Can RCTs help improve the design of CAP?

Luc Behaghel¹, Karen Macours¹ and Julie Subervie²

¹ Paris School of Economics, INRA
² INRA, CEE-M

Abstract

We illustrate how RCTs could be used as a learning tool to shed light on various aspects of the Common Agricultural Policy (CAP). Based on examples drawn from agricultural and social policies in developing and developed countries, we argue that the RCT toolbox has the potential to significantly add to existing approaches to evaluating and designing the CAP.

Keywords: Common agricultural policy; Impact evaluation; Policy design; Field experiments

JEL codes: C93; Q18

1. Introduction

Despite the renewed demand for evaluating agricultural policies in the European Union and the revival of field experiments – also known as randomized controlled trials (RCTs) – in the evaluation of social policies in Europe since the mid-2000s, RCTs have not been applied to the evaluation of the Common Agricultural Policy (CAP) (Colen et al., 2016). This may seem surprising given the role played by RCTs during the “glory days” of agricultural economics until the 1960s (Herberich, Levitt and List, 2009). There are undoubtedly many reasons for this absence, both on the demand side (e.g. the type of evaluations requested or accepted by various stakeholders such as governments, international agencies, farmers, …) and on the supply side (e.g. the incentives available to and the backgrounds of the consultants and academics operating in this area). This paper abstracts from these constraints - each of which undoubtedly deserve careful analyses of their own - and asks a more preliminary question: are RCTs an appropriate tool to address issues relevant to the design of today’s CAP?

While evaluation can serve as both an accountability measure and as a learning tool, this paper focuses on its role as the latter, as RCTs can be particularly powerful instruments in improving our understanding of effective policies. For many aspects of the CAP, the important policy question to
address is not whether a certain type of policy ought to exist, but rather whether specific changes to the design of an existing policy could lead to better efficiency and/or equity outcomes. To the extent that such questions are of interest, RCTs can help provide rigorous answers by randomly introducing potentially relevant variations in the design and testing their implications for a policy’s effectiveness. We consider the two pillars of the CAP, and highlight a (non-exhaustive) list of policy-relevant design issues that could arguably benefit from rigorous empirical evidence: (i) How could the uptake of agri-environmental measures by farmers be increased? (ii) What schemes can induce coordination among farmers on environmental issues in which one farmer’s defection jeopardizes the efforts of all the others such? (iii) What implications could different mechanisms for direct payments have on efficiency and equity? (iv) How can we design agri-environmental contracts to ensure that farmers receive sufficient compensation for their conservation efforts without inducing large efficiency losses? We then discuss how RCTs can shed light on these questions.

The four policy design questions are meant to demonstrate the adaptability of RCTs: are they useful only for a narrow range of issues? It seems hardly surprising that, transposing and extending lessons drawn from the lab by behavioral economists, experiments on information-based interventions or other types of nudges can shed light on a farmer’s propensity to adopt agri-environmental contracts. One may be skeptical, however, about what can be learned from RCTs with respect to three other issues, namely coordination failures, equity – efficiency tradeoffs, and market design. Despite this, we believe that one of the key lessons that emerges from the recent literature that relies on field experiment is the strong adaptability of the RCT toolkit and its usefulness in addressing questions that had been deemed out of reach only a few years ago. To illustrate this claim, we complement the discussion of hypothetical experimental designs that could be applied to the CAP with actual examples from two related policy domains: agricultural policies in developing countries, and social policies in Europe and other contexts.

Our exercise contributes to a small prospective literature on the role that experiments can play in the evaluation of agricultural policies. Heberich et. al (2009) build on observations over the past century to argue that agricultural economists could take “a new train at the depot,” in the form of field experiments inspired by laboratory experiments and economic theory. Colen et al. (2016) provide a detailed review of the state of the art in the evaluation of the CAP, and discuss the comparative advantages of different experimental methods, and their complementarities with other methods. Both papers take an optimistic view of the potential of the RCT toolbox while acknowledging that this potential has not yet been realized. To complement these discussions, we highlight potentially fruitful

---

1 The first pillar consists of direct payments to farmers and market measures, and the second pillar is a rural development policy aimed at balancing territorial development with sustaining a farming sector that is environmentally sound, while simultaneously fostering competitiveness and innovation.
examples of evaluations, inspired by recent successes in neighboring areas. The subsequent sections consider the four policy design questions in turn, and the last section offers several concluding remarks.

