

Choosing Not to Experiment: Regulatory Sandbox Adoption and Policy Instrument Choice in Europe

FIRSA Working Paper No. 2026/02

Alfredo Hernández Sánchez, *PhD*

Vilnius University

July 2026

*Understanding FinTech Regulatory Sandbox
Development in Europe (FIRSA)*

Website: <https://firsa.eu>



Funded by
the European Union



FIRSA Working Papers

This working paper series publishes early research results and policy papers connected to the EU Horizon Marie Skłodowska-Curie Actions ERA Fellowship *FIRSA*. For any comments, please contact:

Alfredo Hernández Sánchez
Institute of International Relations and Political Science
Vilnius University
Vilnius, Lithuania
alfredo.sanchez@tspmi.vu.lt

IMPORTANT:

Author's Original Manuscript (AOM). This is a pre-print version of a manuscript intended for submission to an academic journal. Any remaining errors are the author's responsibility.

Disclaimer

FIRSA is funded by the European Union under grant agreement No. 101180601 titled: *Understanding FinTech Regulatory Sandbox Development in Europe*. Views expressed are those of the author(s) and do not necessarily reflect those of Vilnius University or the European Union.

Published under license [CC BY-NC 4.0](https://creativecommons.org/licenses/by-nc/4.0/) · Copyright © 2026 Alfredo Hernández Sánchez

Table of contents

Abstract	5
Introduction	7
Policy Experimentation and Instrument Choice	10
Varieties of Sandboxing in Europe	13
Explaining Adoption and Non-Adoption	16
Data and Methods	18
Results	22
No single condition explains adoption	22
The adoption solution	23
Contradictory configurations	26
Non-adoption is better structured than adoption	27
Interpreting the Contrasts	29
Lithuania and Latvia against Estonia	29
Austria and Spain against Germany	30
Malta against Finland and Ireland	31
Italy against Belgium, France, and Portugal	31
Validating the Configurational Results	32
Discussion	34
Conclusion	35



Appendix	37
Appendix A: Case-Level Calibration	37
Appendix B: Truth Table and Necessity Diagnostics	38
Appendix C: Regression Results	39
References	40



Abstract

Regulatory sandboxes are widely read as signals of innovation-friendly financial regulation and as evidence of regulatory modernity. Yet only eleven of thirty EU and EEA jurisdictions operate one. This paper treats the European FinTech sandbox map as a problem of policy instrument choice rather than as a mere measure of innovation friendliness. Because initiatives bearing the sandbox label differ widely, the analysis separates supervised live-testing sandboxes from advisory programmes, statutory exemptions, and innovation hubs, and applies a crisp-set Qualitative Comparative Analysis to thirty EU/EEA countries, complemented by case evidence on the decisive contrasts. No condition is close to necessary or sufficient for adoption. Latvia and Lithuania form the only repeated adopter profile, and six of the eleven adopters share their measured profile with non-adopters. Non-adoption follows a clearer institutional logic, in which established alternative facilitation tools, or the absence of supporting administrative conditions, accompany the decision not to sandbox.

Keywords: comparative policy analysis; FinTech; policy instruments; regulatory sandboxes



Practitioner points

- Counts of regulatory sandboxes should not be treated as rankings of regulatory innovation or FinTech friendliness, because similar labels cover very different institutional designs.
- Innovation hubs, supervisory guidance, and ordinary supervision can be deliberate alternatives to live-testing sandboxes rather than weaker substitutes for them.
- Cross-country comparisons should distinguish supervised live testing from advisory cohort programmes, statutory exemptions, and general innovation-facilitation tools before drawing policy lessons.



Introduction

Between 2018 and 2024, eleven EU and EEA jurisdictions established initiatives labelled as FinTech regulatory sandboxes in which firms test financial products with real customers under supervisory monitoring, while nineteen jurisdictions facing the same technologies and the same EU regulatory framework did not ([European Supervisory Authorities 2023](#)). The non-adopters include several of Europe's most digitally capable states and largest FinTech markets, among them Germany, the Netherlands, Sweden, Finland, and Estonia, whereas the adopters include Greece, Hungary, and Slovakia, none of which stands out on conventional measures of FinTech activity or digital government. If sandboxes were simply what innovation-friendly regulators do under FinTech demand-side pressure, the European pattern should look very different. The puzzle addressed here is why supervised regulatory experimentation was institutionalized in some EU and EFTA countries but not in others facing comparable technologies, a common European regulatory framework, and often stronger market pressure.

In policy discourse and in the growing evaluation literature, sandbox adoption is routinely treated as a marker of regulatory modernity. International organizations have promoted the instrument as good regulatory practice ([Jeník and Duff 2020](#)), cross-national studies count sandboxes as evidence of innovation-friendly regulation ([Markellos et al. 2024](#)), and evaluations of the British scheme document real benefits for participating firms ([Cornelli et al. 2024](#)). At the same time, schemes labelled 'sandbox' range from supervised live testing with regulatory waivers to advisory cohort programmes and statutory licensing exemptions, and several non-adopters operate facilitation tools that perform similar functions without the label ([Zetzsche et al. 2017](#); [Buckley et al. 2019](#); [European Supervisory Authorities 2023](#)). Any explanation of sandbox diffusion must therefore first decide what is being explained. This paper defines the outcome narrowly, as supervised live testing, and treats advisory



and statutory-exemption variants as distinct designs within a typology of innovation-facilitation instruments.

Consequently, the paper aims to contribute to the literature on policy instrument choice. Policy instruments embed assumptions about the relationship between regulator and regulated, and their selection is thus a substantive political decision that organizes who bears risk and who obtains access ([Lascoumes and Le Galès 2007](#); [Howlett 2009](#)). Because most jurisdictions operate several facilitation tools at once, the relevant unit of analysis is the policy mix, and the question is why some states added a live-testing instrument to their mix while others deliberately did not ([Howlett and Rayner 2007](#); [Borrás and Edquist 2013](#)). Furthermore, sandboxes are enmeshed within a system of consultancies, international bodies, and regulatory networks promoting them, which makes the gap between the label's diffusion and the design's diffusion analytically informative ([Béland and Howlett 2016](#); [Voß and Simons 2014](#)).

The empirical strategy combines a crisp-set Qualitative Comparative Analysis (QCA) of thirty EU/EEA countries with case-level evidence on the contrasts the analysis identifies as decisive. The QCA draws on the coding of sandbox and innovation-hub arrangements verified against supervisory sources, European Banking Authority payment-institution data, UN e-government indicators, and population data ([Ragin 2008](#); [Schneider and Wagemann 2012](#)). The design is configurational and descriptive in ambition. It uses a conventional consistency threshold, reports contradictory configurations rather than suppressing them, and reads the solutions as maps of the case landscape whose interpretation is then carried by case evidence, not as causal recipes. The two sides of the outcome behave differently in ways that turn out to be substantively relevant. On the adoption side, no common profile unites the eleven adopters; only the Baltic pair of Latvia and Lithuania forms a repeated configuration, and six of the eleven adopters share their measured profile with non-adopters. On the non-adoption side, the solution is markedly stronger, since countries with established alternative facilitation channels, or without the administrative conditions that make



experimentation attractive, declined to sandbox with few exceptions.¹

The paper advances two arguments. The first is that structural conditions underdetermine sandbox adoption. The same national profile is compatible with sandboxing and with its alternatives, which matches case-study findings that policy entrepreneurship and perceived legal constraints decided the Baltic outcomes (Raudla et al. 2024, 2025), and which the policy-entrepreneurship literature would lead one to expect wherever structural conditions are permissive rather than compelling (Mintrom and Norman 2009). The second is that non-adoption has an institutional logic of its own, an alternative-instrument equilibrium in which hubs, guidance, and ordinary supervision do the work a sandbox would otherwise do, rather than being a residual category of regulatory conservatism (Black and Baldwin 2010; Tõnurist, Kattel, and Lember 2017). Separating supervised live testing from advisory, statutory, and legal-exemption variants is a necessary step towards both arguments, since the label otherwise conceals the design variation on which they rest. The setting sharpens the stakes of the choice. Within the single market, an authorization granted in one member state can be passported across the others, so the quality and accessibility of national authorization processes become competitive resources for jurisdictions courting mobile financial firms (Zetzsche et al. 2017; Buckley et al. 2019). A sandbox does not by itself confer passportable authorization, but it advertises and structures the path towards one, which is why instrument choice in this field is simultaneously a supervisory and a positional decision. Governing FinTech uncertainty, as the term is used here, means choosing among instruments that reduce regulatory uncertainty for supervisors and firms alike, whether through bounded live experiments, structured advice, or the ordinary exercise of supervisory discretion.

¹A regression cross-check is consistent with this configurational reading and gives no support to a linear FinTech-pressure account, as adopters are drawn disproportionately from below-median FinTech density.



