainty and optimal measures in a fairly static framework. For the Ostroms, there are no best practices, only a few architectonic principles to build locally resilient orders. They view development as a situated learning process under uncertainty. As encapsulated by Aligica and Boettke (2011: 41), 'knowledge and learning are stabilizers of social order and drivers of social change'. Development is notably rooted in the creative but fallible artisanship of ordinary people. It involves joint problem solving associating diverse collective actors (citizens groups, firms, governmental units, international organisations etc.). Development is a co-production process (Ostrom, 1996), a learning process and an institutional process. Aligica (2005: 164) captures it nicely: 'A viable national economic system needs robust flexible and open institutions and multi-level governance systems that allow for learning and increase adaptive capacity without foreclosing future development options. The concrete policy corollary is that instead of a restricted focus on this or that institution and organization in isolation, one needs to produce a social environment that facilitates the learning and adaptation of organizations, individuals and institutions. [. . .] development does not come from blueprints imposed by exogenous policy makers, but as part of a broader and more complex learning process. [. . .] Forms of institutional transfer that do not seriously take that into account are doomed to fail or to generate unintended consequences difficult to control. A more feasible and realistic target of a development policy would be to create the institutional prerequisites necessary to increase the learning context of social actors, economic firms and political units alike, i.e. not to maximize a given "economic policy target" neatly defined'. Development is a complex and messy process, devoid of panaceas, a source of dilemmas for indigenous and exogenous actors (see E. Ostrom's analysis of the Samaritan dilemma in Gibson *et al*., 2005). It is threatened by predation, crypto-imperialism (Ostrom, 1999) and lock-in into vicious configurations (Ostrom and Bazurto, 2011).

In Duflo's science-based 'benevolent paternalism', the experimental technique works as an 'anti-politics machine' (Ferguson, 1990), social goals being predefined and RCT outcomes settling ideally ambiguities and conflicts. Real-world politics – disregarding or instrumentalising RCTs – and institutions – resulting from social compromises instead of evidence – are thus often

perceived as external disturbances and constraints to economic science and evidence-based policy. This depoliticising stance is at odds with the significance of political economy for the Ostroms, their emphasis on deliberation to co-construct the aspirations and agencies of communities. While Duflo and Banerjee are in line with a technocratic democracy, the Ostroms sustain a Tocquevillean democratic self-governance. For the latter, institutions emanating from democratic processes, far from being straitjackets, are the core of economic processes. They simultaneously constraint and enable human action: Following Commons, 'an institution is collective action in control, liberation and expansion of individual action' (quoted in Kiser and Ostrom, 1982: 65).

While apparently modest, Ostroms' theory goes far beyond the commons to comprehend composite, polycentric orders. It is a wide-ranging institutional analysis (Aligica, 2013; Chanteau and Labrousse, 2013; Cole and McGinnis, 2015), albeit macro-economic developments are absent – the Ostroms having their base-camp in political science. It is a situated theory, contextualising the human condition in its diversity, while Duflo (2012) tend to naturalise the poor (and the rich), endowing them with homogenous a-temporal characteristics. The Ostroms' morphological and combinatorial logic is also reminiscent of Eucken's (1939) institutional approach striving to reconcile empirics (observed historical diversity) and theory (thinking in orders as combination of forms to penetrate this manifold reality). Duflo is much more on the empirical side, reflecting recent tendencies toward pure empiricism inside mainstream economics. Empirical micro-econometrics is seen as the revolutionary alternative to 'the theory-centric macro-fortress [that] appears increasingly hard to defend' (Angrist and Pischke, 2010: 19). Duflo might be dissatisfied with standard neo-classical theory but she seems little aware of alternative theories or afraid to cross the mainstream line. So, her approach remains under-theorised, impeding the cumulation of experimental knowledge.