2. Inducing behavioral changes: nudges and information-based experiments

Farmers are involved in decisions with complex trade-off dynamics and sometimes lack all of the relevant information when considering engaging in new agricultural practices. The complexity and changing nature of some of the environmental regulations and policies, as well as the heterogeneity in conditions and constraints among many farmers in the EU, may well imply that some farmers are not necessarily aware of all the information and practices that may be relevant for their particular situation. This potentially is a bigger constraint among smaller (potentially part-time) farmers or certain farmers in newer member states. Whether such information constrains actually exists, and if so for whom, is clearly an empirical question. Targeting the provision of information regarding eligibility for certain benefits, or advantages of shifting practices, to certain types of farmers could help evaluate such questions.

Even when farmers have access to all the relevant information, the complex and dynamic trade-offs they need to consider for many of their decisions may fosters the use of heuristics and a susceptibility to behavioral inertia and the status quo in agricultural practices. Insights from behavioral science suggests that “nudges” (Thaler and Sunstein, 2008) can then by potential powerful tools to shift decision making. A nudge is not a cash payment. It may consists, for instance, of the provision of information about the social norm (about what others think or do), possibly in conjunction with a particular program. It can also simply consist of the way in which a choice is presented. Empirical evidence shows that nudges can be effective in changing consumer behavior, in particular pro-environmental behaviors (see Schubert (2017) for a review). The question then arises: is it possible to nudge farmers to change their practices, and possibly adopt more pro-environmental (or “greener”) practices?

There are indeed some reasons to believe that AES uptake could be improved with the use of nudges. Some studies suggest that nudges can affect farmers’ intentions to (re-)enroll in AES (Kuhfuss et al.

2 Since 2008, the Dephy farm network of the French Ecophyto program has been experimenting with techniques that reduce the use of pesticides and herbicides. The dissemination of these innovative practices is a major challenge faced by the program.
There is also evidence of a “social identity” mechanism behind individual decisions (Goldstein et al 2008), by which an individual adheres to the descriptive norms of the group of which he considers himself a member. This effect may well apply to farmers, especially among members of a cooperative. Thus, there is reason to believe that this type of psychological lever could be useful in improving the design of agri-environmental policies.

In recent years, behavioral economic research using lab experiments has grown rapidly and has provided evidence of a large number of tools that can be used to encourage green practices. Because these results cannot be directly extrapolated to the field, field experiments are often necessary to validate results derived from the lab. In this respect, the use of lab-in-the-field experiments seems appropriate as an intermediate step to undertake before implementing an RCT, as they can contribute to a better understanding of the role played by context, an essential consideration when implementing any field experiment.

Despite a recent craze for nudges in both the academic and the public sphere (in UK and US in particular), there are almost no experimental studies providing evidence of the impact of nudges on farmers’ decisions in the real world. One exception is a study by Messer et al. (2015), who ran a field experiment in which farmers from Texas, Delaware, and Maryland competed in an auction of conservation contracts that required them to adopt practices that reduced nutrient run-off. The authors show that changing the default option can result in larger bids, and that providing information about what others think of the required practice increases the likelihood that a producer participates in the auction. Wallander et al (2017) and Chabé-Ferret et al (2018) also use social norms in field experiments with the aim of inducing greater farmer participation in AESs, and they make use of RCTs to test the effectiveness of these nudges. Both studies failed to detect a statistically significant impact of these interventions on farmer participation in the AE schemes.

Many alternative nudges remain to be tested in the field in order to improve the design of agri-environmental policies. Ferraro et al. (2017), for example, suggest using a default option in order to increase farmer participation in the US Conservation Reserve Program. The authors also suggest presenting payments in agri-environmental programs using a loss frame, specifically by stating the maximum payment the participant could earn and how much he would lose for every practice not adopted. Both of these suggestions could also work in the European context. Previous evidence thus suggests that insights from behavioral sciences could indeed be used to improve the design of the second pillar of the CAP.