Policy Experimentation and Instrument Choice

A long-standing insight of the policy instruments literature is that governments choose tools, not merely goals, and that these choices are patterned by institutions, capacities, and ideas rather than dictated by the problem at hand ([Lascoumes and Le Galès 2007](#); [Howlett 2009](#)). Instrument choices are nested, in the sense that a decision to add a sandbox sits inside broader commitments about the governance mode of financial supervision, and they are cumulative, in the sense that new tools land in an existing mix whose coherence governments try to preserve ([Howlett and Rayner 2007](#)). For innovation policy specifically, the menu of instruments is wide and the criteria for selection are rarely reducible to mere effectiveness given that instruments also serve as signals to firms and to peer regulators ([Borrás and Edquist 2013](#)). Seen from this angle, the question of why a supervisor adopts a sandbox is a subset of the general question of why organizations add a new instrument to a functioning mix, and the answer need not be the same as the answer to why the instrument exists in the first place.

Sandboxes belong to the family of experimentalist instruments. The experimentation literature distinguishes arrangements that generate systematic learning through bounded trials from arrangements that merely label ordinary discretion as experimental, and it emphasizes that genuine policy experiments require someone to own the risk of failure ([Huitema et al. 2018](#); [Ansell and Bartenberger 2016](#)). In EU governance, experimentalism has typically taken the form of framework goals with local discretion and pooled review, a mode in which learning is organized across levels rather than within a single supervisory scheme ([Sabel and Zeitlin 2008](#)). A live-testing sandbox is a strong commitment within this family. The supervisor authorizes real transactions with real customers under relaxed or closely monitored conditions, takes institutional ownership of controlled risk, and exposes itself to criticism on equal-treatment grounds from firms outside the scheme. Advisory programmes



and contact points sit at the weak end of the same family, since they promise dialogue and guidance without any bounded trial. The political economy of adopting the strong variant therefore differs from that of adopting the mere label, thus a theory of sandbox adoption is in large part a theory of when supervisors accept ownership of experimental risk.

The diffusion literature supplies the baseline expectation against which any such theory must be tested. Instruments spread through learning, competition, imitation, and coercion, and adoption counts alone cannot distinguish among these mechanisms (Shipan and Volden 2008; Gilardi 2010; Maggetti and Gilardi 2016). The sandbox has travelled as a template, popularized by the United Kingdom's Financial Conduct Authority and carried onward by consultancies, international bodies, and regulatory networks with a stake in its spread (Béland and Howlett 2016; Voß and Simons 2014; Jeník and Duff 2020). Networks of this kind move labels more readily than the institutional commitments beneath them, and what an importing regulator makes of the template depends on its own capacities and interpretive frames, so the same model can yield advisory schemes in one jurisdiction and live testing in another (Claire A. Dunlop 2009; Claire A. Dunlop and Radaelli 2013).

Learning and capacity arguments cut in two directions, and this ambivalence matters for the analysis below. Supervisors may adopt sandboxes as learning instruments that give controlled access to information about emerging business models before risks scale, in line with the epistemic reading of financial regulation in which authorities interpret uncertainty through professional communities and institutional mandates (Claire A. Dunlop 2017; James and Quaglia 2023; Kranke and Yarrow 2019). Yet high-capacity authorities are also those best placed to absorb novel business models into ordinary supervision and to conclude that a formal testing scheme adds little, a strategy the responsive-regulation literature treats as legitimate rather than as inaction (Black and Baldwin 2010), while innovation hubs are institutionalized facilitation tools in their own right rather than diminished sandboxes (Tönurist,



[Kattel, and Lember 2017](#)). Policy capacity is in any case multidimensional, since the analytical capacity to run a testing scheme is not the same as the political capacity to defend selective regulatory relief ([Wu, Ramesh, and Howlett 2015](#)). The implication is that capacity should appear on both sides of the outcome, that non-adoption can be an actively maintained equilibrium, and that the decisive differences between similar jurisdictions are likely to run through actors and legal interpretation, particularly where national experimentation proceeds in the shadow of EU law ([Héritier and Lehmkuhl 2008](#)).

The FinTech-specific literature fits within this frame. FinTech confronted European supervisors with a common problem, since new entrants and business models promised competition and efficiency while complicating prudential oversight, consumer protection, and market-integrity enforcement ([Arner, Barberis, and Buckley 2015](#); [Philippon 2017](#)), and national responses varied with state capacities, financial-sector structures, and the platform dynamics of the sector ([Langley and Leyshon 2021, 2023](#)). Evaluation research shows the instrument can matter, as participation in the UK sandbox improved firms' access to funding, plausibly by reducing information asymmetries ([Cornelli et al. 2024](#)), and cross-national work documents widespread but uneven diffusion associated with regulatory capacity and market development ([Markellos et al. 2024](#)). What that literature does not explain is why jurisdictions with similar capacity and exposure chose different instruments. The Baltic states illustrate the point sharply, since three small, digitally ambitious neighbours produced two sandboxes and one deliberate non-adoption, an outcome that case research attributes to policy entrepreneurs, diffusion, and perceived legal constraints rather than to structural differences ([Raudla et al. 2024, 2025](#)).



Varieties of Sandboxing in Europe

A FinTech regulatory sandbox, as understood here, is a formally established scheme in which firms test innovative financial products, services, or business models with real customers, for a limited period, under the monitoring of a financial supervisory authority, with or without waivers from specific regulatory requirements. The definition follows the core elements identified in the legal and regulatory literature on innovation facilitators, which consistently distinguishes testing schemes from advice-giving arrangements ([Zetzsche et al. 2017](#); [Buckley et al. 2019](#)), and it matches the definition used by the European Supervisory Authorities, for whom sandboxes provide a controlled environment for live testing while innovation hubs provide a point of contact and non-binding guidance ([European Supervisory Authorities 2023](#)). Two elements are demanding, live testing with real market participants and structured supervisory involvement in the test, and they deliberately exclude arrangements that share the sandbox vocabulary but not this institutional core. Innovation hubs, the most common facilitation tool in Europe and one that every EU/EEA jurisdiction now operates in some form, are the most important neighbouring instrument, and the regulatory politics of the two differ, because live testing requires the supervisor to take ownership of controlled risk while hubs and advice do not ([Tönurist, Kattel, and Lember 2017](#); [European Supervisory Authorities 2023](#)).

Table 1 organizes the European schemes into five types. The first two, supervisory live-testing sandboxes and statutory testing spaces, meet the definition and jointly constitute the outcome studied below. They differ in legal origin. Most European sandboxes were created by supervisors within existing mandates, while Spain and Italy anchored live testing in dedicated legislation with inter-authority coordination, Spain through Law 7/2020 on the digital transformation of the financial system and Italy through the 2019 growth decree and its 2021 implementing regulation, which established a sandbox operated jointly by the banking, securities, and insurance



authorities under a ministerial FinTech committee.² The third type, the advisory cohort sandbox, uses the sandbox label for structured, thematic advice without live testing, a design of which Ireland's Innovation Sandbox Programme, announced in 2023 and first operated in 2024, is the European case.³ The fourth, the statutory exemption, creates a legal carve-out without any supervised testing process, and Switzerland's 2017 banking-law sandbox, which permits firms to accept public deposits up to one million francs without a banking licence, is of this kind.⁴ The fifth type covers innovation hubs, contact points, and tailor-made supervisory guidance, which function as the main facilitation tool in the nineteen non-adopters ([European Supervisory Authorities 2023](#)).

Three boundary decisions follow from the typology and are applied consistently throughout. Estonia is coded as a non-adopter of supervised live testing. The *Finantsinspektsioon* operates an Innovation Hub, opened in 2021, whose official material explicitly distinguishes the hub from a regulatory sandbox, and the Baltic case literature treats Estonia as the regional non-adopter ([Raudla et al. 2024, 2025](#)). The coding requires care because the 2023 update of the European Supervisory Authorities' innovation-facilitator inventory lists an Estonian sandbox established in 2023 ([European Supervisory Authorities 2023](#)). The scheme in question is the *Finantsinspektsioon*'s test environment, which the supervisor describes as an arrangement supporting the Innovation Hub through agreed long-term monitoring of an innovative solution so as to inform the qualification of the service and the development of super-

²Spain: Ley 7/2020 which created the *espacio controlado de pruebas* administered by the Treasury with the sectoral supervisors. Italy: Decreto-legge 34/2019 (art. 36) and Decreto del Ministero dell'Economia e delle Finanze n. 100/2021, establishing the regulatory sandbox coordinated through the Comitato FinTech with Banca d'Italia, CONSOB, and IVASS. Both schemes are listed as operational statutory sandboxes in [European Supervisory Authorities \(2023\)](#).

³The Central Bank of Ireland describes the Innovation Sandbox Programme as providing regulatory advice and support to selected cohorts working on a common theme before products are placed on the market, without testing involving real customers and without regulatory relief.

⁴Amendment to the Swiss Banking Ordinance (August 2017) which described by FINMA as a licence-exempt innovation space; there is no supervisory testing plan or monitoring component.

