12

Experimental Pathways in EU Law

*Juan Antonio Mayoral**

12.1 INTRODUCTION

Inspired by natural sciences and psychology, experiments have become more dominant in the social sciences as they allow researchers to establish strong causal relationships in complex societal contexts and situations, even within the legal domain. While mainstream observational studies are crucial for revealing correlations determining why certain legal phenomena occur, our capacity to control for all the variables that may influence a particular result in real life is still limited.

The virtue of experiments is that they provide a very useful method for testing hypotheses in controlled settings, where the impact of other relevant factors (or variables) is minimised, isolating the effect of our main explanatory variables of interest. As a result, experiments allow researchers to determine whether one factor directly affects another – that is, that a certain cause produces a certain outcome. This is especially valuable in fields such as economics, psychology, and political science, where complex environments and interactions make causality challenging to discern.

As in the social sciences, law has been no exception. The growing field of empirical legal studies[1] is also shifting towards more experimental approaches due to its advantages for causal inference. Pioneered by Law & Economics scholars, predominantly from the USA, this method has travelled to other legal empirical subfields like international law,[2] and more recently to European Union (EU) law. Its

* This research was funded by the European Research Council CURIAFIDES grant (no. 101089083).

[1] Tom Pavone and Juan A. Mayoral, 'Statistics As If Legality Mattered: The Two-Front Politics of Empirical Legal Studies' in Marija Bartl and Jessica C. Lawrence (eds.), *The Politics of European Legal Research* (Edward Elgar 2022) 78.

[2] Jeffrey L. Dunoff and Mark A. Pollack, 'Experimenting with International Law' (2017) 28(4) *European Journal of International Law* 1317.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 23:04:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009672580.015

adoption in EU law has been driven by political scientists, motivated by their specific research interests and the rigorous methodological standards of their discipline.

The first section of this chapter provides a review of the implementation of experiments that captures the incipient interest in empirical studies seeking to understand and analyse the functioning of EU law and institutions, like the Court of Justice of the European Union (CJEU). However, its application is still limited and circumscribed to social science topics and questions, mostly focused on public opinion. The primary objective of this review is to provide a state of the art on these contributions and their methodological nuances of experimental research within the study of EU law. Additionally, it identifies common and emerging trends and research questions for future empirical investigation of EU law.

The next section assesses the extent to which diverse experimental approaches or types may enrich our understanding of legal decision-making and the functioning, performance, and effectiveness of EU institutions and law. This will offer a prospective analysis of the future role of experiments in EU law research and considers potential ways for refinement and expansion to other research themes.

By exploring these issues, this chapter contributes to the methodological advancement and empirical rigour of the multidisciplinary scholarship devoted to the study of EU law. It offers an introduction for scholars seeking to leverage experimental methodologies in their pursuit of nuanced and evidence-based analyses of EU law. Ultimately, this comprehensive presentation intends to serve as a foundational resource, useful for scholars, policy-makers, and legal practitioners, for advancing empirical research and policy design in the dynamic landscape of EU law.

12.2 EMERGING EXPERIMENTAL INSIGHTS IN EU LAW

EU law scholarship frequently uses the term 'experiment' to describe the European Union as an unprecedented and evolving integration project[3] – a 'trial' among sovereign states in which new supranational institutions and regulations are tested to promote peace, governance, economic growth, and well-being. However, the application and understanding of experimentation as a research method within EU legal studies remains underexplored.

This limitation can be attributed, first, to the persistent resistance from European law schools to introduce empirical methods into their curriculum,[4] despite the

[3] Bruno de Witte, 'The European Union as an International Legal Experiment' in Grainne de Burca and Joseph H.H. Weiler (eds.), *The Worlds of European Constitutionalism* (Cambridge University Press 2012) 19; Arthur Dyevre, Monika Glavina, Nicolas Lampach, Michal Ovádek, and Wessel Wijtvliet, 'Uncertainty and Legal Disintegration: Evidence from Brexit' (31 October 2018), available at https://ssrn.com/abstract=3276189.

[4] Pavone and Mayoral (n 1).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 23:04:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009672580.015

emerging efforts coming from diverse universities, research centres, projects, networks, and conferences in promoting empirical legal training and research. A second factor discouraging the use of experimental methods, as argued by Epstein and Martin,[5] is the conditions necessary to organise experiments: researchers randomly select subjects from the population of interest (*random sampling*) to better reflect the diversity of the population. This technique improves the equal distribution of cofounding variables, thus improving the representativeness of the findings. Researchers also randomly assign these selected subjects to treatment and control conditions (*random assignment*) to establish causation by formulating appropriate comparisons.[6]

Experiments on legal matters hardly achieve these standards. First, *random sampling* is not prioritised in lab experiments as researchers recruit university students or volunteers who do not necessarily represent the broader population. Similarly, field experiments often involve studying individuals or specific groups who are difficult to access, such as officials, judges, lawyers, vulnerable groups, or litigants. Challenges arise from professional, institutional, or ethical barriers, such as the lack of time or interest in legal research, which make it difficult to create random samples from certain communities or groups. Consequently, researchers often rely on non-probabilistic sampling[7] alternatives to ensure that key characteristics like gender, education, experience, or ideology are represented in the sample, allowing for more generalisable conclusions.