---

3 Several governments in developed countries have constituted ‘behavioral insights teams’ within their civil services. In 2014, the U.S. Department of Agriculture (USDA) created the Center for Behavioral and Experimental Agri-Environmental Research (CBEAR). Galizzi (2017) also identifies initiatives by the governments of Australia, Canada, Denmark, Finland, France, Israel, the Netherlands, New Zealand, Norway and Singapore.
For a variety of information interventions or nudges, the evaluation of the scheme’s effectiveness using an RCT seems quite feasible, since large samples can be easily reached and the so-called Stable Unit Treatment Value Assumption (SUTVA) is likely to hold, at least in cases where the nudge or information is customized to each farmer. One limitation of such experiments, however, is that small (albeit valuable) effects may be hard to detect, as the record of the “What Works Centres” in UK shows. Another limitation is that the interpretation of these effects is not always straightforward in the presence of a variety of plausible “behavioral” models and heterogeneous types of behavior (e.g. Duflo et al., 2011). But even if the impact of information nudging is small, it may remain quite cost-effective, as these types of interventions tend to require very little in terms of implementation costs. This is all the more true in countries such as France, where the online application for agricultural subsidies has been mandatory since 2016, and the cost of providing information or nudges through pop-up windows in the Telepac website would thus be virtually zero.

3. Incentives to coordinate: experimenting with collective bonuses

Agri-environment schemes (AES) have become a key mechanism within the CAP for sustaining the environmental services provided by farmland. A number of different schemes have been put in place at the level of individual farms. However, the recent literature has emphasized the potential gains that could be realized by adopting an approach that is characterized by a larger, landscape-wide scale. Accordingly, AES compensation payments are now allowed by EU regulation to be paid to groups of farmers (Regulation N° 1305/2013, article 28, cited in Westerink et al., 2017). Such payments, often referred to as "collective bonuses", raise important design questions. Contract theory highlights the various potentially counterproductive mechanisms at play. An obvious rationale for setting incentives at the landscape level is that many of the outcomes for which the schemes are designed (e.g., wildlife conservation, water quality and storage) are also applicable at the landscape level. While actions at the individual farm level could certainly contribute to these objectives, their impact on landscape-level results would arguably be weaker, thus yielding imperfectly aligned incentives. Employing "multi-tasking" models (Holmstrom and Milgrom, 1991) may also induce farmers to focus on better-incentivized tasks at the expense of other tasks that could be more important with respect to the environmental objectives identified. Collective bonuses can also create collective action problems, such as free-riding. The free-rider problem can, however, be mitigated to some extent by institutional arrangements or peer pressure. In short, the appropriate design of collective bonuses is a difficult theoretical question involving numerous plausible mechanisms whose effects may be context
dependent. Empirical evidence is needed in order to weigh the relative strengths of these mechanisms and to assess their net effect. Relatedly, given the evidence in several instances that incentives may have perverse effects, transparent and convincing evidence that collective bonuses have the expected positive effects in the contexts of interest will be key to their political viability.

Although to our knowledge RCTs have not yet been used to evaluate agri-environmental schemes at the landscape level, two examples suggest that they may prove a useful tool in this regard. In Uganda, Jayachandran et al. (2017) analyze the impact of financial incentives for forest owners to maintain the integrity of their forestland, thus providing experimental evidence of the effectiveness and cost-effectiveness of Payments for Ecosystem Services. Even though the incentives in this study are set at the farm level, where each individual is given the opportunity to enroll to receive payments if he refrains from clearing trees, the randomization of treatment assignments takes place at the village level such that the impacts that are measured encompass any "leakages" (externalities) and collective dynamics within villages. The authors also use random variation in the proximity between treatment and control villages to account for potential spillovers across villages. As a result, outcomes in this study are measured at the landscape level. Though there is no evidence of spillovers and limited evidence of leakages, program enrollment is low (32 percent), seemingly due to limited program awareness, leading the authors to question whether the program could be better marketed. Collective incentives could indeed be a way to do so, as they could induce farmers to advertise the program to other farmers in their village. In other words, everything in this experimental design is in place to analyze alternative collective schemes, and the results suggest that such schemes are indeed worth trying.