Both authors are pragmatic but in different respects. Duflo is pragmatic in the common sense of the word: she is looking for 'what works', assessing the immediate practical consequences of an action rather than drawing action from theory or dogma. Ostrom is more in line with the philosophical pragmatism of James, Dewey and Peirce (Aligica, 2013; Groenewegen, 2011), who influenced Veblen and Commons, two important references in Ostrom's work. Her conception of the public is directly related to the one by Dewey (McGinnis, 2011), an author that left a profound mark on Vince Ostrom (Aligica, 2013); the Ostrom's view on scientific communities reminds of Dewey's community of enquiry. The role of recalcitrant matter of fact, of puzzles is much alike the abductive reasoning expressed by Peirce (1905) or, in Ostrom's (1997: 99) words: 'An openness to uncertainty, social dilemmas, anomalies, and puzzles as presenting problematics, allows for learning, innovation, and basic advances in knowledge to occur'. The Ostroms' theoretical pluralism is akin to the pragmatist idea that they are many sound ways to conceptualise the world. As shown by

https://doi.org/10.1017/S1744137415000429 Published online by Cambridge University Press

Aligica (2013: 167), 'pragmatist elements have been a constant tacit presence in their work. [...] A non-positivist, fallibilist and realist epistemology, a social philosophy of knowledge processes and social coordination, a penchant for normative pluralism and experientialism, this is a cluster typically associated with the pragmatist tradition'. Experimentalism and RCTs also have roots in this very tradition: Peirce himself introduced randomisation in experiments in 1883 (Hacking, 1988). Yet, the way RCTs are conducted and interpreted by J-PAL scholars, imbued with positivism, distance them from this epistemic foundation. Positivism is not inherent to RCTs or evidence-informed policy. Other uses display more affinities with pragmatist underpinnings. For instance, Connolly (2015) from the Campbell Collaboration advocates a critical realist epistemological foundation for RCTs and to shift the discourse from 'what works' to more nuanced accounts of 'what works? For whom? In what contexts?', by using mixed methods and involving the practitioners extensively.

Furthermore, both Duflo and Ostrom use the figure of the artisan to describe the task of the social scientist, again with distinct meanings: The first sees the economist as a plumber and a social engineer fixing social devices, the other as a tinkering craftsman patiently learning from trial and errors and extracting knowledge from local uses. Despite the fact that RCT proves mostly relevant for the small subset of simple, short-run, micro interventions, Duflo and her colleagues established it as an all-purpose device. Their success is consonant with the dominant conception of economics as a tool-centred knowledge relatively insensitive to historical and geographical variations (Hirschman, 1981), with the decline of specificity-oriented paradigms, notably in development economics and with the weakening of area studies within the economic profession (Fourcade, 2006: 160). Technicity substitutes for scientificity, a drift that Ostrom (1982, preface) diagnosed and criticised early in the case of political science. For her, 'the language of data analysis and method [should not] dominate the language of theory construction' (Aligica, 2013: 169). This contrasts with Duflo's technical reductionism, which is sometimes associated with psychological reductionism. Interdisciplinary in the mainstream is currently tantamount to connexions with behavioural psychology (or neuro-economics). It is because of such preconceptions, together with the tenures of senior J-PAL members in top-flight economic departments, that J-PAL economics can be labelled as mainstream.

Duflo's claim of RCT as the gold standard, her methodological monism diverges with Ostrom's emphasis on the bounded relevance and fruitfulness of any method and the associated necessity of a controlled combination of methods. Moreover, Duflo's approach is congruent with the dilution of development matters into a myriad of micro-issues in the wake of the Post-Washington consensus and the Millennium development goals (Labrousse, 2010). It goes also well together with the trend towards piecemeal analyses witnessed by mainstream economics since the end of the 1990s, i.e. the period when Duflo's first research

took place. On the other hand, Ostrom's interdisciplinary framework was initiated in the 1960's and grew within a dynamic but 'provincial' political science faculty, a discipline more sensitive to historicity than economics; her emphasis on structures, complexity and evolution may be influenced by the episteme of genetic structuralism *à la* Piaget culminating in the 60–70's and the concomitant emergence of complex system theory, i.e. a crucial period in the development of the IAD framework. Ostrom's relative marginality measured up to Ivy League economics found expression in the avalanche of economists' harsh comments on her Nobel prize, standing out against the chorus of praise that followed Duflo's Clark Medal one year later. Hence, these approaches follow diverging pathways that carry traces of their respective emergence contexts.