This limits the generalisability of lab and field experiments, focusing instead on *random assignment* to establish causality. Nevertheless, randomisation is not always easy to achieve in the case of field experiments, as intervening in real-world legal procedures can influence legal outcomes, potentially violating principles of equality, fairness, and justice (e.g., assigning certain litigants to specific legal procedures while withholding them from others).

This explains why the first EU experimental legal studies were conducted outside of courts, focusing instead on public opinion and attitudes towards the functioning of EU law and the institutions shaping it such as the Court or the Commission. These studies have been facilitated by the introduction of online survey platforms, which provide easy access to large and representative samples of the general public while significantly reducing the costs and complexities of implementing random

[5] Lee Epstein and Andrew D. Martin, 'Quantitative Approaches to Empirical Legal Research' in Peter Cane, and Herbert M. Kritzer (eds.), *The Oxford Handbook of Empirical Legal Research* (Oxford University Press 2012) 911.

[6] See James N. Druckman, Donald P. Green, James H. Kuklinski, and Arthur Lupia, 'Experiments: An Introduction to Core Concepts' in James N. Druckman, Donald P. Greene, James H. Kuklinski, and Arthur Lupia (eds.), *Cambridge Handbook of Experimental Political Science* (Cambridge University Press 2011) 15.

[7] Jason D. Rivera, 'When Attaining the Best Sample Is Out of Reach: Nonprobability Alternatives When Engaging in Public Administration Research' (2018) 25(3) *Journal of Public Affairs Education* 314.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 23:04:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009672580.015

assignment to experimental conditions in scenario-based experiments.[8] Nevertheless, the main caveat here is whether the results obtained in a controlled and simplified experiment carried out on a particular sample of citizens are in fact generalisable to the more complex, real-world scenarios and real-world population, since we ask the participants to imagine situations. In this regard, the 'external validity' of the experiment can be ensured by accurately designing scenarios as proxies for real-world legal situations applicable across judges with different backgrounds.[9]

Against this background, an incipient number of political scientists have begun to use experiments as a tool for testing theories of compliance and enforcement of EU law, EU legal interventions, regulatory decisions, and CJEU rulings and their impact in the public.[10] These studies converge around the EU's challenges in enforcing norms and maintaining legitimacy by analysing the complex dynamics between EU institutions, Member States, and citizens. While enforcement is critical for upholding EU norms, these studies underline the importance of mitigating backlash and fostering public support when applying EU remedies. In this regard, they emphasise the need for EU institutions to balance the effective enforcement of legal values, such as judicial independence and rule of law, simultaneously maintaining legitimacy in the eyes of the general public to avoid undermining trust in EU governance.

[8] Shanto Iyengar, 'Laboratory Experiments in Political Science' in James N. Druckman, Donald P. Greene, James H. Kuklinski, and Arthur Lupia (eds.), *Cambridge Handbook of Experimental Political Science* (Cambridge University Press 2011) 73; Dunoff and Pollack (n 2).

[9] Rose McDermott, 'Experimental Methods in Political Science' (2002) 5(1) *Annual Review of Political Science* 31.

[10] Stuart J. Turnbull-Dugarte and Daniel Devine, 'Can EU Judicial Intervention Increase Polity Scepticism? Quasi-Experimental Evidence from Spain' (2021) 29(6) *Journal of European Public Policy* 865; Sivaram Cheruvu and Joshua C. Fjelstul, 'Improving the Efficiency of Pretrial Bargaining in Disputes over Noncompliance with International Law: Encouraging Evidence from the European Union' (2021) 29(8) *Journal of European Public Policy* 1249; Mikael Rask Madsen, Juan A. Mayoral, Anton Strezhnev, and Erik Voeten, 'Sovereignty, Substance, and Public Support for European Courts' Human Rights Rulings' (2022) 116(2) *American Political Science Review* 419; Jonathan Zeitlin, David van der Duin, Theresa Kuhn, Maria Weimer, and Martin Dybdahl Jensen, 'Governance Reforms and Public Acceptance of Regulatory Decisions: Cross-National Evidence from Linked Survey Experiments on Pesticides Authorization in the European Union' (2023) 17 *Regulation & Governance* 980; Øyvind Stiansen, Daniel Naurin, Asimina Michailidou, and Adriana Purzycka, 'Enforcing the Rule of Law in the EU: Effects on Public Opinion' (2024) 32(2) *Journal of European Public Policy* 522; Sivaram Cheruvu and Jay N. Krehbiel, 'Do Preliminary References Increase Public Support for European Law? Experimental Evidence from Germany' (2024) 78(1) *International Organization* 170; Dimiter Toshkov, Honorata Mazepus, Nikoleta Yordanova, and Darinka Piqani, 'Enforcement and Public Opinion: The Perceived Legitimacy of Rule of Law Sanctions' (2024) 32(2) *Journal of European Public Policy* 550; Jerg Gutmann, Jaroslaw Kantorowicz, and Stefan Voigt, 'How Do Citizens Define and Value the Rule of Law? A Conjoint Experiment in Germany and Poland' (2024) 32(4) *Journal of European Public Policy* 899.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 23:04:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009672580.015