The second example comes from the incentive literature in the economics of education. Quite naturally, given the collective nature of the education production function at the classroom and school level, researchers have been experimenting with various forms of individual and collective bonuses for teachers and students as incentives to increase student performance. Empirical evaluations in this literature underscore the complexity of the mechanisms at play, and specifically their dynamic nature. For instance, Muralidharan and Sundararaman (2011) compare the effectiveness of teacher- and school-level incentives in India: while the two compare well in the first year, teacher-level incentives outperform school-level incentives in the second year, suggesting collective action problems might have come into play. Other systematic investigations such as those conducted by Roland Fryer or John List suggest that the literature in the economics of education remains on the steep segment of the learning curve regarding how to best harness the power of incentives in this area (e.g., Fryer et al., 2012). Much could certainly be learned from similar efforts with respect to AES.
4. Fighting farmer poverty: experimenting with transfers?

The first pillar of the CAP has long consisted of direct payments to farmers, providing a safety net against volatile prices. Such subsidies are a core component of the CAP and are largely uncontested as a principle of the program. Prima facie, it may therefore appear that there is no useful role for RCTs in the evaluation of these policies. Yet beyond the normative questions regarding the principle of the payments per se, many questions exist regarding the specific design of the different interventions within the first pillar, over which member countries do exercise some degree of freedom. Ensuring that transfers reach the targeted groups without inducing large efficiency losses matters for the sustainability of the CAP over the long term. The equity considerations underlying the payments of the first pillar are also receiving more attention given differences in allocations between older and newer member states.

Incidentally, RCTs have proved to be a very useful tool in shedding light on the advantages and disadvantages of direct payments to farmers and other rural households in developing countries. The evaluation of aconditional cash transfer program in Mexico (*PROGRESA*, now named Prospera) was one of the first large-scale RCTs conducted in the developing world, and the experimental evidence of its impacts was crucial in assuring continued support through changing political climates. In fact, the program rapidly expanded nationwide in Mexico and triggered similar programs across Latin America, where they now reach 25 percent of the total population. Such programs have also spread to Africa, Asia, and even Europe and the United States over the last 20 years (Fiszbein and Schady, 2009; Robles, Rubio and Stampini, 2015).

The initial evaluation of the program in Mexico, which tested the effectiveness of one particular design, was followed by a large literature experimentally comparing alternative design variations. These studies helped policymakers understand, for instance, whether making transfers conditional, labeling them, or making them entirely unconditional impacted their effectiveness (Baird, Ozler and McIntosh, 2011; Benhassine et al, 2015). They also tested whether varying the recipient in the household (men vs. women) affected outcomes (Akresh, de Walque and Kazianga, 2016; Almas et al 2018), and whether making payments conditional on actions or achievements was more effective (Barrera-Osorio et al, 2011). Such studies have now been carried out in many different settings (e.g. Malawi, Morocco, Burkina Faso, Bosnia, Colombia), and frequently in close collaboration with government ministries. The agencies implementing these studies often have valid questions regarding how to optimally design a transfer program for a specific objective or context.

Once it is acknowledged that the answers to such questions are not obvious beforehand, experimentation becomes an ethical policy procedure. While such RCTs do not call into question the
benefits themselves, nor the justification for the policy goals they aim to achieve, they are well suited to provide empirical evidence on potential trade-offs between different objectives that alternative design options may imply. Learning from these experiences suggests that RCTs have potential as useful instruments in comparing alternative direct payment schemes under the first pillar of the CAP, and in investigating how different designs may imply different trade-offs between the possible objectives therein.