References

Acemoglu, D. (2010), 'Theory, General Equilibrium, and Political Economy in Development Economics', *Journal of Economic Perspectives*, **24**(3): 17–32.

Aligica, P. (2003), *Rethinking Institutional Analysis and Development: The Bloomington School. Interviews with Vincent and Elinor Ostrom*, George Mason University, Arlington (VA).

Aligica, P. (2005), 'Institutional Analysis and Economic Development Policy: Notes on the Applied Agenda of the Bloomington School', *Journal of Economic Behavior and Organization*, **57**(2): 159–165.

Aligica, P. (2013), *Institutional Diversity and Political Economy. The Ostroms and Beyond*, Oxford: Oxford University Press.

Aligica, P. and P. Boettke (2011), 'The Two Social Philosophies of Ostroms' Institutionalism', *The Policy Studies Journal*, **39**(1): 29–49.

Angrist, J. D. and J. S. Pischke (2010), 'The Credibility Revolution in Empirical Economics: How Better Research Design is Taking the Con out of Econometrics', *NBER Working Paper*, 15794, Cambridge (Mass.).

Ashraf, N., X. Giné, and D. Karlan (2008), 'Finding Missing Markets (and a disturbing epilogue): Evidence from an Export Crop Adoption and Marketing Intervention in Kenya', http://poverty-action.org/sites/default/files/FindingMissingMarkets.pdf (accessed the January 28, 2014).

Banerjee, A. (2007), *Making Aid Work*, Cambridge, MA: MIT Press.

Banerjee, A. and E. Duflo (2009), 'L'approche Expérimentale en économie du Développement', *Revue d'économie Politique*, **119**(5): 691–726.

Banerjee, A. and E. Duflo (2011), *Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty*, New York: Public Affairs.

Bernard, T., J. Delarue, and J. D. Naudet (2012), 'Impact Evaluations: A Tool for Accountability? Lessons from Experience at Agence Française de Développement', *The Journal of Development Effectiveness*, **4**(2): 314–327.

Berriet-Solliec, M., P. Labarthe, and C. Laurent, (2014), 'Empirical validity of the evaluation of public policies: models of evaluation and quality of evidence', *Evaluation*, **2**(2): 195–213.

Boettke, P., L. Palagashvili, and J. Lemke (2013), 'Riding in Cars with Boys: Elinor Ostrom's Adventures with the Police', *Journal of Institutional Economics*, **9**(4): 407–425.