Table 1: A typology of FinTech innovation-facilitation instruments in Europe

Type	Defining feature	Cases (year established)
Supervisory live-testing sandbox	Live testing with real customers under supervisory monitoring, created within the supervisor's mandate	Denmark (2018), Hungary (2018), Lithuania (2018), Norway (2019), Austria (2020), Malta (2020), Greece (2021), Latvia (2021), Slovakia (2022)
Statutory testing space	Live testing anchored in dedicated legislation with inter-authority coordination	Spain (2020), Italy (2021)
Advisory cohort sandbox	Sandbox label applied to structured, cohort-based advice without live testing or waivers	Ireland (2024)
Statutory exemption	Legal carve-out from licensing requirements without a supervised testing process	Switzerland (2017)
Alternative facilitation only	Innovation hub, contact point, or tailor-made supervisory guidance without a sandbox scheme	Remaining 18 EU/EEA jurisdictions, including Estonia, Finland, Germany, the Netherlands, and Sweden

visory positions.⁵ Because the arrangement lacks a bounded, supervisor-approved live test, it falls outside the outcome definition, and the sensitivity of every result to recoding Estonia as an adopter is reported in the validation section. Ireland is coded as a non-adopter of live testing despite the sandbox label, because the Central Bank of Ireland's programme is advisory and cohort-based, and it enters the analysis as a design variant rather than as an outcome case. Switzerland is retained in the dataset as a statutory-exemption case but excluded from the main QCA sample, because the FinTech-pressure measure derives from the register of institutions authorized under the EU's revised Payment Services Directive (PSD2), a common authorization regime that does not extend to Switzerland, so the measure is not comparable for a non-EEA jurisdiction.

Figure 1 shows the timing of the thirteen named schemes. Live-testing adoption came in two loose waves, an early cluster in 2018 and 2019 comprising Denmark, Hungary,

⁵The test environment involves no approved testing plan, no bounded testing period, and no regulatory relief; firms operate under the ordinary legal framework while the supervisor observes.

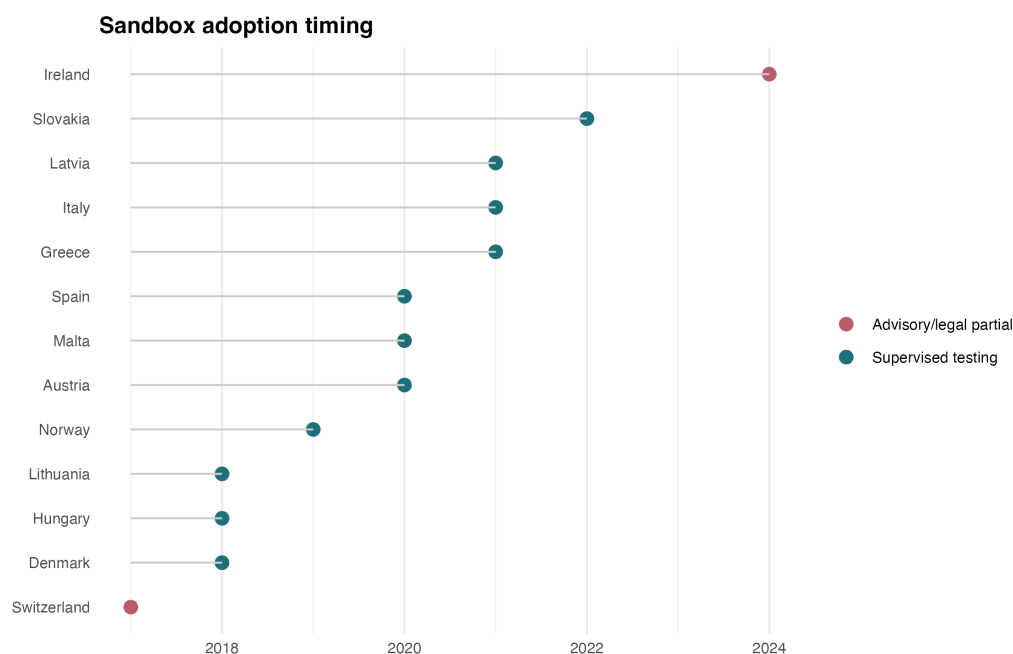


Figure 1: Sandbox adoption timing by design type

Lithuania, and Norway, and a broader second cluster between 2020 and 2022 that includes the two statutory testing spaces alongside Austria, Malta, Greece, Latvia, and Slovakia. Switzerland's statutory exemption predates both waves, and the late arrival of Ireland's advisory programme, well after the peak of European sandbox creation, already suggests that the label travels more easily than the live-testing design.

Explaining Adoption and Non-Adoption

Building on the literatures reviewed above, four expectations guide the analysis. Each is stated with the conditions meant to capture it, and Table 2 records what the proxies can and cannot measure. One term of art recurs throughout and is defined here. Dependent market economies (DMEs) are the post-socialist political economies of



Central and Eastern Europe whose growth model rests on externally owned, export-oriented integration into European production and finance chains (Nölke and Vliegenthart 2009); the broad DME coding used below includes the Baltic and Southeast European cases, whose trajectories differ in ways the analysis returns to (Bohle and Greskovits 2012).

E1 (jurisdictional positioning). Small jurisdictions competing for mobile financial activity within the single market have the strongest incentives to adopt visible regulatory instruments that signal accessible, credible supervision, the more so where FinTech firms are already present and the state has administrative capacity to run a testing scheme. The expected adopter profile combines small-state status with FinTech pressure and digital capacity. Because passporting raises the strategic value of national authorization and supervisory reputation, this is competition through credible supervision rather than through deregulation (Zetzsche et al. 2017).

E2 (supervisory learning). High-capacity supervisors may adopt sandboxes as learning instruments that give controlled access to information about emerging business models before risks scale (Claire A. Dunlop and Radaelli 2013; James and Quaglia 2023). This pathway does not require intense market pressure, and it should appear as digital capacity combined with prior innovation-facilitation engagement, with FinTech pressure irrelevant or absent.

E3 (institutional translation). In post-socialist dependent market economies, FinTech policy can serve as a modernization signal and a catch-up device. Dependent integration alone should not be expected to produce adoption, since post-socialist trajectories differ considerably (Bohle and Greskovits 2012), and the same structural position can support reliance on EU templates, partial or non-supervisory testing arrangements, or a judgement that ordinary regulation suffices. The expectation is therefore conditional, in that the DME condition should matter only in combination with enabling conditions such as prior hubs or digital capacity.



E4 (alternative-instrument equilibrium, non-adoption). Non-adoption is expected to display regularities of its own rather than being a residual left over once adoption is explained. Where established hubs, contact points, or ordinary supervisory dialogue already reduce uncertainty, or where demand and administrative capacity are limited, a sandbox adds little and may raise equal-treatment concerns, an expectation grounded in the reading of hubs and responsive supervision as full-fledged regulatory strategies (Black and Baldwin 2010; Tönurist, Kattel, and Lember 2017). High-capacity jurisdictions in particular can absorb facilitation into ordinary supervision, so capacity should appear on both sides of the outcome.

A fifth and simpler hypothesis is retained as a baseline. The default account in the diffusion literature is functional, holding that instruments spread to where the problem pressure is, and the cross-national sandbox literature indeed associates adoption with market development and regulatory capacity (Shipan and Volden 2008; Markellos et al. 2024). Testing this baseline explicitly serves two purposes. It guards against attributing to configurations a pattern that a single pressure variable could explain more economically, and it disciplines the interpretation of the configurational results, since any conjunctural reading is more credible once the additive alternative has been examined on the same data rather than dismissed by assumption (Maggetti and Gilardi 2016).

Data and Methods

The analysis covers the thirty EU/EEA jurisdictions, the EU-27 plus Iceland, Liechtenstein, and Norway, for which the PSD2-based FinTech measure is comparable, while Switzerland is discussed qualitatively. The outcome, SANDBOX, is supervised live-testing sandbox adoption by the end of 2024 as defined in the typology, which yields eleven adopters among the thirty cases. Outcome coding was reviewed against official supervisor sources, the ESAs innovation-facilitator inventory, and the

Table 2: Mechanisms, expected configurations, and proxy limitations

Mechanism	Expected configuration	Condition(s)	Main proxy limitation
E1: Jurisdictional positioning	Small state with FinTech presence and administrative capacity	SMALL, FINTECH, DIGICAP	PSD2 register covers payment, e-money, and account-information entities only; the six-million population cut-off is a pragmatic dichotomization, not a theoretically derived boundary
E2: Supervisory learning	Digital capacity and prior facilitation engagement, without requiring FinTech pressure	DIGICAP, HUB	EGDI measures e-government readiness, not financial supervisory capacity; hub timing captures formal, not informal, engagement
E3: Institutional translation	Post-socialist dependent integration combined with enabling conditions	DME (with HUB or DIGICAP)	Broad DME coding groups Baltic and Southeast European trajectories that differ in financial positioning
E4: Alternative-instrument equilibrium	Established hubs or high capacity without a sandbox; or absence of enabling conditions	HUB, DIGICAP (for non-adoption)	Hub coding does not measure the intensity or quality of alternative facilitation

secondary literature ([European Supervisory Authorities 2023](#); [Raudla et al. 2024, 2025](#)), and the two recodes relative to the original project file are Estonia and, for the live-testing outcome, Ireland.