By mostly using (scenario-based) survey experiments, these studies investigate whether enforcement actions and rulings influence public attitudes, finding that the content of decisions often outweighs procedural concerns.[11] Such studies also observed the prevalence of personal traits like education, religiosity, and political orientation in EU law enforcement, rather than national or cultural differences. Moreover, it was found that supranational enforcement rarely triggers substantial public backlash that can be strategically mobilised by EU opposers, provided that interventions align with public values or are framed effectively. For example, while additional information about EU actions does little to affect public opinion,[12] public awareness of widespread support for rule of law increases the perceived legitimacy of sanctions.[13] More specifically for courts, studies have shown the extent to which CJEU legitimacy is embedded in the legitimacy and support of their national counterparts, showing how Member States' courts play a crucial role in legitimising CJEU rulings by fostering public support for expansive interpretations of EU law.[14]

While survey experiments have mainly been used to test public reactions to hypothetical scenarios on law enforcement and compliance, some quasi-experimental designs[15] have been implemented to analyse the impact of EU policies[16] or court rulings[17] in real-world settings, improving the external validity or generalisation of the findings compared to artificial scenarios. In this regard, quasi-experiments are more feasible (when observational data is available on the issues) and useful to overcome problems of randomisation due to impractical or unethical reasons (see example above on litigants).[18]

In quasi-experiments, the objects of study (individuals, countries, law cases, policies, etc.) are assigned to groups based on existing conditions, natural events, or non-random criteria.[19] Cheruvu and Fjelstul use a quasi-experimental design on observational data to estimate the effect of the EU Pilot programme on the efficiency of pretrial bargaining during infringement procedures.[20] Using infringement cases as the unit of analysis, they define the participation of certain countries in the

[11] Madsen et al. (n 10).
[12] Øyvind Stiansen et al; Toshkov et al. (n 10).
[13] Toshkov et al. (n 10).
[14] Cheruvu and Krehbiel (n 10).
[15] Yongnam Kim and Peter Steiner, 'Quasi-experimental Designs for Causal Inference' (2016) 51(3–4) *Educational Psychologist* 395.
[16] Cheruvu and Fjelstul (n 10).
[17] Turnbull-Dugarte and Devine (n 10).
[18] Thomas E. Nelson, Sarah M. Bryner, and Dustin M. Carnahan, 'Media and Politics' in James N. Druckman, Donald P. Greene, James H. Kuklinski, and Arthur Lupia (eds.), *Cambridge Handbook of Experimental Political Science* (Cambridge University Press 2011) 201.
[19] Hilke AM Grootelaar and Kees van den Bos, 'Chapter 2: Conducting Experiments and Surveys in the Field of Administrative Justice – On the Importance of Fair Procedures in Governance' in Willem H. van Boom, Pieter Desmet, and Peter Mascini (eds.), *Empirical Legal Research in Action* (Edward Elgar 2018) 38.
[20] Cheruvu and Fjelstul (n 10).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 23:04:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009672580.015

EU Pilot as a treatment.[21] Since Member States self-select into the programme, the treatment is not randomly assigned as it would be in a controlled trial. The treatment group consists of cases involving participants in the EU Pilot and the control group consists of cases involving Member States not participating in the programme. Once the allocation is done, they apply a *difference-in-difference* statistical analysis that allows them to estimate a treatment effect on the Member States participating the in EU Pilot compared with those who are not part of the EU Pilot. While researchers rely on statistical methods to account for these differences produced in quasi-experiments, causal inferences are weaker because groups may differ systematically in ways unrelated to the treatment, introducing potential confounding variables or alternative explanations that might affect the internal validity and reliability of the results compared to experimental designs.[22]

Similarly, Dyevre et al.[23] treat Brexit as a quasi-natural experiment to evaluate how political uncertainty discouraged British litigants and judges to invoke EU law and to refer cases to the CJEU, accelerating the process of legal disintegration. For that purpose, they implemented a *difference-in-difference* design to reproduce the conditions of a randomised experiment by comparing post-intervention change in the dependent variable (referral activity) in the treatment group (British courts) to a control group (courts in the rest of the EU) and compared the UK's referral activity before and after the Brexit referendum. In this design, the treatment group is exposed to a certain policy intervention under the key assumption that, in the absence of a Brexit referendum, the unobserved differences between the treatment and control group will remain constant over time.