Boyer, R., B. Chavance, and O. Godard (eds.), (1991), *Les Figures de L'irréversibilité en Économie*, Paris: EHESS.
Boyer, R., U. Hiroyasu, and A. Isogai (eds.), (2011), *Diversity and Transformations of Asian Capitalisms*, London: Routledge.
Bureau, M.-C., F. Sarfati, J. Simha, and C. Tuchszirer (2013), 'L'expérimentation dans l'action publique', *Travail et Emploi*, **135**: 41–55.
Chanteau, J. P. and A. Labrousse (2013), 'L'institutionnalisme Méthodologique d'Elinor Ostrom: Quelques Enjeux et Controverses', *Revue de la régulation*, **14** Autumn http://regulation.revues.org/10555.
Cohen, J. and P. Dupas (2007), 'Free Distribution or Cost-Sharing? Evidence from a Randomized Malaria Prevention Experiment', *Brookings Global Economy and Development Working Paper 11*, December.
Cole, D. and M. McGinnis (eds.), (2015), *Elinor Ostrom and the Bloomington School of Political Economy*, Lanham: Lexington Books.
Connolly, P. (2015), 'The Campbell Collaboration, and the application of evidence-based practice in education', *Workshop on Evidence Based Policy*, Académie des Sciences, Paris, 11th September.
Crawford, S. and E. Ostrom (1995), 'A Grammar of Institutions', *American Political Science Review*, **89**(3): 582–600.
Crawford, S. and E. Ostrom (2005), "A Grammar of Institutions", in E. Ostrom (ed.), *Understanding Institutional Diversity*, Princeton NJ: Princeton University Press: 137–174.
De Angelis, C. and B. Fontanarosa, (2008), 'Impugning the Integrity of Medical Science: The Adverse Effects Of Industry Influence', *Journal of the American Medical Association*, **299**(15): 833–1835.
Deaton, A. (2009), 'Instruments of Development: Randomization in the Tropics, and the Search for the Elusive Keys to Economic Development', Manuscript, Princeton, January 2009.
Desrosières, A. (2008), *L'argument Statistique*, I & II, Paris: Presse de l'École des Mines.
Devaux-Spatarakis, A. (2014), *La méthode expérimentale par assignation aléatoire : un instrument de recomposition de l'interaction entre sciences sociales et action publique en France ?*, PhD Thesis in Political Science, Bordeaux: University of Bordeaux.
Duflo, E. (2003), 'Poor but Rational?', *MIT Working Paper*, Cambridge, MA: MIT Press, http://econ-www.mit.edu/files/747, last accessed February 15, 2013.
Duflo, E. (2009), *Expérience, Science et Lutte Contre La pauvreté*, Paris: Fayard.
Duflo, E. (2010a), *Le Développement Humain. Lutter Contre la Pauvreté* (I), Paris: Seuil.
Duflo, E. (2010b), *La Politique de L'autonomie. Lutter Contre la Pauvreté* (II), Paris: Seuil.
Duflo, E. (2012), 'Human Values and the Design of the Fight against Poverty. Paternalism *versus* Freedom?', *Tanner Lecture*, May 2, Harvard University.
Duflo, E. and M. Kremer (2003), 'Use of Randomization in the Evaluation of Development Effectiveness', Paper prepared for the World Bank Operations Evaluation Department, Conference on Evaluation and Development Effectiveness, Washington DC, 15–16 July.
Earl, P. and J. Potts (2011), 'A Nobel Prize for Governance and Institutions – Oliver Williamson and Elinor Ostrom', *Review of Political Economy*, **23**(1): 1–24.
Eucken, W. (1939), *Die Grundlagen der Nationalökonomie*, 6th ed., Berlin-Heidelberg: Springer-Verlag.
Fagot-Largeault, A. (2010), *Médecine et Philosophie*, Paris: PUF.