QCA is used here because it fits the structure of the problem. It treats each case as a configuration of conditions and asks which combinations of conditions are regularly associated with an outcome, rather than estimating the average effect of each variable across cases ([Ragin 2008](#); [Schneider and Wagemann 2012](#)). With a binary outcome and only eleven adopters among thirty cases, the sample is too small for stable multivariable estimation. The outcome is a discrete institutional choice, and the theoretical expectations sketched above are explicitly conjunctural, in that small-state incentives are expected to matter together with capacity and market presence rather than separately. QCA also permits causal asymmetry, meaning that the explanation of non-adoption need not be the mirror image of the explanation of adoption, which

matters for a study whose central claim concerns the distinct logic of non-adoption. Configurational methods of this kind are an established tool in comparative policy analysis for exactly such medium-N instrument-adoption questions ([Schneider and Wagemann 2012](#); [Dusa 2019](#)).

Five conditions are used, each calibrated as a crisp set, where every country is coded as either in or out of each set; the reasons for preferring this binary calibration over graded fuzzy membership are given below. FINTECH captures FinTech pressure through the number of payment, electronic-money, and account-information institutions authorized under the EU’s common payments framework, the revised Payment Services Directive (PSD2), per million residents, from the European Banking Authority register extraction ([European Banking Authority 2026](#)), dichotomized at the sample median. Because the directive creates a single authorization regime across the EU/EEA, entity counts from its register are comparable across the sample in a way that national FinTech statistics are not. DIGICAP captures digital state capacity through the 2024 UN E-Government Development Index (EGDI) ([United Nations Department of Economic and Social Affairs 2024](#)), also cut at the sample median, a relative threshold in an already high-capacity region. SMALL codes populations at or below six million, using World Development Indicators population data ([Bank 2024](#)); the data source supplies the measure, not the threshold, and the six-million line is a pragmatic cut-off rather than a theoretically derived boundary, chosen because it keeps together the jurisdictions commonly discussed as small states in the European debate, from the microstates through the Baltic states to Denmark, Slovakia, Finland, Norway, and Ireland. HUB (labelled EARLY_HUB in Figure 2) codes whether an innovation hub existed by 2018, before most sandbox adoptions, capturing prior facilitation infrastructure. DME codes broad post-socialist dependent market integration as defined above, with the Baltic cases retained in the broad coding but discussed separately ([Nölke and Vliegenthart 2009](#)). Calibration rules are reported in Table 3 and the full case-level memberships in Appendix A.

Table 3: Calibration of the QCA conditions

Condition	Threshold	Membership rule
FINTECH	10.5	1 if PSD2 entities per million are at or above the EU/EEA sample median
DIGICAP	0.879	1 if EGDI 2024 is at or above the EU/EEA sample median
SMALL	6	1 if population is at or below 6 million
HUB	2018	1 if an innovation hub existed by 2018
DME	1	1 if coded as broad post-socialist dependent market economy

Two methodological choices require justification. The first concerns calibration. Crisp sets are used in the main analysis for three reasons. The outcome is a discrete institutional choice, two of the five conditions (early hub presence and DME status) are dichotomous by construction, and the remaining proxies are not measured precisely enough to defend fine-grained fuzzy membership scores. Crisp calibration also keeps every coding decision auditable at the case level, this is critical given that interpretive weight falls on individual cases. A fuzzy-set alternative, in which the quasi-continuous conditions would receive graded membership scores between 0 and 1 ([Schneider and Wagemann 2012](#)), was considered and set aside, because graded scores would add apparent precision to blunt proxies without changing which side of each calibration threshold the cases fall on, and a fuzzy analysis of a binary outcome with two dichotomous conditions would leave the set-theoretic structure of the evidence unchanged.

The second choice concerns minimization, the step in which QCA reduces the observed combinations of conditions to a compact solution. Its starting point is the truth table, an inventory of every combination of conditions observed in the data together with the outcomes of the countries sharing it. Truth tables for adoption and non-adoption are minimized at a consistency threshold of 0.75, meaning that a combination of conditions counts towards the solution only if at least three-quarters of the countries sharing it display the outcome, and no logical remainders are included, so the solutions summarize only configurations actually observed among the



thirty cases (Schneider and Wagemann 2012; Dusa 2019). Contradictory configurations, identical measured profiles that contain both adopters and non-adopters, are reported and interpreted rather than resolved by threshold manipulation. They are reported because they carry real information. When countries that measure the same on every condition nonetheless chose differently, the measured conditions cannot be the whole explanation, and the analysis has to turn to case evidence for what separates them.

The proxies have known limits, summarized in Table 2, and stating them plainly clarifies what the analysis can and cannot show. The PSD2 register counts firms in the regulated payments sector, so it misses venture funding, crypto activity, and start-up density, and a country can host a lively FinTech scene that the measure understates. The EGDI measures how digitalized government services are, not how sophisticated the financial supervisor is, so it is a proxy for the state's general digital capability rather than a direct measure of supervisory capacity (Wu, Ramesh, and Howlett 2015). Hub timing captures whether formal facilitation infrastructure existed, not how intensively it was used. The DME family is historically grounded but broad. These limits are why the solutions are read as descriptive maps to be interpreted with case evidence, and why the validation section examines how much the headline results depend on individual coding decisions.

Results

No single condition explains adoption

In set-theoretic terms, a condition is necessary if virtually all adopters display it and sufficient if virtually all countries displaying it adopt. Neither holds for any condition here. The highest necessity consistencies, at 0.64, belong to digital capacity, early hub presence, the *absence* of above-median FinTech density, and the absence



of DME status, and all fall far below conventional benchmarks (full diagnostics in Appendix B). The FinTech-pressure baseline fails most clearly, since only four of the eleven adopters (Latvia, Lithuania, Malta, and Denmark) have above-median PSD2 density, while high-density jurisdictions such as Luxembourg, Cyprus, Liechtenstein, the Netherlands, and Sweden did not adopt. The European pattern therefore lends no support to the view that sandboxes emerge wherever FinTech market pressure is strongest, and whatever explains adoption must be sought in combinations of conditions rather than in any single one.

Figure 2 displays the raw material of the analysis and can be read without any set-theoretic apparatus. Each row is a country, each of the first five columns is a condition, the final column is the outcome, and a dark cell indicates membership, so Latvia's row, dark on all six columns, reads as a small, digitally capable, FinTech-dense post-socialist state with an early hub that adopted a sandbox. Two features of the figure anticipate the formal results. Reading down the SANDBOX column, the dark cells are scattered across very different row patterns, which is the visual form of the finding that adopters share no common profile, since the adopter set contains the full-profile Baltic pair, small high-capacity Nordic states, two large statutory cases, and three adopters (Greece, Hungary, Slovakia) with mostly light rows. Comparing rows with identical condition patterns but different outcome cells reveals the contradictions analysed below, of which Germany against Austria is the clearest example. The absence of a common adopter profile, visible before any minimization, is itself the first substantive finding.

The adoption solution

The next two subsections present the two truth-table solutions, first for adoption and then for non-adoption, and they are best read together, because the paper's central empirical result is the asymmetry between them. For readers new to QCA, a solution



Figure 2: Condition memberships and the live-testing outcome

is a list of condition combinations, joined by logical OR, that regularly go together with the outcome; its consistency is the share of countries covered by the solution that actually display the outcome, and its coverage is the share of outcome cases the solution accounts for. Minimizing the adoption truth table at the 0.75 threshold yields four perfectly consistent terms:

$$\begin{aligned}
 \text{SANDBOX} \Leftarrow & \text{FINTECH} \cdot \text{DIGICAP} \cdot \text{SMALL} \cdot \text{HUB} \cdot \text{DME} && (\text{Latvia, Lithuania}) \\
 & + \text{FinTech} \cdot \text{DIGICAP} \cdot \text{SMALL} \cdot \text{hub} \cdot \text{dme} && (\text{Norway}) \\
 & + \text{FinTech} \cdot \text{digicap} \cdot \text{small} \cdot \text{HUB} \cdot \text{DME} && (\text{Hungary}) \\
 & + \text{FinTech} \cdot \text{digicap} \cdot \text{small} \cdot \text{hub} \cdot \text{dme} && (\text{Greece})
 \end{aligned}$$

where upper case indicates presence and lower case absence of a condition. Solution consistency is 1.000 but raw coverage is only 0.455, so the solution accounts for five of the eleven adopters (Table 4). The gap between the two figures should be read at face value, since at a conventional threshold the data support exactly one repeated adoption configuration, the Baltic pair, together with three single-case terms that describe individual countries rather than classes of cases and should not carry generalizing weight.