Recently, researchers in EU law started to apply experiments to study EU legal decision-making with law students[24] as a strategy to address the challenges of accessing legal professionals. Building on behavioural legal research, Ovádek's paper examines how the framing of legal arguments on the application of EU legislation (e.g., attaching political motivations to them) affects an argument's perceived legitimacy and attractiveness.[25] His 'apolitical hypothesis' suggests that adding political reasons to legal arguments reduces their appeal to legal

[21] According to the EU Pilot website, it 'is a mechanism for informal dialogue between the Commission and the Member State concerned on issues relating to potential non-compliance with EU law. It can be used before launching a formal infringement procedure. Info: https://single-market-scoreboard.ec.europa.eu/enforcement-tools/eu-pilot_en.

[22] Christopher F. Karpowitz and Tali Mendelberg, 'An Experimental Approach to Citizen Deliberation' in James N. Druckman, Donald P. Greene, James H. Kuklinski, and Arthur Lupia (eds.), *Cambridge Handbook of Experimental Political Science* (Cambridge University Press 2011) 258.

[23] Dyevre et al. (n 3).

[24] Arthur Dyevre and Michal Ovádek, 'Experimental Legal Methods in the Classroom' (2020) 16(1) *Utrecht Law Review* 1.

[25] Michal Ovádek, 'The Apolitical Lawyer: Experimental Evidence of a Framing Effect' (2019) 48 *European Journal of Law and Economics* 385.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 23:04:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009672580.015

professionals. The experimental results confirm this hypothesis, showing that a political frame made law students 12–24 per cent more likely to select the 'apolitical' legal option.

Although still in its early stages, all these research contributions exemplify how experiments are gradually gaining relevance and attractiveness in EU empirical legal scholarship. This impulse is driven, in part, by the need for improved methodologies that enhance the explanatory power of our theories on the legal integration of Europe.[26] As we will see in the next section, while experimental design and implementation presents a great challenge and requires reflection on causal relations in the EU legal domain, the discipline is moving towards more experimentation.

12.3 EXPLORING THE UNTAPPED EXPERIMENTAL POTENTIAL IN EU LAW

This section presents different experimental designs available for studies of EU law. It describes their value in aiding our understanding of how EU law and actors operate in practice, as well as their associated challenges, in order to inspire researchers in the field to consider experiments for exploring dynamics in the legal realm.

12.3.1 *Scenario/Vignettes Survey Experiments*

The use of scenarios in survey research is not new in the empirical research of EU law where vignettes have been used to describe hypothetical real-life situations in which judges choose to follow certain courses of action with regards to EU law application.[27] The experimental use of scenarios has been promoted thanks to studies listed in the previous section investigating the public's reaction to CJEU rulings, sanctions, and EU legislation.

The key distinction between experimental and non-experimental survey scenarios lies in their design. In non-experimental surveys, all participants respond to the same vignette(s) where critical information varies, and their reactions are recorded. In contrast, experimental surveys require participants to be randomly assigned to at least two different scenarios: one that includes the treatment and another serving as a control.

[26] Arthur Dyevre, 'EU Judicial Behaviour Research: A Look Back and a Look Ahead' (2024) 25(3) *European Politics and Society* 468.

[27] Janet Finch, 'The Vignette Technique in Survey Research' (1987) 21 *Sociology* 105; Tobias Nowak, Fabian Amtenbrink, Marc Hertogh, and Mark Wissink, *National Judges as European Union Judges: Knowledge, Experiences and Attitudes of Lower Court Judges in Germany and the Netherlands* (Eleven International 2011); Urszula Jaremba, *National Judges as EU Law Judges: The Polish Civil Law System* (Brill 2013).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 23:04:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009672580.015

For example, consider a study conducted[28] to examine how Polish judges would react to a case of compatibility between EU and Polish law if the Polish Constitutional Court (PCC) had restricted the application of EU law. The 113 judges participating were presented with the following two scenarios:

