Using Experiments to Design and Evaluate the CAP: Insights from an Expert Panel

Le recours à des expérimentations pour concevoir et évaluer la PAC : perspectives d'un groupe d'experts

Experimente zur Gestaltung und Bewertung der GAP: Erkenntnisse eines Expertengremiums

Daniele Curzi, Sylvain Chabé-Ferret, Salvatore Di Falco, Laure Kuhfuss, Marianne Lefebvre and Alan Matthews

New challenges ahead call for new instruments to evaluate the CAP

Over the last two decades, the agricultural economics research literature has considerably evolved. The focus of earlier studies on CAP impacts was mostly on agricultural production, prices on agricultural markets, incomes and social welfare (Thoyer and Préget, 2019). However, instruments used by agricultural policies in Europe and elsewhere have changed dramatically, in part due to the shift of emphasis toward new issues such as climate change.

> **"** Les expériment-ations économiques ont des caractéristiques uniques qui en font un instrument supplément-aire potentiellement important dans la boîte des outils économiques traditionnellement utilisés pour concevoir et évaluer la PAC. **"**

Within that context, there has been increasing interest in using experimental approaches to respond to these changing needs for agricultural policy design and evaluation. The use of experiments has been an essential component of the so-called 'credibility revolution' in empirical economics. Policymakers are more and more interested in the insights coming from experimental methods, and they are increasingly willing to use these results to implement large-scale public policies.

Despite this interest, the use of these innovative methodologies to analyse the CAP is rather scant. The European Association of Agricultural Economics conference held in August 2021 hosted a panel discussion aimed at providing a broad overview of the potential contributions and limits of experimental tools for CAP design and evaluation. We summarise here the main insights that emerged during this session, starting from a brief presentation of the different experimental approaches, followed by some examples of their application in the study of the CAP. Finally, we discuss the potential hurdles to the use of experimental results to design actual policies, and some potential solutions to these problems.

What is the quantitative economist's toolkit?

The analysis of policy interventions can rely on different methods:

- Randomized Control Trials (RCTs), where in the population participating in the experiment there is a random selection of individuals receiving the intervention, while the other individuals not receiving the treatment represent the control group. Since these two groups are identical, any difference in a given outcome of interest is due to the intervention (i.e. treatment). Within this framework we can find both laboratory and field experiments, where individual behaviours are studied in particular settings aimed at capturing their response to the experimental design.

- Natural experiments, which are studies set in non-controlled conditions, which rely on the randomness of policy implementation to try to mimic the conditions of RCTs without explicit randomisation.

- Observational methods, which try to account explicitly for differences between people receiving the policy and people not receiving the policy without resorting to randomisation or to a natural experiment.

- Structural models where there is a use of explicit economic modelling to predict the effect of various policy scenarios. Choice experiments are a special type of

structural model in which participants in the experiment are asked hypothetical questions about their preferences for various policy options that can be used to fit a model that can help predict the take up of various policy options.

> **"** Ökonomische Experimente haben einige einzigartige Eigenschaften, die sie zu einer potenziell wichtigen Ergänzung des traditionellen wirtschaftswissen-schaftlichen Instrumentariums zur Gestaltung und Bewertung der GAP machen. **"**

All these methods require assumptions to provide reliable conclusions. RCTs have some unique characteristics, as the crucial assumptions needed to properly test the effect of a policy are enforced by the design of the experiment. This is the reason why RCTs are considered as the gold standard to give the most convincing answers to research questions aimed at analysing the effect of an intervention (e.g. a policy). The results of RCTs can indeed be used to form credible estimates of the cost-effectiveness and benefit-cost ratios of various programmes, enabling policymakers to select the most efficient policy options.

An important characteristic of RCTs is that they allow for surprises in the results: the impact of the intervention may be different from the expectations. If the results of the experiment suggest that the intervention tested is not very effective, the policy may be changed before its actual implementation. RCTs are therefore particularly useful during the design and piloting phase of an intervention; as a tool for iterative evaluation, since they can provide an early understanding of potential problems, which could be then addressed in the second phase of the policy implementation. RCTs are also very useful in that they can serve to benchmark the other type of methods and to quantify their biases.