Table 4: Adoption solution terms (consistency threshold 0.75, conservative solution)

Configuration	Cases	Consistency	Raw coverage
FINTECH * DIGICAP * SMALL * HUB * DME	Latvia; Lithuania	1.000	0.182
fintech * DIGICAP * SMALL * hub * dme	Norway	1.000	0.091
fintech * digicap * small * hub * dme	Greece	1.000	0.091
fintech * digicap * small * HUB * DME	Hungary	1.000	0.091

Note: Upper case indicates presence, lower case absence. Overall solution consistency 1.000; coverage 0.455. HUB denotes an innovation hub established by 2018.



The Baltic term gives the clearest joint support to E1 and E3. Latvia and Lithuania combine every condition, including dependent digital integration, and are the only cases to do so. The configuration matches the positioning account of small jurisdictions converting credible supervision into a competitive asset, but the Baltic case literature shows the profile was permissive rather than determining, since Estonia shares almost all of it, and the actual decisions turned on policy entrepreneurs and perceptions of legal room (Raudla et al. 2024, 2025; Mintrom and Norman 2009). Norway fits E2 in its purest form, as a small, high-capacity supervisor that adopted a sandbox without above-median FinTech density and without a prior hub, a pattern that suggests controlled learning rather than market response. Hungary, an early adopter in 2018 whose central bank has pursued a visibly entrepreneurial innovation agenda, is consistent with the translation logic of E3, since it combines DME status with an early hub. Greece is the limiting case, because its term is defined by the absence of every condition, which says only that the available structural proxies do not capture whatever drove the Bank of Greece's 2021 adoption; post-crisis supervisory modernization is plausible, but establishing it is a matter for process evidence rather than for this design.

Contradictory configurations

Six adopters (Austria, Denmark, Italy, Malta, Slovakia, and Spain) are not covered by the solution because each sits in a configuration also occupied by non-adopters. Table 5 lists these contradictory rows, which carry much of the comparative information in the data, and the section that follows this one takes the most instructive of them to the case level.

Not every contradiction can bear interpretation. Denmark against Iceland (small, FinTech-dense, high-capacity, no early hub) and Slovakia against Slovenia (small

Table 5: Contradictory configurations: identical measured profiles with divergent outcomes

Configuration	Adopters	Non-adopters	Consistency
fintech * DIGICAP * small * HUB * dme	Austria, Spain	Germany	0.67
fintech * digicap * SMALL * hub * DME	Slovakia	Slovenia	0.50
FINTECH * DIGICAP * SMALL * hub * dme	Denmark	Iceland	0.50
FINTECH * DIGICAP * SMALL * HUB * dme	Malta	Finland, Ireland	0.33
fintech * digicap * small * HUB * dme	Italy	Belgium, France, Portugal	0.25

DME states) are single-case contrasts on which a cross-national design cannot adjudicate, and they are treated here as material for future process research rather than as evidence for any expectation, a status Denmark's FT Lab and Slovakia's 2022 adoption share with Greece. What the contradictory rows collectively establish is nonetheless robust, namely that for most of the adopter set, measured structural conditions underdetermine the outcome.

Non-adoption is better structured than adoption

The finding that adoption resists structural explanation would be of limited interest on its own, since it could simply mean the conditions were badly chosen. What gives it substance is that the same five conditions organize the other side of the outcome rather well, which matters for anyone who reads the European sandbox map as a ranking of regulatory attitude, because it implies that the informative question is not who sandboxes but who declines to. Minimizing the non-adoption truth table at the same 0.75 threshold produces a markedly stronger solution, with consistency 0.933 and coverage 0.737, covering fourteen of the nineteen non-adopters (Table 6).

The terms describe recognizably different non-adoption logics, supporting E4. One

Table 6: Non-adoption solution terms (consistency threshold 0.75, conservative solution)

Configuration	Covered cases	Deviant case	Consistency	Raw coverage
FINTECH * digicap * SMALL * dme	Cyprus; Luxembourg; Liechtenstein	–	1.000	0.158
FINTECH * digicap * small * DME	Czechia; Poland	–	1.000	0.105
digicap * small * hub * DME	Bulgaria; Czechia; Romania	–	1.000	0.158
DIGICAP * SMALL * hub * DME	Croatia; Estonia	–	1.000	0.105
fintech * digicap * small * HUB * dme	Belgium; France; Portugal	Italy	0.750	0.158
FINTECH * DIGICAP * small * HUB * dme	Netherlands; Sweden	–	1.000	0.105

Note: Overall solution consistency 0.933; coverage 0.737. Italy is the only deviant case, discussed in the text.

is the alternative-instrument equilibrium of large, high-capacity FinTech markets, in which the Netherlands and Sweden combine above-median FinTech density, high digital capacity, and early hubs, and channel facilitation through tailor-made supervision and innovation hubs rather than testing schemes. A related term captures established hub models in Belgium, France, and Portugal, with Italy as the solution's only deviant case, its statutory sandbox cutting against an otherwise stable hub equilibrium. A second logic concerns financial centres, since Luxembourg, Cyprus, and Liechtenstein host dense regulated payment sectors yet rely on licensing and guidance, a reminder that jurisdictional competition does not require a sandbox and a caution against strong readings of E1. The remaining terms cover the post-socialist cases outside the Baltic profile, namely Czechia, Poland, Bulgaria, Romania, Croatia, and Estonia, and these are not cases of regulatory inaction. Poland's supervisor operates an innovation hub and a virtual testing environment while its sandbox project has remained suspended, and Czechia combines a central-bank contact point with an agency-run testing programme launched in 2024, which sits outside the supervisory live-testing definition used here. The pattern supports the conditional reading of E3,



in that dependent market integration without the Baltic enabling conditions has so far gone together with hubs and partial or non-supervisory arrangements rather than with supervised live testing.

The five uncovered non-adopters (Germany, Finland, Iceland, Ireland, and Slovenia) all sit in the contradictory rows of Table 5, so they are structurally indistinguishable from adopters within the measured space. Germany and Finland are the paradigmatic high-capacity non-adopters, with measured profiles resembling those of Austria, Spain, or Malta, and their non-adoption is more plausibly an instrument choice than a capacity constraint. The asymmetry between the two solutions is the paper's central empirical result, since the conditions that fail to predict who sandboxes succeed considerably better at describing who does not.

Interpreting the Contrasts

The configurational results identify where similar countries chose differently, and this section takes the four most instructive contrasts to the case level. The vignettes are compact by design, drawing on supervisory documents and the secondary literature, and each is tied to a specific row of the truth table. Their purpose is to show that the mechanisms behind identical measured profiles differ in ways the conditions cannot register, which is the substantive content of the underdetermination claim.

Lithuania and Latvia against Estonia

The sharpest minimal contrast in the data involves a consistent row rather than a contradictory one. Estonia differs from Latvia and Lithuania on exactly one measured condition, since its innovation hub arrived in 2021 rather than by 2018, yet the substantive difference is a deliberate institutional choice. Lithuania moved first



and most aggressively, embedding its 2018 sandbox in a broader central-bank strategy to attract FinTech firms through fast licensing and a dedicated newcomer programme, while Latvia followed only in 2021, after several years in which its hub carried the facilitation load. Estonia, the region's digital-government showcase, built a communication-based hub, added a monitoring-oriented test environment in 2023, and explicitly declined the live-testing model. Case research based on interviews with the participating officials attributes the divergence to policy entrepreneurs inside the Lithuanian central bank, to diffusion through regulatory networks, and to Estonian perceptions that legal room and administrative payoff were both lacking, rather than to any structural difference the region's statistics can register (Raudla et al. 2024, 2025). The episode matches the general finding of the policy-entrepreneurship literature that permissive structural conditions leave the decision to identifiable actors (Mintrom and Norman 2009), and it makes Estonia arguably the strongest single case against equating digital-state capacity with sandbox adoption.

Austria and Spain against Germany

The high-capacity early-hub row containing Austria, Spain, and Germany speaks directly to E2 and its limits, because the same supervisory-learning profile produced three different institutional routes. Austria's supervisor created a sandbox within its own mandate in 2020, attaching it to an existing FinTech contact point. Spain took the statutory route in the same year, anchoring live testing in Law 7/2020 with a Treasury-administered testing space in which the sectoral supervisors participate. Germany, operating with the same capacity endowment and an earlier contact-point infrastructure, chose neither, and BaFin has consistently presented ordinary supervision and case-by-case guidance as providing equivalent support without the equal-treatment problems a selective testing scheme raises.⁶ Capacity, on this evidence,

⁶BaFin's position is documented in its published responses on regulatory sandboxes, which argue that the German supervisory framework accommodates innovative business models without a



enables experimentation without requiring it, which is the pattern E2 cannot generate on its own, and the German stance illustrates the responsive-regulation argument that facilitation can be absorbed into ordinary supervision as a deliberate strategy (Black and Baldwin 2010).