> **Scenario 1 without the intervention of the PCC:** 'You are uncertain whether or not a national provision conflicts with an EU provision. In this case, the national provision is central to the resolution of the case. However, one of the litigants invokes a CJEU ruling stating that the national legislation is contrary to EU law and not applicable. Consequently . . .'
>
> **Scenario 2 with the intervention of the PCC:** 'You are uncertain whether or not a national provision conflicts with an EU provision. In this case, the national provision is central to the resolution of the case. However, one of the litigants invokes a CJEU ruling stating that the national legislation is contrary to EU law and not applicable. *By contrast, the Constitutional Court has ruled that this EU provision should be applied restrictively because it is affecting fundamental national legal rules or values.* Consequently . . .' (emphasis added)

For each scenario, the judges were presented with three courses of action to choose from:

1. I would secure the national provision from the CJEU's interpretation./I would follow the Constitutional Court's interpretation.
2. I would interpret national law in accordance with EU law.
3. I would follow the CJEU's interpretation and apply EU law instead of the national law.

The researchers then checked whether the judges' responses changed based on the intervention of the PCC. The proportion of judges who said they would secure the national provision from the CJEU (option A) increased from 11.50 per cent for scenario 1 to 58.26 per cent for scenario 2.

However, this strategy has several problems. One of them is so-called 'order effects';[29] the responses might be affected by the sequencing of the vignettes. For instance, the participants might frame scenario 2 in contrast to scenario 1, exaggerating their responses to differentiate the two situations and, hence, amplifying the effect of the PCC's intervention as well.[30] A second issue arising from the lack of random assignment is the difficulty in establishing strong causality. Without randomisation

[28] Juan A. Mayoral, 'Game of Courts: The Effect of Constitutional Judicial Conflicts on Polish Judges' Cooperation with the CJEU' in Clara Rauchegger and Anna Wallerman (eds.), *The Eurosceptic Challenge: National Interpretation and Implementation of EU Law* (Hart 2019) 213.

[29] Fritz Strack, 'Order Effects in Survey Research: Activation and Information Functions of Preceding Questions' in Norbert Schwarz and Seymour Sudman (eds.), *Context Effects in Social and Psychological Research* (Springer 1992) 23.

[30] Arto T. Selkälä and Ulf-Dietrich Reips, 'A New Two-Dimensional Question Format in Web Survey Design: Assimilation and Contrast Effects' (2023) 16(2) *Methodological Innovations* 186.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 23:04:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009672580.015

of judges between scenarios, it remains unclear whether changes in the judges' responses are genuinely due to the PCC intervention or merely a result of the contrast with the preceding scenario.

Scenario survey experiments can fix these situations by ensuring that participants are randomly assigned either to scenario 1 (as a control group) or 2 (as a treatment group), enhancing the causal link between the judges' responses and the intervention of the PCC. Although the setting in which individuals respond is not fully controlled, this method still allows for some manipulation of how participants are assigned to different conditions. These experiments are also highly flexible, as they can be conducted outside of labs or courthouses and distributed in many ways (on paper, via interviews, or online), making it much easier to reach the target respondents. However, this is also its main disadvantage as respondents complete the survey in uncontrolled environments (e.g., at home, at work, alone or with someone, on a computer or mobile, etc.), which introduces variability in how they engage with the experiment and where interruptions, misunderstanding of scenarios, or a lack of interest can influence their responses.

Scenario survey experiments can be applied in several modes. The literature demonstrates how these methods are used to explain the public's attitudes towards EU law and the functioning and interventions of EU institutions. By varying, for instance, the content of court rulings, researchers evaluate how these differences influence perceptions of legitimacy or public support.[31] Another option is short stories where individuals are required to imagine that they are part of a narrative and respond to questions.[32] In these scenarios, researchers can play with the type of information that they present to the participants and ask them to place themselves in a situation, context, or interaction (court hearing, in a undemocratic country, in an exchange with lawyers) or add information about the behaviour of other actors, such as judges, litigants, politicians, and so on.

Another possibility is the use of quasi-experimental settings when expected or unexpected events occur during a regular survey. This approach helps to identify causal effects of important events on survey outcomes. With this technique, researchers do not have control over participants' group assignment as this is instead determined by the event in question. For example, Turnbull-Dugarte and Devine[33] applied quasi-experimentation to study the impact of CJEU rulings on public opinion, taking as a reference the announcement of the salient and highly politicised *Junqueras* ruling[34] in Spain. The content of the ruling was published on 19 December 2019, during the fieldwork period for Wave 9 of the European Social Survey (ESS) conducted between 8 November 2019 and 27 January 2020.