Some examples of the use of experimental methods to improve the CAP

Although experiments may provide additional insights to the understanding and design of the CAP, only a few contributions so far have made use of some of these methodologies. Experiments have shown that nudging, i.e. a non-monetary incentive aimed at influencing individual decisions through the provision of information or a change in the framing of the information provided (Thaler and Sunstein, 2008), may represent a cost-effective tool to drive farmers to adopt greener practices. An example is given by a RCT conducted in France, where 10,000 information letters describing Agri-environmental schemes (AES) were sent to randomly selected farmers (Chabé-Ferret *et al.*, 2021). The experiment resulted in 73 new AES contracts among the participants. Clearly, this is not a big number *per se*, but what is interesting is the cost-effectiveness of the programme. Considering the cost for sending letters to the participants, the cost for each new AES contract was about 90€. Thus, it is possible to consider this programme very cost-effective. Other interesting aspects emerging from this experiment involve potential spillover effects of the programme, which further increase its cost effectiveness. This is because other farmers not participating in the experiment have actually signed new AES contracts after receiving information about the programmes from their neighbours.

Another RCT conducted among 200 farmers in South-West France was aimed at testing whether social comparison represents a way to nudge farmers into saving water (Chabé-Ferret *et al.,* 2019). In this experiment, each farmer received weekly information about her individual as well as group water consumption. The results revealed that the programme was effective in reducing water consumption, especially among those farmers who made more use of water for irrigation. However, and quite surprisingly, the experiment also suggested that the information sent to farmers led to an increase in water consumption among those who did not consume water at all before the programme started.

Nudging has also been used in an experiment aimed at testing whether social norms may be a driver of farmers' decision-making process when dealing with the decision of whether to maintain practices after the end of the programme subsidising them (Kuhfuss *et al.,* 2016a). The experiment was conducted using an online survey among about 400 farmers. The results suggest that farmers are more likely to maintain AES practices when informed that others would also do that. This shows that the social norm is important in affecting farmers' decisions, and therefore nudges can be an important way to achieve this goal. However, a more recent experiment (Kuhfuss *et al.,* 2022) suggests that the effect of some nudges may be insignificant in comparison with monetary incentives.

One important issue related to the implementation of AES is related to their cost-effectiveness. In this respect a recent choice experiment measured farmers' preferences for participation in results-based payment schemes (Tanaka *et al.,* 2022). Although conducted in Japan, the experiment is still informative in the EU context. The idea behind the experiment is that, paying farmers based on the achievement of positive environmental results instead of paying them for the adoption of specific practices, could be more effective to reach the environmental target. It would also give farmers more flexibility to choose the practices that are more suitable to their farm characteristics. However,

farmers may also perceive this as a risky option. Therefore, it is important to measure *ex-ante* farmers' preferences and likely participation. The findings of the experiment suggest that preferences for results-based payment schemes are heterogeneous. Despite uncertainty regarding payments, some farmers would be willing to participate. However, the participation rate decreased when payments were conditioned to higher environmental objectives. When results-based schemes were associated with an eco-label, then the level of payments required by farmers to participate was lowered, pointing towards the potential synergies between public incentives and labelling approaches that would need to be further investigated.

From experiments to policy

What happens to experimental results when scaling-up the intervention? Given the potential role of experiments for better policy understanding and design, it is important to understand whether experimental results are actually confirmed when the policy tested in an experimental setting is applied to the general population. The results from the literature trying to address this question present mixed evidence. Some policies seem to scale up very well, for instance in the case of education and social protection policies. However, there is large and disconcerting evidence of many experimental results failing when mapped into a more general framework. This is the case for instance with many environmental/energy programmes, especially in developing countries that were not successful when implemented on a large scale, as had been expected from the results of the small-scale experiments. This might also be the case of some nudges when implemented at scale, as demonstrated recently by DellaVigna and Linos (2022). Clearly, these events are very disappointing: first, because many of these experiments are very costly and thus, when failing to scale up, it suggests a

The use of information letters in an economic experiment proved to be a cost-effective way to improve farmers' participation in agri-environmental schemes.

waste of financial resources; second, there is also reputational damage, which may raise the scepticism of policymakers toward experimental results; third, and perhaps more importantly, the research fails to achieve its target and this can have negative consequences for the economic welfare of participants.