Malta against Finland and Ireland

The full small-state profile without DME, shared by Malta, Finland, and Ireland, makes the same point for E1, since small-state competitive incentives proved compatible with three different instrument choices. Malta institutionalized live testing in 2020 as part of a wider strategy of regulatory specialization in financial and digital services. Finland has not established a domestic live-testing sandbox and continues to rely on its hub-based model. Ireland adopted the sandbox label without the live-testing design, and the advisory cohort programme its central bank launched in 2024 delivers structured guidance to selected firms without market testing or regulatory relief. The Irish case is consistent with the possibility that the sandbox label carries signalling value even where the underlying design remains advisory (Béland and Howlett 2016; Voß and Simons 2014).

Italy against Belgium, France, and Portugal

The early-hub-only row, in which Italy stands against Belgium, France, and Portugal, isolates what is distinctive about the Italian case. All four are larger states whose supervisors built hub or contact-point infrastructures in the mid-2010s and settled into a facilitation equilibrium around them. Among them only Italy legislated a sandbox, embedding it in a national FinTech statute and an inter-authority committee that

dedicated testing regime (<https://www.bafin.de>). Germany appears in European Supervisory Authorities (2023) with an innovation hub only.



coordinates the banking, securities, and insurance supervisors. The statutory route shifted the formal venue of adoption from supervisory discretion toward legislation and inter-authority coordination, and Spain's law-based testing space has the same signature. In these cases at least, where hub equilibria were entrenched at the supervisory level, live testing arrived through legislative rather than supervisory channels, a pattern consistent with the multilevel argument that national experimentation is shaped by which institutional venue takes ownership of it ([Héritier and Lehmkuhl 2008](#)).

Validating the Configurational Results

Because both the outcome coding and the analytic thresholds involve judgement, this section examines how much the headline results depend on them, and it adds a regression cross-check for readers who want the additive alternative examined on the same data. The intent is practical, since a configurational claim is only worth interpreting if it survives reasonable changes to the decisions that produced it.

The results are not artefacts of the outcome boundary or the threshold. Broadening the outcome to the inclusive coding, which adds Ireland's advisory programme, leaves the adoption solution terms unchanged at the 0.75 threshold, since Ireland joins the already contradictory row it shares with Malta and Finland, and it leaves the non-adoption terms identical. Raising the non-adoption threshold to 0.80 removes only the hub-equilibrium term containing the Italian deviant and yields a perfectly consistent five-term solution covering eleven non-adopters. Lowering the adoption threshold to two-thirds adds the row containing Austria and Spain with Germany as a deviant and lifts coverage to 0.64, at the cost of admitting exactly the contradictions Table 5 reports, so the stricter threshold is retained. Only seventeen of thirty-two possible configurations are observed, which means limited diversity is substantial, and no counterfactual minimization is attempted.



The Estonia coding receives a dedicated check, given its weight in the Baltic contrast and its listing as a sandbox operator in the ESAs inventory. Recoding Estonia as an adopter and re-minimizing both truth tables changes remarkably little. The adoption solution keeps its structure, with the Baltic term absorbing Estonia and simplifying to $\text{FINTECH} \cdot \text{DIGICAP} \cdot \text{SMALL} \cdot \text{DME}$, since the hub-timing condition no longer distinguishes the covered cases, and solution consistency remains 1.000 with coverage rising to 0.500. The non-adoption solution retains six terms with consistency 0.929 and coverage 0.722, and Italy remains the only deviant. The substantive conclusions, one repeated Baltic adopter profile and a well-structured non-adoption landscape, therefore hold under either coding, although the minimal contrast between Estonia and its neighbours naturally disappears if Estonia is counted as an adopter. The main coding is retained because the Estonian test environment lacks the bounded live test that defines the outcome, but readers who weigh the ESAs listing differently lose only the vignette, not the results.

A small regression exercise checks whether the configurational reading misses a simple additive pattern. With thirty cases and eleven positives it is descriptive only, and the full tables and coefficient plot are reported in Appendix C. A logit of adoption on standardized log FinTech density, standardized digital capacity, and the small-state, early-hub, and DME indicators finds no reliable additive association, as the FinTech coefficient is negative (log-odds -0.80 , standard error 0.59), digital capacity is essentially zero, and small-state status has the largest but highly imprecise positive coefficient. The model's classification errors point in the same direction, since it misses eight of the eleven adopters and falsely predicts adoption for Ireland, Germany, Finland, Slovenia, and Croatia, precisely the cases the QCA identifies as structurally indistinguishable from adopters. The cross-check is thus consistent with the configurational account rather than a rival to it, in that no linear structural story fits these data, and the FinTech-pressure baseline can be set aside.



Discussion

For the policy instruments literature, the findings give empirical content to the claim that instrument choice is a substantive political decision rather than a technical one ([Lascoumes and Le Galès 2007](#)). The European sandbox record shows a single template producing supervisor-run schemes, statutory testing spaces, advisory programmes, and deliberate abstention across jurisdictions whose measured profiles are often identical, which is difficult to reconcile with any functional account in which problem pressure selects the tool. The more useful frame is the policy mix, since every jurisdiction in the sample operates some facilitation instrument, and the observed choice is between mixes that include a live-testing component and mixes that do the same work through hubs, guidance, and ordinary supervision ([Howlett and Rayner 2007](#); [Tönurist, Kattel, and Lember 2017](#)). Non-adoption, on this reading, is a designed feature of a mix rather than the absence of design, and the strength of the non-adoption solution suggests that such equilibria are systematically maintained rather than accidental. The claim should be kept at its proper altitude, since the QCA maps where these equilibria lie and the vignettes indicate, without demonstrating, how they are maintained; establishing the mechanisms would require process evidence of the kind the Baltic case studies provide.

For diffusion research, the results illustrate why adoption counts mislead when the diffusing object is a label rather than a design ([Maggetti and Gilardi 2016](#)). The sandbox travelled through an active instrument constituency, and what different jurisdictions imported ranged from the full live-testing commitment to the name alone, with Ireland's advisory programme as the limiting case ([Béland and Howlett 2016](#); [Voß and Simons 2014](#)). Cross-national indices that count sandboxes as markers of innovation-friendly regulation therefore aggregate institutionally distinct objects, and rankings built on them risk misclassifying both the high-capacity non-adopters, whose abstention reflects instrument choice, and the label adopters, whose schemes



involve no experimentation. The practical implication for comparative work is that the unit worth counting is the institutional commitment, the bounded live test with supervisory ownership of risk, rather than the vocabulary.

For the experimentation literature, the European record suggests that the demanding variant of policy experimentation diffuses on a different track from experimental rhetoric (Huitema et al. 2018; Ansell and Bartenberger 2016). Live testing concentrated in two kinds of settings, small open jurisdictions with positioning incentives and larger states where live testing was created by statute rather than by the supervisor, while the region's most capable administrations relied on hubs and ordinary supervision without establishing a domestic live-testing sandbox. The point has forward-looking relevance, because EU legislation has begun to mandate regulatory sandboxes in adjacent domains, most prominently the requirement that every member state establish an artificial-intelligence regulatory sandbox.⁷ The FinTech experience implies that harmonizing the label will not harmonize the institution, and that mandated schemes will land in national instrument mixes whose existing equilibria differ in exactly the ways documented here.

Conclusion

This paper asked why some EU and EFTA countries adopt FinTech regulatory sandboxes while others rely on alternative facilitation tools. Answering the question required separating supervised live testing from advisory programmes, statutory exemptions, and innovation hubs. Two findings stand out. The first is that structural conditions underdetermine adoption, since no condition is close to necessary or sufficient, adopters are drawn disproportionately from below-median regulated FinTech density, and only the Latvian and Lithuanian profile recurs. Most adopters

⁷Regulation (EU) 2024/1689 (the Artificial Intelligence Act), art. 57, requires each member state to ensure the establishment of at least one AI regulatory sandbox by August 2026.



share their measured profile with non-adopters, and the decisive differences, Estonia against its Baltic neighbours, Germany against Austria and Spain, Finland and Ireland against Malta, ran through instrument choice rather than structural endowment. The second is that non-adoption has a well-structured institutional logic of its own, in which high-capacity markets absorb facilitation into hubs and ordinary supervision, financial centres rely on licensing and guidance, and post-socialist jurisdictions outside the Baltic profile combine hubs with partial or non-supervisory testing arrangements.

These claims are bounded. Crisp sets, imperfect proxies, and substantial limited diversity mean the solutions describe the case landscape rather than establish causal recipes, and the single-case adoption terms, Greece, Hungary, and to a lesser extent Slovakia and Denmark, require process evidence that cross-national conditions cannot supply. The measurement agenda follows directly, since supervisory capacity should be measured more directly than through e-government indices, and FinTech pressure through broader indicators than the PSD2 firm register.

The findings nonetheless carry implications beyond the sandbox debate. Within the single market, national innovation-facilitation instruments are positional goods, because passporting raises the strategic value of national authorization and supervisory reputation, which is why small jurisdictions may loom large among adopters and why the label spreads faster than the design. Reading sandbox counts as rankings of regulatory modernity mistakes an instrument choice for a capability. The European FinTech sandbox map is therefore better read as a record of how differently situated supervisors chose to govern the same uncertainty, some through bounded experiments and most through dialogue and ordinary supervision.

