Why and How: Public-Sphere Persuasion and Institutional Legitimacy*

Laure Athias†

July 11, 2026

Abstract

Norms acquire institutional legitimacy when recognized as rationally justifiable. I develop a model of public-sphere persuasion in which individuals evaluate candidate norms along two dimensions, merit and implementability. Endorsement is bundle-based: a norm must be judged desirable and institutionally feasible. In Weberian culture—one expecting strict separation between technocratic administration and normative authority—state agents internalize a role conception providing them a cost advantage in implementability arguments and a handicap in merit ones, so peer and state-agent persuasion need not operate symmetrically. A second cultural margin is symmetric: both channels' persuasive effort depends on how far the prevailing grammar of legitimacy makes public justification consequential. I characterize the equilibrium persuasion architecture: when each channel bundles merit and implementability internally, peer and state-agent exposures are substitutes, while a positive interaction is possible only under cross-channel fit. I test the model using Swiss panel data on support for equal institutional treatment of in-group and out-group members, combining content-weighted exposure and workplace-switch designs. Both peer and state-agent exposure raise endorsement while their interaction is negative. Private-to-public non-mission switches raise endorsement by 10 percentage points. Where Weberian culture is stronger, proxied by state secularity support, the state-agent effect weakens and the interaction turns positive—a sign reversal possible only under cross-channel fit. Where culture favors self-legislation, proxied by exposure to French republican state-building, channels' own effects strengthen and the interaction deepens. Cultural change is an endogenous process of norm legitimation whose pace and direction reflect a persuasion architecture shaped by that same cultural environment.

***JEL* Codes:** D02, D63, D71, D73, D83, P16, Z13.

*I am grateful to James A. Robinson for discussions that have nurtured this paper and for providing an intellectual home. I also thank Marius Brülhart, Martin Castillo-Quintana, Jacob Hariri, Leander Heldring, Rachel Kranton, Eduardo Montero, Raul Sanchez de la Sierra, Judith Scheele, Mathias Thoenig, Bruno Ventelou, Nicolas Werquin, Giorgio Zanarone, and Yanos Zylberberg for helpful comments at different stages of the project. I also thank the participants in the HEC Lausanne Public Economics and Policy Seminar, University of Chicago Substantivist Economics and Societies conference, 13th World Congress of the Econometric Society, 29th Annual SIOE Conference, 23rd session of the Institutional and Organizational Economics Academy. I declare no research support, financial relationships, or other potential conflicts of interest.

†Laure Athias, University of Lausanne (Email: laure.athias@unil.ch).

"If spirit is something like the achievement of a collective normative like-mindedness, then there are such norms and we are subject to them only because we have legislated such norms," and yet they can retain "real, genuinely binding, and justifiable normative force."
Robert B. Pippin (2008), p. 17

"The validity of the claims to legitimacy may be based on rational grounds—resting on a belief in the 'legality' of patterns of normative rules and the right of those elevated to authority under such rules to issue commands. [...] In the case of legal authority, obedience is owed to the legally established impersonal order."
Max Weber (1922 [1978, pp. 215–216])

"The modern idea of order, in contradistinction to the medieval Christian ideal, was seen from the beginning as for the here and now. But it definitely migrates along a path, running from the more hermeneutic to the more prescriptive."
Charles Taylor (2004), p. 7

1 Introduction

Institutional legitimacy matters because it affects institutional performance: it shapes voluntary compliance and enforcement frictions (Tyler (2006)) and, more broadly, conditions whether materially similar institutions work across environments (Moscona et al. (2026)). In economics, legitimacy—when it appears—typically enters either through cultural beliefs and values that sustain institutional equilibria (Greif (1994), Tabellini (2008), Acemoglu and Robinson (2025)), or through direct preferences over rights, liberty, or democratic procedures (Besley and Persson (2019), Gratton and Lee (2024)). This paper studies the prior process through which legitimacy is formed: how a candidate norm comes to be recognized as worthy of institutionalization. A long tradition in political philosophy and anthropology argues that formal institutions are not merely technical devices. They give normative meanings institutional form, making the rules they enact intelligible and potentially legitimate (e.g., Castoriadis (1987), Taylor (2004), Malinowski (1922), Sahlins (1976), Sahlins (2000)). In modern societies, unlike premodern orders grounded in inherited, sacred, or customary authority, norms can claim institutional legitimacy only insofar as they are rationally justifiable (Habermas (1990), Pippin (2008)). Legitimacy therefore turns on recognition: whether citizens come to see a candidate norm as rationally justifiable. That recognition is formed—or fails to form—through public justification. Norms therefore cannot be treated only as inherited background conditions; they enter as objects of public justification, evaluated both on their merit and on their institutional implementability.