Ferguson, J. (1990), *The Anti-Politics Machine: 'Development', Depoliticization and Bureaucratic Power in Lesotho*, Cambridge: Cambridge University Press.
Fligstein, N. and J. Zhang (2011), 'A New Agenda for Research on the Trajectory of Chinese Capitalism', *Management and Organization Review*, **7**(1): 39–62.
Foucault, M. (1980), *Power/Knowledge: Selected Interviews and Other Writings, 1972–1977*, New York: Pantheon.
Fourcade, M. (2006), "The Construction of a Global Profession: The Transnationalization of Economics", *American Journal of Sociology*, **112**(1): 145–195.
Gibson, C. C., K. Andersson, E. Ostrom, and S. Shivakumar, (2005), *The Samaritan's Dilemma: The Political Economy of Development Aid*, Oxford: Oxford University Press.
Glewwe, P., M. Kremer, S. Moulin, and E. Zitzewitz (2004), "Retrospective vs. prospective analyses of school inputs: the case of flip charts in Kenya," *Journal of Development Economics*, **74**(1), 251–68.
Gomel, B. and E. Serverin (2013), 'L'expérimentation Sociale Aléatoire en France en Trois Questions', *Travail et Emploi*, **135**: 57–71.
Groenewegen, J. (2011), 'The Bloomington School and American Institutionalism', *The Good Society*, **20**(1): 15–36.
Hacking, I. (1988), 'Telepathy: Origins of Randomization in Experimental Design', *Isis*, **79**(3): 427–451.
Heilmann, S. (2008), 'Policy Experimentation in China's Economic Rise', *Studies in Comparative International Development*, **43**(1): 1–26.
Hirschman, A. O. (1958), *The Strategy of Economic Development*, Yale: Yale University Press.
Hirschman, A. O. (1981), *Essays in Trespassing: Economics to Politics and Beyond*, Cambridge: Cambridge University Press.
Hirshleifer, J. (1985), 'The Expanding Domain of Economics', *American Economic Review*, **75**(6): 53–56.
Hodgson, G. (2001), *How Economics forgot History*, London: Routledge.
Janssen, M., F. Bousquet, and E. Ostrom (2011), 'A Multimethod Approach to Study the Governance of Social-Ecological Systems', *Natures, Sciences, Sociétés*, **19**(4): 382–394 www.nss-journal.org.
Jatteau, A. (2013), 'Expérimenter le Développement ? Des économistes et leurs Terrains', *Genèses*, **4**(93): 8–28.
Keel, O. (2011), *La Médecine des Preuves - Une Histoire de L'expérimentation Thérapeutique par Essais Cliniques Contrôlés*, Montréal: Presses Universitaires de Montréal.
Kirman, A. (1992), 'Whom or What Does the Representative Individual Represents?', *Journal of Economic Perspectives*, **6**(2): 117–136.
Kiser, L. and E. Ostrom (1982), 'The Three Worlds of Action: A Metatheoretical Synthesis of Institutional Approaches', in E. Ostrom (ed.), London, Sage, 179–222.
Miguel, E. and M. Kremer (2004), 'Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities', *Econometrica*, **72**(1): 159–217.
Labrousse, A. (2010), 'Nouvelle économie du Développement et Essais Cliniques Randomisés : Une Mise en Perspective d'un Outil de Preuve et de Gouvernement', *Revue de la régulation*, **7**, Spring, http://regulation.revues.org/7818.
Labrousse, A. and P. Zamora (2013), 'Expérimentations de Terrain et Politiques Publiques du Travail et de l'emploi. Apports récents et mises en perspective – Introduction', *Travail et Emploi*, **135**: 5–13.

Laurent, C., J. Baudry, M. Berriet-Solliec, M. Kirsch, D. Perraud, B. Tinel, A. Trouvé, N. Allsopp, and P. Bonnafou (2009), 'Pourquoi s'intéresser à la notion d'evidence-based policy?', *Tiers-Monde*, **200**(4): 853–873.

Lee, A. and M. L. Done (1999), 'The Use of Nonpharmacologic Techniques to Prevent Postoperative Nausea and Vomiting: A Meta-Analysis', *Anesthesia and Analgesia*, **88**(6): 1362.

Low, B., E. Ostrom, C. Simon, and J. Wilson (2003), 'Redundancy and Diversity: Do They Influence Optimal Management?', in F. Berkes, J. Colding, and L. Folke (eds.), *Navigating Social-Ecological Systems: Building Resilience for Complexity and Change*, New York: Cambridge University Press, pp. 83–114.

Marks, H. (1997), *The Progress of Experiment: Science and Therapeutic Reform in the United States, 1900–1990*, Cambridge: Cambridge University Press.

Masquelet, A. (2006), *Le Raisonnement Médical*, Paris: PUF.

McCartney, M. (2004), 'Dynamic *versus* Static Efficiency: The Case of Textile Exports from Bangladesh and the Developmental State', *Post-autistic Economics Review*, **26**, August, http://www.paecon.net/PPAEReview/issue26/McCartney26.htm.

McGinnis, M. D. (2011) 'An Introduction to IAD and the Language of the Ostrom Workshop: A Simple Guide to a Complex Framework for the Analysis of Institutions and Their Development', *Policy Studies Journal*, **39**(1): 169–183.

Meager, R. (2015), 'Understanding the Impact of Microcredit Expansions: A Bayesian Hierarchical Analysis of 7 Randomised Experiments', *Preliminary Working Paper*, MIT, September 9.

Monnier, E. (1992), *Évaluations de l'action des Pouvoirs Publics*, Paris: Economica.

Morvant-Roux, S., I. Guérin, M. Roesch, and J. Y. Moisseron (2014), 'Adding Value to Randomization with Qualitative Analysis: The Case of Microcredit in Rural Morocco', *World Development*, **56**, 302–312.