What are the threats to the scalability of experimental results?

There are three broad reasons why things may fail at scale (List, 2022).

1. The existence of a 'false positive'. This may happen when an experiment reveals an incorrect favourable result. This effect is compounded by publication bias, where only statistically significant results are published in academic journals. As a result, the published academic literature gives an overoptimistic estimate of the impacts of some programmes.

2. The effect of the intervention diminishes when moving from small scale to large scale. This has been called the voltage drop effect, which has been frequently

observed in the implementation phase of research. This potential problem raises concern about the cost-effectiveness of these experiments. This problem may occur because the quality of the intervention varies when moving from small-scale to large-scale population. This may be due to significant differences between the actual intervention with respect to the experiment carried out within an academic environment; or because stakeholders have not received proper guidance on how to customise the intervention within the community-based population without compromising its cost-effectiveness.

3. The population and the situation under investigation in the experiment are not representative. This may happen when there are some participants' characteristics that make them particularly predisposed to exhibit a stronger behaviour than in the population at large. This may happen for instance when people are asked if they want to participate in the experiment, and these people may be more driven to do so. Sometimes in field experiments there is an 'encouragement design', where participants decide whether or not they want to be part of the experiment. In this case, there is not true randomisation of the treatment; what is actually randomised is the information about the treatment. Another problem is that sometimes experiments involve graduate students, who have a strong incentive to comply with the experimental protocol. When the experiment is then mapped out in the general public the compliance with the experimental design may reduce significantly.

How to use insights gained from experiments in order to inform agricultural policy design?

In order to avoid, or at least limit, failures in the scalability of the experimental results, there are a number of potential solutions that may be adopted. First, the

Economic experiments have been used to test approaches for water savings in agriculture.

implementation of a series of well-powered (ideally independent) replications of the experiments. This would empower experimental findings, and policymakers would probably have higher incentive to apply the experimental design to a general population. A potential problem is that researchers must be given the incentive to replicate experiments, as it can be difficult to receive funding for replications from agencies or policymakers. An additional problem is that the system of incentives for academics should change as well. This is because there is no incentive to work on the replication of an experiment; it is not seen as a valuable exercise and its publication potential is much lower than that of the original experiment.

To improve the scalability of experimental results, an important issue to consider is the conformity of the general policy to the original experimental programme. For political reasons or other constraints, when moving from small scale to large scale there are often some changes in the intervention, which makes it hard to confirm the scalability of the effect. At the same time, taking into consideration the constraints faced by policymakers on the ground is a key requirement to be able to conduct large-scale experiments. Conducting large coordinated experiments across all of Europe, with a network combining

researchers and policymakers might be a way to deal with these issues. In this respect, the initiatives that are part of the REECAP (Research Network on Economic Experiments for the Common Agricultural Policy) network might enable us to get there.

Another important issue to consider when designing an experiment or when assessing its results, is the understanding of heterogeneous effects; namely the extent to which an experiment may have effects that differ across the involved individuals. This requires experiments to look not only at the average treatment effect, but also to estimate the interactions that may be of interest; and thus to understand the effects that experiments may have on the population where there may be some different relevant features. However, these kinds of analyses should be conducted on very large samples to be meaningful.

How can experiments help improve the CAP further?

Experimental methods potentially constitute an important addition to the traditional economics toolkit used to study the CAP. Several challenges remain and we present here four main research avenues.

1. To address the problem of the scalability of small-case experimental results, in addition to the above-mentioned insights, it is important to conduct

experiments with larger samples, in a wider range of contexts. This is even more important in the new CAP delivery model, where there is substantial flexibility in the policy application at the Member State level, and where thus it would be important to have a better understanding of the effect of a policy under different scenarios. Experimental results on different options could contribute to the elaboration of national variations of the EU policy.