5. Targeting under adverse selection: experimenting with innovative agri-environmental contract designs

Voluntary agri-environmental contracts that offer fixed payment schemes suffer from two major problems. First, farmers who face the lowest costs for complying with environmental requirements are more likely to enter the program; in cases where the program would pay some farmers for doing nothing differently from what they would have done in the absence of payment, adverse selection may induce large windfall effects (Chabé-Ferret and Subervie 2013). Second, those with the highest costs of participating are less likely to enter the program, though they are precisely those whose engagement would have the greatest contribution to the program’s effectiveness (Kuhfuss et al. 2014). One option for reducing the efficiency losses due to adverse selection is to shift from fixed payment schemes to auction mechanisms. Procurement auctions have been in place in other countries for many years (e.g. the US Conservation Reserve Program, established in the 1980s, as well as several pilot programs in Australia). However, the context in which conservation auctions are implemented may impact their effectiveness (Lundberg et al. 2018; Ferraro 2008). It is therefore impossible to anticipate the additional gains from auctions (compared to fixed payment contracts) in the European context; here again, RCTs could provide a useful way to address this question.

For several decades, researchers have used experimental auctions to estimate consumer willingness-to-pay for new products (Corrigan et al., 2009). In a demand-revealing auction mechanism similar to the Vickrey (1961) and Becker-DeGroot-Marschak (BDM) (1964) auctions, bids indeed provide a direct measure of auction participants’ willingness-to-pay for the good being sold. There is a substantial literature dealing with the implementation of such designs in university experimental economics labs (Berry et al. 2015 and references therein). More recently, a number of studies in developing countries have also demonstrated that these insights can have broader relevance for addressing design trade-offs in actual policy applications. Hoffmann, Barrett and Just (2009) used a BDM design to measure the gap between willingness-to-pay and willingness-to-accept for bed nets in Uganda. Berry, Fischer and
Guiteras (2015) also use a BDM mechanism to estimate the willingness to pay for water filters in Ghana, while Guiteras and Jack (2017) use the same mechanism to investigate how workers respond to different contractual arrangements in the context of informal day labor markets in rural Malawi. Finally, and more closely related to the topic we address here, Jack (2013) uses a Vickrey auction to explicitly take into account landholders’ willingness-to-accept an afforestation contract in Malawi. She moreover runs a RCT to demonstrate that landholders who received a tree planting contract as a result of bidding in the auction kept significantly more trees alive over a three-year period than did landholders who received the contract through a lottery. Much would certainly need to be learned from a systematic analysis of the demand for AES in European countries, especially taking into account the heterogeneity that is likely to be present with respect to this demand.

6. Concluding comments

The lessons from this prospective exercise can be summarized as follows. First, many insights from laboratory experiments in the behavioral sciences could be used to improve the design of the second pillar of the CAP, and RCTs would be useful in evaluating how well these approaches can contribute to achieving higher participation rates in real world conditions. Second, the value of collective incentives is a challenging theoretical question due to the many mechanisms that are potentially at play in determining the ultimate impacts of these incentives. RCTs may therefore be of particular use in weighing the relative strengths of plausible mechanisms in specific field settings. Third, ensuring that transfers reach the targeted groups without inducing large efficiency losses is an important consideration in fostering the sustainability of the CAP; this question, too, can be usefully addressed by RCTs. Finally, the potential for adverse selection is high in current schemes where farmers receive the same payment regardless of the opportunity costs of conservation they face. Here again, experiments can be used for both eliciting individual willingness-to-accept an agri-environmental contract and for separately identifying the effects of farmer selection and payment on the provision of environmental services.

In this paper, we strongly advocate for experimentation, aiming to demonstrate that RCTs have significant potential for improving the design of the CAP. We do not view RCTs as a way of evaluating the CAP per se, but rather as a tool for evaluating different design alternatives for which there are no obvious ex ante expectations. We moreover argue that RCTs could be used as a tool to test new schemes with complementary interventions in an effort to enable the CAP to reach its target audience and objectives.
The lack of studies on this topic may indicate the presence of political economic issues – such as the fear that the tool may do a disservice to its promoters and damage the credibility of CAP itself. In this context, it could be useful to follow the guidance offered by Campbell and Stanley (1969) in order to avoid such conflicts and misperceptions: real decisions often imply choosing between plan A and plan B rather than putting the global architecture of a policy in question; when such choices can be made following a documented and transparent process, the global architecture may be continuously improved, and the policy reinforced.

References


Ferraro, P. J. (2008), 'Asymmetric information and contract design for payments for environmental services', Ecological Economics 65(4), 810 - 821.