[31] Madsen et al. (n 10) 66.
[32] Kees van den Bos, *Empirical Legal Research: A Primer* (Edward Elgar 2020).
[33] Turnbull-Dugarte and Devine (n 10).
[34] Case C-502/19 *Junqueras* ECLI:EU:C:2019:1115.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 23:04:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009672580.015

The unexpected character of the CJEU ruling created a quasi-experiment with (naturally) exogenous random assignment of exposure to the CJEU's decision.[35] Randomly selected ESS respondents interviewed before the ruling served as the control group, while those interviewed after its announcement constituted the treatment group.

Stiansen et al.[36] also benefited from unexpected events during their survey to study how Polish citizens' views on the *Law and Justice* (PiS) government's judicial reforms are influenced by information concerning the battle between the EU and the Polish government over these measures. While data was being collected in Poland, the European Commission announced an infringement proceeding against Poland for violating EU law with respect to a new law targeting Russian meddling with the upcoming Polish elections. This unexpected development provided an opportunity to design a quasi-experiment, allowing the authors to test the extent to which Polish citizens take cues from EU enforcement actions.

12.3.2 *Laboratory Experiments*

Lab(oratory) experiments comprise a sample of individuals assigned to a hypothetical framing or scenario in a controlled situation and with controlled procedures in order to study a decision, behaviour, or opinion. Such experiments might simulate court trial decisions, judicial dialogues, or situations where different actors interact in the legal domain. In the lab, researchers can randomly allocate the conditions to the participants, as in survey experiments, but in a more carefully controlled environment where researchers limit the impact of external factors to ensure the experiment's integrity. These situations are set up so that only the variables of interest or treatment are allowed to vary, while other potentially confounding factors are kept constant.

Normally, samples of law students are used as they are easier to access. However, lab experiments might also involve non-student subjects like lawyers or judges, especially when studying the legal decision-making of these professionals. These experiments are also referred to as an *artefactual field experiment* as they confront an imaginary lawsuit, and participants must take action or a decision as in real life.[37]

In the worst-case scenario – that is, if access to legal professionals is significantly limited – experiments can be conducted with comparable populations. These may include legal advisors in courts, candidates from judicial training schools, lawyers, or law students. Several studies have compared undergraduates and political and legal elites, and they have found that judges and law students differ systematically in their

[35] Jordi Muñoz, Albert Falcó-Gimeno, and Enrique Hernández, 'Unexpected Event during Survey Design: Promise and Pitfalls for Causal Inference' (2020) 28(2) *Political Analysis* 186.
[36] Øyvind Stiansen et al. (n 10).
[37] Grootelaar and van den Bos (n 19).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 23:04:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009672580.015

ability to apply legal rules.[38] Nevertheless, several techniques on preparation, framing, and training of the participants might reduce these differences by, for instance, offering a brief informational sessions on EU law to the participants to give some practical legal information and guidance on legal reasoning.[39] Another way to improve the readiness of students before the survey is by practising EU law cases, similar to those that students will encounter in the experiment, to ensure they feel confident and familiar with legal reasoning.

Normally, experiments collect participants' self-reports on their responses to various experimental scenarios and measure their reaction times to each one. Additionally, integrating biometric technologies into lab experiments allows researchers to capture physical responses, such as those measured by a pupilometer. This device assesses cognitive effort by tracking pupil dilation, as greater task demands (like complex legal cases) normally lead to increased pupil size.[40]

Lab experiments, due to the random assignment of the selected subjects to treatment and control conditions in a controlled environment, provide an advantage for establishing causal relationships that improve internal validity, compared to field and scenario experiments. Nevertheless, this is often achieved at the cost of external validity or generalisation. It is important to indicate that even if experiments might be abstract, they should not be unrealistic. Although researchers try to approximate common real-world circumstances as much as they can in their labs, they are more interested in testing whether the experiment design gives a specific outcome. In this regard, several strategies might be applied to demonstrate that the effects found in lab experiments happen in the real world, such as combining lab experiments with field experimentation, frequency surveys, or interviews to show that the observed effect is found in other data and its interpretation is solid across methods. Additionally, the external validity of the experiment is also ensured by the accurate design and pre-test in collaboration with legal experts to create experiments that resemble real-life legal situations.

To date, except for Ovádek's in-class experiment,[41] no research publications using lab experiments on EU law exist, which leaves a vast untapped field for exploring legal behaviour, attitudes, and decision-making. Building on the Law & Economics scholarship,[42] experimental labs can be used to understand how legal reasoning,

[38] Dunoff and Pollack (n 2); Susan D. Hyde, 'Experiments in International Relations: Lab, Survey, and Field' (2015) 18(1) *Annual Review of Political Science* 403; Holger Spamann and Lars Klöhn, 'Can Law Students Replace Judges in Experiments of Judicial Decision-Making?' (2024) 1(1) *Journal of Law and Empirical Analysis* 158.

[39] Ovádek (n 25).