2. The results of different experiments should be aggregated through meta-analyses (a statistical technique for combining data from multiple studies on a particular topic), which provide a useful tool for improving the reliability of experimental findings. Experiments should be part of the policy evaluation toolbox as a complement to other methods (Cohen *et al.*, 2016). With a better integration of experiments in larger research projects or evaluation work, one could highlight the value-added of experimental insights: isolating the effect of a policy from other factors due to randomisation, and providing insights into the complex puzzle of farmers' decision-making.

3. While experiments related to the behavioural drivers of farmers' reactions to policies have flourished, there are still a number of unexplored areas. Most of these experiments concern agri-environmental policies, while other areas within the CAP also deserve attention. For example, one needs to know more about the drivers and levers for farmers to invest in animal welfare, or farmers' adoption of risk management tools, including the possibility of both self and private insurance to address potential risks.

4. Experiments may also be useful to evaluate monetary incentives related to the CAP. For example, Randomised Control Trials (RCTs) could be used to study whether proposing group-bonuses to farmers when they

Economic experiments can potentially represent an important addition to the traditional economics toolkit used to study the CAP and could be used to inform agricultural policy design.

contract AES are cost-effective (choice experiments have suggested that they could be extremely cost-effective (Kuhfuss *et al.*, 2016b)). These bonuses are supplementary payments to the usual AES contribution, and are given to recipient farmers provided that a certain threshold of collective participation in the policy is reached. When considering this particular payment scheme, it is worth considering farmers' private transaction costs; this is because they need to know each other's compliance costs to determine 'side-payments' from low compliance cost participants to high compliance cost participants in order to ensure their participation in the programme; thus receiving a bonus-payment for the group.

What are the main hurdles to a wider use of experiments to study the CAP?

A more intense use of experiments to study the CAP is clearly conditional on an increase in the acceptability of these methodologies. The inclusion of stakeholders and policymakers in the evaluation process may be helpful in this respect. Stakeholders such as farmers can be the subject of the experiment (those whose decisions are analysed), but they

can also take part in the design phase of the experiment. Pre-design qualitative work with stakeholders is becoming increasingly prevalent, in particular to select attributes in choice experiments, through traditional focus groups or individual interviews (Armatas *et al.*, 2014).

> " Economic experiments have some unique characteristics that make them a potentially important addition to the traditional economics toolkit used to design and evaluate the CAP. "

The first hurdle that we face is the EU's General Data Protection Regulation or GDPR, that regulates what types of data can be collected on participants in experiments. Implementing a field experiment generally requires the constitution of a database of individual farmers, with information on their location and other possibly identifying information. The GDPR, in its Article 6.1(e, f), acknowledges that

research is a sufficient motive for retaining personal data. In its Article 14, the GDPR requires that individuals are informed that their personal data will be collected, and can access, modify and erase it. This requirement can prove hard to implement in large field experiments, especially for the control group, which might not be in contact with the experimenters. Informing the individuals that they are part of an experiment might also affect their behaviour and seriously bias the results of the experiment. Fortunately, Article 14.5.b of the GDPR offers exceptions to the information requirements, when the data are not collected directly from the individuals.

The use, at all stages of research projects, of data management plans that guarantee compliance with the GDPR regulation, while making open science possible, will add to the trust required to allow access to individual level data as well as increase the reliability of experimental results. Another hurdle is to convince citizens and policymakers of the benefits of evidence-based policymaking, where RCTs play a major role. Hopefully, the mounting pressure of the climate and biodiversity crisis will help us focus on finding the most cost-effective solutions. The more RCTs we conduct, even at a small scale, the more we will be able to illustrate the benefits of this approach. The key here is finding policymakers motivated to evaluate the real impacts of their policies. Fortunately, there is an increasing number of these individuals.

Funding Information

Further Reading

■ Armatas, C. A., Venn, T. J. and Watson, A. E. (2014). Applying Q-methodology to select and define attributes for non-market valuation: a case study from Northwest Wyoming, United States. *Ecological Economics*, **107**: 447–456.