Appendix

Appendix A: Case-Level Calibration

Table 7: Case-level crisp memberships (main QCA sample, n = 30)

Country	Live-testing sandbox	Inclusive coding	FINTECH	DIGICAP	SMALL	HUB	DME
Austria	1	1	0	1	0	1	0
Denmark	1	1	1	1	1	0	0
Greece	1	1	0	0	0	0	0
Hungary	1	1	0	0	0	1	1
Italy	1	1	0	0	0	1	0
Latvia	1	1	1	1	1	1	1
Lithuania	1	1	1	1	1	1	1
Malta	1	1	1	1	1	1	0
Norway	1	1	0	1	1	0	0
Slovakia	1	1	0	0	1	0	1
Spain	1	1	0	1	0	1	0
Belgium	0	0	0	0	0	1	0
Bulgaria	0	0	0	0	0	0	1
Croatia	0	0	0	1	1	0	1
Cyprus	0	0	1	0	1	1	0
Czechia	0	0	1	0	0	0	1
Estonia	0	0	1	1	1	0	1
Finland	0	0	1	1	1	1	0
France	0	0	0	0	0	1	0
Germany	0	0	0	1	0	1	0
Iceland	0	0	1	1	1	0	0
Ireland	0	1	1	1	1	1	0
Liechtenstein	0	0	1	0	1	1	0
Luxembourg	0	0	1	0	1	0	0
Netherlands	0	0	1	1	0	1	0
Poland	0	0	1	0	0	1	1
Portugal	0	0	0	0	0	1	0
Romania	0	0	0	0	0	0	1
Slovenia	0	0	0	0	1	0	1

Table 7: Case-level crisp memberships (main QCA sample, $n = 30$) (*continued*)

Country	Live-testing sandbox	Inclusive coding	FINTECH	DIGICAP	SMALL	HUB	DME
Sweden	0	0	1	1	0	1	0

Switzerland (statutory exemption, 2017) is excluded from the QCA sample because PSD2 density is not comparable outside the EU/EEA.

Appendix B: Truth Table and Necessity Diagnostics

Table 8: Observed truth table for the live-testing outcome

FINTECH	DIGICAP	SMALL	HUB	DME	n	Adopters	Consistency	Cases
0	0	0	0	0	1	1	1.00	Greece
0	0	0	0	1	2	0	0.00	Bulgaria; Romania
0	0	0	1	0	4	1	0.25	Belgium; France; Italy; Portugal
0	0	0	1	1	1	1	1.00	Hungary
0	0	1	0	1	2	1	0.50	Slovakia; Slovenia
0	1	0	1	0	3	2	0.67	Austria; Germany; Spain
0	1	1	0	0	1	1	1.00	Norway
0	1	1	0	1	1	0	0.00	Croatia
1	0	0	0	1	1	0	0.00	Czechia
1	0	0	1	1	1	0	0.00	Poland
1	0	1	0	0	1	0	0.00	Luxembourg
1	0	1	1	0	2	0	0.00	Cyprus; Liechtenstein
1	1	0	1	0	2	0	0.00	Netherlands; Sweden
1	1	1	0	0	2	1	0.50	Denmark; Iceland
1	1	1	0	1	1	0	0.00	Estonia
1	1	1	1	0	3	1	0.33	Finland; Ireland; Malta
1	1	1	1	1	2	2	1.00	Latvia; Lithuania

Table 9: Necessity diagnostics for adoption and non-adoption

Condition	Consistency (adoption)	Coverage (adoption)	Consistency (non-adoption)	Coverage (non-adoption)
FINTECH	0.364	0.267	0.579	0.733
~FINTECH	0.636	0.467	0.421	0.533
DIGICAP	0.636	0.467	0.421	0.533
~DIGICAP	0.364	0.267	0.579	0.733
SMALL	0.545	0.400	0.474	0.600
~SMALL	0.455	0.333	0.526	0.667
HUB	0.636	0.389	0.579	0.611
~HUB	0.364	0.333	0.421	0.667
DME	0.364	0.364	0.368	0.636
~DME	0.636	0.368	0.632	0.632

Appendix C: Regression Results

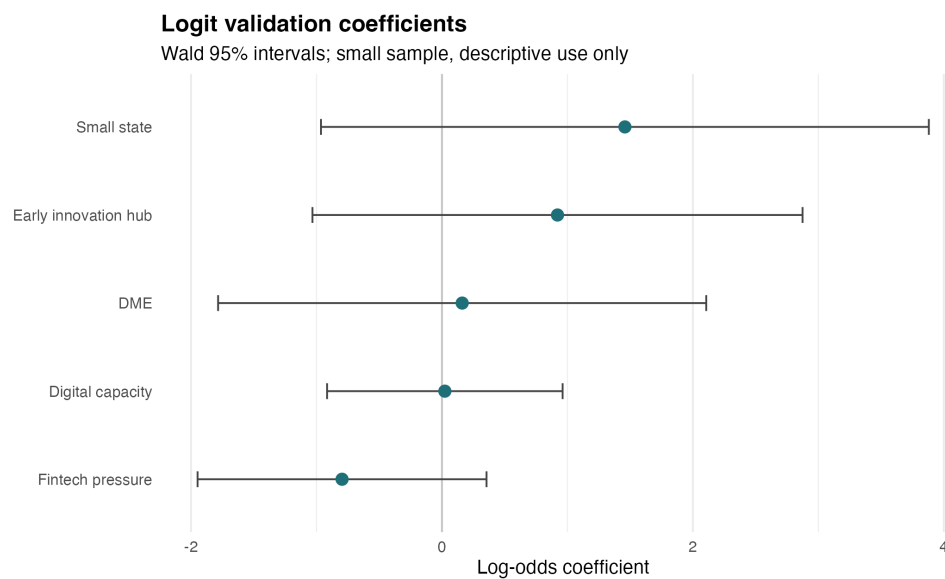


Figure 3: Logit coefficients for the adoption cross-check

The linear probability version of the full model has an R-squared of 0.076. Classified at a 0.5 threshold, the full logit correctly identifies three of eleven adopters (Slovakia, Latvia, Austria) and falsely predicts adoption for Ireland, Slovenia, Croatia, Germany, and Finland.

Table 10: Logit models for the live-testing outcome ($n = 30$)

Model	Term	Log-odds	Std. error	Odds ratio	p value
Logit core	FinTech pressure (log, std.)	-0.593	0.532	0.55	0.266
	Digital capacity (std.)	0.051	0.417	1.05	0.902
	Small state	0.935	1.025	2.55	0.362
Logit full	FinTech pressure (log, std.)	-0.796	0.587	0.45	0.175
	Digital capacity (std.)	0.023	0.479	1.02	0.962
	Small state	1.458	1.236	4.30	0.238
	Early innovation hub	0.921	0.997	2.51	0.355
	DME	0.161	0.993	1.17	0.871

References

- Ansell, Christopher K., and Martin Bartenberger. 2016. “Varieties of Experimentalism.” *Ecological Economics* 130: 64–73. <https://doi.org/10.1016/j.ecolecon.2016.05.016>.
- Arner, Douglas W., Janos Nathan Barberis, and Ross P. Buckley. 2015. “The Evolution of Fintech: A New Post-Crisis Paradigm?” *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.2676553>.
- Bank, World. 2024. “World Development Indicators.” Washington DC. <https://databank.worldbank.org/source/world-development-indicators>.
- Béland, Daniel, and Michael Howlett. 2016. “How Solutions Chase Problems: Instrument Constituencies in the Policy Process.” *Governance* 29 (3): 393–409. <https://doi.org/10.1111/gove.12179>.
- Black, Julia, and Robert Baldwin. 2010. “Really Responsive Risk-Based Regulation.” *Law & Policy* 32 (2): 181–213. <https://doi.org/10.1111/j.1467-9930.2010.00318.x>.
- Bohle, Dorothee, and Béla Greskovits. 2012. *Capitalist Diversity on Europe’s Periphery*. Ithaca, NY: Cornell University Press.
- Borrás, Susana, and Charles Edquist. 2013. “The Choice of Innovation Policy