[40] Emanuel Querino, Lafaiete dos Santos, and Giuliano Ginani et al., 'Cognitive Effort and Pupil Dilation in Controlled and Automatic Processes' (2015) 6(1) *Translational Neuroscience* 168.

[41] Ovádek (n 25).

[42] Christoph Engel, 'Experiments' in Lee Epstein, Gunnar Grendstad, Urška Šadl, and Keren Weinshall (eds.), *The Oxford Handbook of Comparative Judicial Behaviour* (Oxford University Press 2024) 265.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 23:04:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009672580.015

doctrines, methods of interpretation, or even personal (e.g., EU identity) or political characteristics impact case decisions on EU law. Using the example of methods of interpretation, we could create a lab experiment where participants (ideally judges) were randomly assigned to one or more of the methods of interpretation[43] used by the CJEU in a ruling and ask them whether they would comply or not with it. This lab experiment would give more clarity to the important issue of judicial compliance by also establishing strong causal claims on the impact of the methods of interpretation on judges' behaviour and, hence, its relevance for the legal integration of Europe.

Lab experiments are also suitable for reviewing current research questions on the extent to which legal behaviour or decision-making is influenced by factors such as the identity of the litigants (e.g., Member State, Commission, business, or individuals),[44] the precedents set by earlier CJEU rulings,[45] among others. Translating these discussions into lab experiments could refine longstanding questions or uncover new nuances in our theoretical understanding of the legal construction of Europe.[46]

12.3.3 *Field Experiments*

Empirical research in EU law can benefit from changes, interventions, or differences occurring in the fabric of EU law (legal/policy settings, jurisdictions, etc.).[47] Normally, a field experiment takes place in a court where we can introduce new modes of legal reasoning, institutional reforms, staff, and so on, and see how it affects the way judges deal with their cases before and after this change or intervention is made. We can formulate field experiments where we establish *within-subjects* and *between-subjects* comparisons, each with its pros and cons.[48] For example, a suitable *within-subjects* comparison would involve judges who previously rendered judgments on EU law matters without specialised training in the field, compared to their performance after completing a specialised course in EU law. We can also

[43] Koen Lenaerts and José Antonio Gutiérrez-Fons, 'To Say What the Law of the EU Is: Methods of Interpretation and the European Court of Justice' (2014) 20(2) *Columbia Journal of European Law* 3.

[44] Silje Hermansen, Tommaso Pavone, and Louisa Boulaziz, 'Leveling and Spotlighting: How International Courts Refract Private Litigation to Build Institutional Legitimacy' *British Journal of Political Science*. 2025;55:e158.

[45] Urška Šadl and Ioannis Panagis, 'The Force of EU Case Law: An Empirical Study of Precedential Constraint' (2016), iCourts Working Paper Series No. 68, available at https://ssrn.com/abstract=2787119.

[46] Dyevre (n 26).

[47] Christoph Engel, 'Chapter 3: Experimental Criminal Law: A Survey of Contributions from Law, Economics, and Criminology' in Willem H. van Boom, Pieter Desmet, and Peter Mascini (eds.), *Empirical Legal Research in Action* (Edward Elgar 2018) 23.

[48] Van den Bos (n 32).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 23:04:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009672580.015

conduct a *between-subjects comparison* by examining judges from courts that attended these specialised courses and comparing them to those who did not.

As a main advantage, *within-subject* comparisons offer a straightforward analysis by examining the same individuals across different conditions. However, they can generate *test-retest* effects, where participants' previous responses influence their later answers. In contrast, *between-subject* comparisons avoid this issue by comparing different groups. However, they may reduce comparability due to differences between the groups, especially if random assignment is not possible, as naturally occurring groups may vary in uncontrollable ways.

Due to the difficulty of setting experiments in courtrooms, field experiment design frequently depends on naturally occurring differences, turning them into quasi-experiments where researchers lack control over participant selection. This is important in experiments because it ensures that the treatment and control groups are comparable, which strengthens the validity of causal inferences. While randomisation provides the gold standard for causal inference, quasi-experiments, even without random assignment, often provide sufficient causal evidence by demonstrating consistent effects across similar comparisons between groups that differ naturally in exposure to a treatment.[49]

Some of the reviewed works adopted this type of design to measure the impact of events like Brexit on national courts' preliminary references[50] and the EU Pilot programme on infringement procedures.[51] Using the infringement procedure as a policy field, Cheruvu and Fjelstul did not randomly assign Member States to a treatment group (participants in the EU Pilot programme) or a control group (no participation in EU Pilot). Instead, the variation occurred naturally. Ideally, researchers would use randomised controlled experiments in field settings, where participants are randomly assigned to experimental or control groups. However, such opportunities are rare and difficult in policy and legal contexts, despite their methodological advantages.