■ Behaghel, L., Macours, K. and Subervie, J. (2019). How can randomised controlled trials help improve the design of the common agricultural policy? *European Review of Agricultural Economics*, **46**(3): 473–493.

■ Chabé-Ferret, S., Le Coent, P., Reynaud, A., Subervie, J. and Lepercq, D. (2019). Can we nudge farmers into saving water? Evidence from a randomised experiment. *European Review of Agricultural Economics*, **46**(3): 393–416.

■ Chabé-Ferret, S., Le Coent, P., David-Legleye V. and Delannoy V. (2021). *Can nudges increase the effectiveness of payments for ecosystem services? Evidence from a randomized experiment, mimeo*, Toulouse School of Economics.

■ Cohen, L., Gomez Y Paloma, S., Latacz-Lohmann, U., Lefebvre, M., Préget, R. and Thoyer, S. (2016). Economic experiments as a tool for agricultural policy evaluation: Insights from the European CAP. *Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie*, **64**(4): 667–694.

■ DellaVigna S. and Linos E. (2022). RCTs to scale: Comprehensive evidence from two nudge units. *Econometrica*, **90**(1): 81–116.

■ Kuhfuss, L., Préget, R., Thoyer, S. and Hanley, N. (2016b). Nudging farmers to enrol land into agri-environmental schemes: the role of a collective bonus. *European Review of Agricultural Economics*, **43**(4): 609–636.

■ Kuhfuss, L., Préget, R., Thoyer, S., Hanley, N., Le Coent, P. and Désolé, M. (2016a). Nudges, social norms, and permanence in agri-environmental schemes. *Land Economics*, **92**(4): 641–655.

■ Kuhfuss, L., Préget, R. Thoyer, S., de Vries, F.P. and Hanley, N. (2022). Enhancing spatial coordination in payment for ecosystem services schemes with non-pecuniary preferences. *Ecological Economics*, **192**(107271).

■ Lefebvre, M., Barreiro-Hurlé, J., Blanchflower, C., Colen, L., Kuhfuss, L., Rommel, J. and Thoyer, S. (2021). Can economic experiments contribute to a more effective CAP? *EuroChoices*. **20**(3): 42–49.

■ List J. (2022). *The Voltage Effect*. New York: Currency.

■ Tanaka, K., Hanley, N. and Kuhfuss, L. (2022). Farmers' preferences toward an outcome-based payment for ecosystem service schemes in Japan. *Journal of Agricultural Economics*. Published Online. Available online at: https://doi.org/10.1111/1477-9552.12478.

■ Thaler, R. H. and Sunstein, C. R. (2008). *Nudge: Improving Decisions about Health, Wealth, and Happiness*. New Haven, CT: Yale University Press.

■ Thoyer, S. and Préget, R. (2019). Enriching the CAP evaluation toolbox with experimental approaches: Introduction to the special issue. *European Review of Agricultural Economics*, **46**(3): 347–366.

Daniele Curzi, Assistant Professor, University of Milan, Italy. (iD) https://orcid.org/0000-0002-1596-5126
Email: daniele.curzi@unimi.it

Sylvain Chabé-Ferret, Assistant Professor, Toulouse School of Economics, France.
Email: sylvain.chabe-ferret@tse-fr.eu

Salvatore Di Falco, Full Professor, University of Geneva, Switzerland.
Email: salvatore.difalco@unige.ch

Laure Kuhfuss, Environmental Economist, Social, Economic and Geographical Sciences Group (SEGS), James Hutton Institute, United Kingdom.
Email: laure.kuhfuss@hutton.ac.uk

Marianne Lefebvre, Assistant Professor, University of Angers, France.
Email: marianne.lefebvre@univ-angers.fr

Alan Matthews, Professor Emeritus of European Agricultural Policy, Trinity College Dublin, Ireland.
Email: alan.matthews@tcd.ie