- Instruments.” *Technological Forecasting and Social Change* 80 (8): 1513–22. <https://doi.org/10.1016/j.techfore.2013.03.002>.
- Buckley, Ross P., Douglas W. Arner, Robin Veidt, and Dirk A. Zetzsche. 2019. “Building Fintech Ecosystems: Regulatory Sandboxes, Innovation Hubs and Beyond.” *Washington University Journal of Law & Policy* 61: 55–98. <https://doi.org/10.2139/ssrn.3455872>.
- Cornelli, Giulio, Sebastian Doerr, Leonardo Gambacorta, and Ouarda Merrouche. 2024. “Regulatory Sandboxes and Fintech Funding: Evidence from the UK.” Edited by Christine Parlour. *Review of Finance* 28 (1): 203–33. <https://doi.org/10.1093/rof/rfad017>.
- Dunlop, Claire A. 2017. “The Irony of Epistemic Learning: Epistemic Communities, Policy Learning and the Case of Europe’s Hormones Saga.” *Policy and Society* 36 (2): 215–32. <https://doi.org/10.1080/14494035.2017.1322260>.
- Dunlop, Claire A. 2009. “Policy Transfer as Learning: Capturing Variation in What Decision-Makers Learn from Epistemic Communities.” *Policy Studies* 30 (3): 289–311. <https://doi.org/10.1080/01442870902863869>.
- Dunlop, Claire A., and Claudio M. Radaelli. 2013. “Systematising Policy Learning: From Monolith to Dimensions.” *Political Studies* 61 (3): 599–619. <https://doi.org/10.1111/j.1467-9248.2012.00982.x>.
- Dusa, Adrian. 2019. *QCA with r: A Comprehensive Resource*. Cham: Springer International Publishing. <https://doi.org/10.1007/978-3-319-75668-4>.
- European Banking Authority. 2026. “Registers and Data: Payment and Electronic Money Institutions.” 2026. <https://www.eba.europa.eu/risk-analysis-and-data/registers>.
- European Supervisory Authorities. 2023. “Joint ESAs Report on Innovation Facilitators: Update on the Functioning of Innovation Facilitators – Innovation Hubs and Regulatory Sandboxes.” ESA 2023 27. European Supervisory Authorities (EBA, EIOPA, ESMA). https://www.esma.europa.eu/sites/default/files/2023-12/ESA_2023_27_Joint_ESAs_Report_on_Innovation_Facilitators_2023.pdf.



- Gilardi, Fabrizio. 2010. “Who Learns from What in Policy Diffusion Processes?” *American Journal of Political Science* 54 (3): 650–66. <https://doi.org/10.1111/j.1540-5907.2010.00452.x>.
- Héritier, Adrienne, and Dirk Lehmkuhl. 2008. “The Shadow of Hierarchy and New Modes of Governance.” *Journal of Public Policy* 28 (1): 1–17. <https://doi.org/10.1017/S0143814X08000755>.
- Howlett, Michael. 2009. “Governance Modes, Policy Regimes and Operational Plans: A Multi-Level Nested Model of Policy Instrument Choice and Policy Design.” *Policy Sciences* 42 (1): 73–89. <https://doi.org/10.1007/s11077-009-9079-1>.
- Howlett, Michael, and Jeremy Rayner. 2007. “Design Principles for Policy Mixes: Cohesion and Coherence in ‘New Governance Arrangements’.” *Policy and Society* 26 (4): 1–18. [https://doi.org/10.1016/S1449-4035\(07\)70118-2](https://doi.org/10.1016/S1449-4035(07)70118-2).
- Huitema, Dave, Andrew Jordan, Stefania Munaretto, and Mikael Hildén. 2018. “Policy Experimentation: Core Concepts, Political Dynamics, Governance and Impacts.” *Policy Sciences* 51 (2): 143–59. <https://doi.org/10.1007/s11077-018-9321-9>.
- James, Scott, and Lucia Quaglia. 2023. “Epistemic Contestation and Interagency Conflict: The Challenge of Regulating Investment Funds.” *Regulation & Governance* 17 (2): 346–62. <https://doi.org/10.1111/rego.12457>.
- Jeník, Ivo, and Schan Duff. 2020. “How to Build a Regulatory Sandbox: A Practical Guide for Policy Makers.” Technical Guide. Washington, DC: CGAP/World Bank. <https://documents1.worldbank.org/curated/en/126281625136122935/pdf/How-to-Build-a-Regulatory-Sandbox-A-Practical-Guide-for-Policy-Makers.pdf>.
- Kranke, Matthias, and David Yarrow. 2019. “The Global Governance of Systemic Risk: How Measurement Practices Tame Macroprudential Politics.” *New Political Economy* 24 (6): 816–32. <https://doi.org/10.1080/13563467.2018.1545754>.
- Langley, Paul, and Andrew Leyshon. 2021. “The Platform Political Economy of FinTech: Reintermediation, Consolidation and Capitalisation.” *New Political*



- Economy* 26 (3): 376–88. <https://doi.org/10.1080/13563467.2020.1766432>.
- . 2023. “FinTech Platform Regulation: Regulating with/Against Platforms in the UK and China.” *Cambridge Journal of Regions, Economy and Society* 16 (2): 257–68. <https://doi.org/10.1093/cjres/rsad005>.
- Lascoumes, Pierre, and Patrick Le Galès. 2007. “Introduction: Understanding Public Policy Through Its Instruments—from the Nature of Instruments to the Sociology of Public Policy Instrumentation.” *Governance* 20 (1): 1–21. <https://doi.org/10.1111/j.1468-0491.2007.00342.x>.
- Maggetti, Martino, and Fabrizio Gilardi. 2016. “Problems (and Solutions) in the Measurement of Policy Diffusion Mechanisms.” *Journal of Public Policy* 36 (1): 87–107. <https://doi.org/10.1017/S0143814X1400035X>.
- Markellos, Raphael N., Sean F. Ennis, Bryn Enstone, Anastasios Manos, Dimitrios Pazaitis, and Dimitrios Psychoyios. 2024. “Worldwide Adoption of Regulatory Sandboxes: Drivers, Constraints and Policies.” *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4764911>.
- Mintrom, Michael, and Phillipa Norman. 2009. “Policy Entrepreneurship and Policy Change.” *Policy Studies Journal* 37 (4): 649–67. <https://doi.org/10.1111/j.1541-0072.2009.00329.x>.
- Nölke, Andreas, and Arjan Vliegenthart. 2009. “Enlarging the Varieties of Capitalism: The Emergence of Dependent Market Economies in East Central Europe.” *World Politics* 61 (4): 670–702. <https://doi.org/10.1017/S0043887109990098>.
- Philippon, Thomas. 2017. “The FinTech Opportunity.” BIS Working Paper No. 655. Bank for International Settlements. <https://www.bis.org/publ/work655.pdf>.
- Ragin, Charles C. 2008. *Redesigning Social Inquiry: Fuzzy Sets and Beyond*. Chicago: University of Chicago Press.
- Raudla, Ringa, Egert Juuse, Vytautas Kuokstis, Aleksandrs Cepilovs, and James W. Douglas. 2024. “Regulatory Sandboxes and Innovation Hubs for FinTech: Experiences of the Baltic States.” *European Journal of Law and Economics*, December. <https://doi.org/10.1007/s10657-024-09830-y>.



- Raudla, Ringa, Egert Juuse, Vytautas Kuokštis, Aleksandrs Cepilovs, Vytenis Cipinys, and Matti Ylönen. 2025. “To Sandbox or Not to Sandbox? Diverging Strategies of Regulatory Responses to FinTech.” *Regulation & Governance* 19 (3): 917–32. <https://doi.org/10.1111/rego.12630>.
- Sabel, Charles F., and Jonathan Zeitlin. 2008. “Learning from Difference: The New Architecture of Experimentalist Governance in the EU.” *European Law Journal* 14 (3): 271–327. <https://doi.org/10.1111/j.1468-0386.2008.00415.x>.
- Schneider, Carsten Q., and Claudius Wagemann. 2012. *Set-Theoretic Methods for the Social Sciences: A Guide to Qualitative Comparative Analysis*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139004244>.
- Shipan, Charles R., and Craig Volden. 2008. “The Mechanisms of Policy Diffusion.” *American Journal of Political Science* 52 (4): 840–57. <https://doi.org/10.1111/j.1540-5907.2008.00346.x>.
- Tõnurist, Piret, Rainer Kattel, and Veiko Lember. 2017. “Innovation Labs in the Public Sector: What They Are and What They Do?” *Public Management Review* 19 (10): 1455–79. <https://doi.org/10.1080/14719037.2017.1287939>.
- United Nations Department of Economic and Social Affairs. 2024. “United Nations e-Government Survey 2024: Accelerating Digital Transformation for Sustainable Development.” New York: United Nations Department of Economic; Social Affairs. <https://publicadministration.un.org/egovkb/en-us/Reports/UN-E-Government-Survey-2024>.
- Voß, Jan-Peter, and Arno Simons. 2014. “Instrument Constituencies and the Supply Side of Policy Innovation: The Social Life of Emissions Trading.” *Environmental Politics* 23 (5): 735–54. <https://doi.org/10.1080/09644016.2014.923625>.
- Wu, Xun, M. Ramesh, and Michael Howlett. 2015. “Policy Capacity: A Conceptual Framework for Understanding Policy Competences and Capabilities.” *Policy and Society* 34 (3): 165–71. <https://doi.org/10.1016/j.polsoc.2015.09.001>.
- Zetzsche, Dirk A., Ross P. Buckley, Douglas W. Arner, and Janos Nathan Barberis. 2017. “Regulating a Revolution: From Regulatory Sandboxes to Smart

Regulation.” *Fordham Journal of Corporate & Financial Law* 23 (1): 31–103.
<https://doi.org/10.2139/ssrn.3018534>.