Finally, we should stress the potential that field experiments have for studying the implementation and effect of EU legislation on the behaviour of Member States or other actors (e.g., corporations, workers, citizens, etc.). By designing field experiments or quasi-experiments, we could assess the extent to which certain environmental regulations, consumer protection laws, or competition rules have an impact or achieve their objectives depending on heterogeneous national or sub-national preferences, capacities, and conditions.[52] For instance, a quasi-experimental setting

[49] William R. Shadish, M.H. Clark and Peter M. Steiner, 'Can Nonrandomized Experiments Yield Accurate Answers? A Randomized Experiment Comparing Random and Nonrandom Assignments' (2008) 103(484) *Journal of the American Statistical Association* 1334.

[50] Dyevre et al. (n 3).

[51] Cheruvu and Fjelstul (n 10).

[52] Asya Zhelyazkova, Eva Thomann, Eva Ruffing, and Sebastiaan Princen, 'Differentiated Policy Implementation in the European Union' (2023) 47(3) *West European Politics* 439.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 23:04:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009672580.015

could address corporate compliance with EU law depending on the enforcement mechanisms deployed by national authorities to enforce this regulation. In this regard, a treatment group might consist of firms operating in Member States with active enforcement of the selected legislation – that is, where monitoring and penalties are visibly enforced – while a control group could be made of firms in Member States with weaker enforcement mechanisms or delayed implementation timelines.

An observational design of the same situation would analyse compliance levels across all firms in different Member States, identifying statistical associations or correlations between variables without controlling for confounders. This lack of control increases the risk of bias from unobserved variables.[53] This risk is mitigated in the quasi-experiment design, which compares two groups of firms operating with strong and weak enforcement. By employing techniques such as difference-in-differences, this research strategy allows for causal inference while controlling for time-invariant unobserved confounders.[54]

12.4 CONCLUSION

Legal experimentation offers a powerful empirical methodology for understanding and improving the design and effectiveness of EU law. It enables researchers to test new initiatives or changes that might affect legal decision-making or behaviour, measure their impact, and identify the most effective solutions for broader application.

This method is particularly useful in the European context, where EU law often operates in diverse and complex real-world settings due to the decentralised, and sometimes discretionary, national enforcement of EU law. This method might provide useful insights into what works under certain settings, helping judiciaries, national governments and, most importantly, EU institutions to identify more efficient legal interventions. Such an evidence-based approach can bridge the gap between legal design and actual legal decision-making and implementation of EU law, addressing problems of effectiveness, uniform judicial application, and compliance.

Building on the earlier example of judicial training, a more precise field experiment with straightforward policy implications could be designed to investigate the impact of judicial training on the uniform application of the Digital Services Act (DSA) across Member States in collaboration with the European Judicial Training Network (EJTN). In this experiment, judges in the treatment group would attend a

[53] Marie-Claire E. Aussems, Anne Boomsma, and Tom A.B. Snijders, 'The Use of Quasi-experiments in the Social Sciences: A Content Analysis' (2011) 45 *Quality and Quantity* 21.

[54] Alan S. Gerber, Donald P. Green, and Edward H. Kaplan, 'The Illusion of Learning from Observational Research' in Ian Shapiro, Rogers M. Smith, and Tarek E. Masoud (eds.), *Problems and Methods in the Study of Politics* (Cambridge University Press 2004) 251.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 23:04:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009672580.015

training module emphasising strict enforcement of the DSA, focusing on transparency and accountability requirements for online services. In the control group, judges would receive no training before ruling on EU law cases related to the enforcement of the DSA. This would help researchers and policy-makers to assess whether specialised judicial training influences how national judges apply the DSA and the extent to which it contributes to its uniform implementation across Member States.

Despite these advantages, experiments in the legal realm face significant challenges, especially in the case of lab experiments. Random selection and random assignment are particularly difficult to achieve, not only due to ethical issues and the safeguarding of legal principles of equality, fairness, and justice, but also due to the contextual and heterogeneous nature of national legal systems where EU law is applied by judges and national authorities. Differences in institutional structures, legal cultures, and procedural norms across jurisdictions can complicate the design and generalisability of experiments. These challenges necessitate creative methodological solutions and careful consideration of contextual factors to ensure meaningful and credible results.

In sum, while challenging, conducting experiments in the field of EU law unquestionably has the potential to transform and advance EU law research. By leveraging experimentation, researchers can contribute to a deeper understanding of EU legal decision-making, rulings, and legislation and their impact, providing valuable tools for crafting judicial proceedings and laws that are not only effective but also adaptable to the complexities of the European Union's legal and policy diversity.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 23:04:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009672580.015

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 23:04:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009672580.015