Myrdal, G. (1957), *Economic Theory and Underdeveloped Regions*, London: Duckworth and Co.

Ostrom, E. (1965), *Public Entrepreneurship: A Case Study in Ground Water Basin Management*, PhD Thesis in political science, Los Angeles: UCLA.

Ostrom, E. (ed.), (1982), *Strategies of Political Inquiry*, London: Sage.

Ostrom, E. (1990), *Governing the Commons: The Evolution of Institutions for Collective Actions*, Cambridge: Cambridge University Press.

Ostrom, E. (1996), 'Crossing the Great Divide: Coproduction, Synergy, and Development', *World Development*, **24**(6): 1073–1087.

Ostrom, E. (1998), "A behavioral approach to the rational choice theory of collective action: Presidential address", *American Political Science Review*, **92**(1): 1–22.

Ostrom, E. (1999), 'Coping With Tragedies of the Commons', *Annual Review of Political Science*, **2**(1): 493–535.

Ostrom, E. (2005), *Understanding Institutional Diversity*, Princeton, NJ: Princeton University Press

Ostrom, E. (2007a), 'A Diagnostic Approach for Going Beyond Panaceas', *Proceedings of the National Academy of Sciences*, **104**(39): 15181–15187.

Ostrom, E. (2007b), 'Challenges and Growth: The Development of the Interdisciplinary Field of Institutional Analysis', *Journal of Institutional Economics*, **3**(3): 39–264.

Ostrom, E. (2011), 'Background on the Institutional Analysis and Development Framework', *Policy Studies Journal*, **39**(1): 7–27.

Ostrom, E., C. Gibson, S. J. Shivakumar, and K. Andersson (2001), *Aid, Incentives, and Sustainability: An Institutional Analysis of Development Cooperation*, Stockholm: Sida.
Ostrom, E., R. Parks, and G. Whitaker (1977), *Policing Metropolitan America*, Washington DC: US Government Printing Office.
Ostrom, E. and X. Basurto (2011), "Crafting Analytical Tools to Study Institutional Change", *Journal of Institutional Economics*, **7**(3): 317–343.
Ostrom, V. (1973), 'Order and Change Amid Increasing Relative Ignorance', *Working Paper n°W73-1*, Bloomington: Indiana University.
Ostrom, V. (1980), 'Artisanship and Artifact', *Public Administration Review*, **40**(4): 309–317.
Ostrom, V. (1997), *The Meaning of Democracy and the Vulnerabilities of Democracies*, Ann Arbor: University of Michigan Press.
Ostrom, V. (1999), 'Crypto-Imperialism, Predatory States, and Self-governance', in M. McGinnis (ed.), *Polycentric Governance and Development*, Ann Arbor: University of Michigan Press, pp. 166–185.
Peirce, C. S. (1905), *The Collected Papers of Charles Sanders Peirce*, C. Hartshorne, P. Weiss (eds.), Cambridge, MA: Harvard University Press, vol. 5 (1935).
Perroux, F. (1961), *L'économie du XX^e siècle*, Paris: PUF.
Poteete, A., M. A. Janssen, and E. Ostrom (2010), *Working Together: Collective Action, the Commons, and Multiple Methods in Practice*, Princeton: Princeton University Press.
Quentin, A. and I. Guérin (2013), 'La Randomisation à l'épreuve du Terrain – L'exemple du Projet de Microassurance Sky au Cambodge', *Revue Tiers Monde*, **213**(1): 179–200.
Ramsey, S. and J. Scoggins (2008), 'Practicing on the Tip of an Information Iceberg? Evidence of Under-publication of Registered Clinical Trials in Oncology', *The Oncologist*, **13**(9): 925–929.
Rodrik, D. (2009), "*The New Development Economics: We Shall Experiment, But How Shall We Learn?*", in J. Cohen and W. Easterly, (eds)., *What Works in Development? Thinking Big and Thinking Small*, Washington D.C.: Brookings Institution Press.
Schulz, K. F. (1995), 'Subverting Randomization in Controlled Trials', *JAMA: Journal of the American Medical Association*, **274**(18), 1456–1458.