Summary

Using Experiments to Design and Evaluate the CAP: Insights from an Expert Panel

Over the last twenty years the Common Agricultural Policy (CAP) has considerably evolved, by introducing new objectives and instruments to address the increasing number of challenges ahead. These changes call for the use of innovative tools to analyse agricultural policy design and evaluation. During the last European Association of Agricultural Economics conference, a panel of experts presented their points of view on how experiments can enhance the CAP evaluation toolkit. In this article we summarise the main insights emerging during this session. We present a review of the different existing experimental approaches followed by some examples of their application in the study of the CAP and a discussion of the potential hurdles to the use of experimental results to design actual policies. From the different contributions it emerges that experimental approaches may represent effective tools to improve the design and evaluation of the CAP. However, a potential hurdle to their wider use is that experimental results often fail to be confirmed when applied to large-scale policy interventions. In this article we discuss some potential solutions to the main problems affecting the scalability of experimental results, and we provide some insights on how experiments can help to improve the CAP further.

Le recours à des expérimentations pour concevoir et évaluer la PAC : perspectives d'un groupe d'experts

Au cours des vingt dernières années, la politique agricole commune (PAC) a considérablement évolué en introduisant de nouveaux objectifs et instruments pour relever le nombre croissant de défis à venir. Ces changements appellent l'utilisation d'outils innovants pour analyser la conception et l'évaluation des politiques agricoles. Lors de la dernière conférence de l'Association européenne d'économie agricole, un panel d'experts a présenté son point de vue sur la manière dont les expérimentations peuvent enrichir la boîte à outils d'évaluation de la PAC. Dans cet article, nous résumons les principales idées qui ont émergé au cours de cette session. Nous présentons une revue des différentes approches expérimentales existantes, suivie de quelques exemples de leur application dans l'étude de la PAC, et une discussion des obstacles potentiels à l'utilisation des résultats expérimentaux pour concevoir des politiques réelles. Des différentes contributions, il ressort que les approches expérimentales peuvent représenter des outils efficaces pour améliorer la conception et l'évaluation de la PAC. Cependant, un obstacle potentiel à leur utilisation plus large est que les résultats expérimentaux ne sont souvent pas confirmés lorsqu'ils sont appliqués à des interventions politiques à grande échelle. Dans cet article, nous discutons de certaines solutions potentielles aux principaux problèmes de généralisation des résultats expérimentaux, et nous donnons un aperçu de la manière dont les expérimentations peuvent aider à améliorer davantage la PAC.

Experimente zur Gestaltung und Bewertung der GAP: Erkenntnisse eines Expertengremiums

In den letzten zwanzig Jahren hat sich die Gemeinsame Agrarpolitik (GAP) durch die Einführung neuer Ziele und Instrumente zur Bewältigung der immer zahlreicheren Herausforderungen erheblich weiterentwickelt. Diese Veränderungen erfordern den Einsatz innovativer Instrumente zur Analyse der Gestaltung und Bewertung der Agrarpolitik. Auf der letzten Konferenz der *European Association of Agricultural Economics* (EAAE) hat eine Gruppe von Experten ihre Ansichten darüber dargelegt, wie Experimente das Instrumentarium zur Bewertung der GAP verbessern können. In diesem Artikel fassen wir die wichtigsten Erkenntnisse aus dieser Sitzung zusammen. Wir geben einen Überblick über die verschiedenen bestehenden experimentellen Ansätze, gefolgt von einigen Beispielen ihrer Anwendung bei der Untersuchung der GAP und einer Diskussion über die potenziellen Hürden bei der Nutzung experimenteller Ergebnisse für die Gestaltung aktueller Politiken. Aus den verschiedenen Beiträgen geht hervor, dass experimentelle Ansätze wirksame Instrumente zur Verbesserung der Gestaltung und Bewertung der GAP darstellen können. Ein potenzielles Hindernis für ihren breiteren Einsatz ist jedoch, dass experimentelle Ergebnisse bei der Anwendung auf groß angelegte, politische Maßnahmen oft nicht bestätigt werden können. In diesem Artikel erörtern wir einige mögliche Lösungen für die Hauptprobleme, die die Skalierbarkeit experimenteller Ergebnisse beeinträchtigen und wir geben Einblicke, wie Experimente zur weiteren Verbesserung der GAP beitragen können.