This paper develops a tractable theory of how institutional legitimacy emerges through persuasion in the public sphere—the arena of public reasoning in which shared judgments about public matters are formed[1]. Because the object of persuasion is a candidate norm for public institutionalization rather than private approval, the relevant question is whether the norm is worthy of institutionalization. Citizens therefore evaluate it along two latent dimensions: *merit* (m), capturing whether the norm is normatively desirable, and *implementability* (h), capturing whether it can be

[1] Taylor (2004) argues that the public sphere is central to institutional legitimacy in modern Western societies, and defines it as a common space in which members of society meet, both face-to-face and through various media, to discuss matters of common interest and "form a common mind about these" (Taylor (2004), p. 83). Habermas (1989) traces the emergence of this rational and critical discursive space of public justification to 18th-century bourgeois society, particularly in Europe, through institutions such as coffeehouses and salons, literary journals and newspapers, and learned societies and clubs, and locates it between the private realm and the state.

administered, enforced, and sustained in practice[2]. Receivers endorse when the posterior expected value of institutionalizing the norm exceeds a threshold. This expected value can be purely material or include non-material components; what matters is that receivers evaluate the consequences of realizing a candidate norm as a rule of public action. Endorsement is thus bundle-based: persuasion is strongest when citizens receive matched reasons that answer both *why* the norm should hold and *how* it can be carried out. I abstract from the extensive margin of competition across candidate norms and focus on the intensive margin: conditional on a norm being on the public agenda, I ask how persuasion shapes endorsement of it as a legitimate rule of public action.

A potential source of asymmetry between producers in the public sphere is institutional role. Legal-rational bureaucracy in the sense of Weber (1922) attaches state agents' authority to office rather than person and makes it valid only within legally defined competences: state agents are structurally required, in the exercise of their office, to justify what can be done, by which rules, and through which administrative means—the normative order being what the office presupposes rather than adjudicates[3]. Where the surrounding culture is Weberian—expecting a strict separation between technocratic administration and normative authority—state agents internalize the legal-rational role conception and carry it into public discourse: their marginal cost of producing precise implementability arguments falls relative to peers, while assertions about what is desirable fall outside the register of claims their role equips them to produce and become correspondingly costlier for them to voice. This motivates distinguishing two persuasion channels in the public sphere—peers and state agents—that need not operate symmetrically. I model each channel as an aggregate producer of public reasons, summarizing the content supplied by the relevant set of actors. Exposure scales informational content: greater reach delivers more of a channel's arguments, and repeated encounters with reasons of given precision make them more persuasive in citizens' posteriors. Each channel chooses how much precision to supply in each dimension, trading off convex effort costs against an internalized legitimacy payoff that combines instrumental gains from compliance with the normative value of living under rules they recognize as publicly justified. Two features of the surrounding cultural environment discipline this producer problem: the Weberian culture, which governs how deeply state agents internalize the legal-rational role conception and hence the cost asymmetry across dimensions, and the prevailing grammar of legitimacy, which governs how much of the common legitimacy payoff producers internalize.

This yields an equilibrium persuasion architecture. When channels supply sufficiently complete bundles internally, peer and state-agent exposures are substitutes: because each channel already provides both merit and implementability reasons, additional exposure to one channel reduces the marginal value of exposure to the other. When channels specialize across dimensions, by contrast,

[2]The model does not presume that citizens explicitly debate the primitives (m, h). Public-sphere exchanges typically proceed through concrete cases, policy controversies, and institutional proposals. Such discourse indirectly conveys information about desirability and implementability, and endorsement updates as public reasons accumulate. Formally, (m, h) are sufficient statistics for the informational content of these arguments in the learning problem.

[3]Read together, Weber's and Taylor's epigraphs contain the argument in miniature. Weber's states what legitimacy is and where authority to act comes from: validity rests on a *belief* in the legality of patterns of *normative* rules—legitimacy as recognition of a normative order as valid—while office confers the right to issue commands under those rules, the authorized apparatus of execution. In this configuration the normative order enters as settled: the office presupposes it and carries it out. Taylor's epigraph names what modernity does to order itself: the modern idea of order migrates from the hermeneutic—a template through which the world is read—to the prescriptive—a project demanding realization. Candidate norms in modern societies are prescriptive objects: they must be argued for, not read off. The legal-rational office, however, stands to the enacted order in a hermeneutic relation: equipped to articulate how a given order is carried out, not to adjudicate why it should bind. The distance between these two stances is the source of the dimension-specific cost structure formalized in the model—implementability arguments cheap, merit arguments costly—and its depth is cultural: the asymmetry is greatest where the surrounding culture holds office-holders most strictly to the hermeneutic stance, and absent where it does not.

receivers must assemble matched argumentative bundles across sources. Specialization therefore creates cross-channel fit and complementarity at low exposure. The two architectures are not symmetric in what they can produce: under within-channel bundling, the exposure interaction is non-positive at every exposure level, so a positive interaction, wherever observed, is possible only under cross-channel fit. This architecture also has dynamic implications. Because endorsement is bundle-based, implementability specialization makes persuasion contingent on cross-channel completion: state-agent exposure supplies the institutional *how* but leaves the normative *why* incomplete, while peer exposure supplies the normative *why* but leaves institutional implementability incomplete. Consensus is therefore harder to build when one type of reason is missing, but dynamic feedback can generate tipping-point dynamics once merit and implementability arguments jointly circulate, even if the normative *why* remains weakly disciplined. Without specialization, the update map inherits only the concavity of the endorsement link, and cultural change is gradual.

I take the model to geocoded individual-level panel data from the Swiss Household Panel (SHP). Switzerland is a useful setting for two reasons. Strong decentralization makes the canton the tier at which candidate norms are institutionally realized, grounding the canton-year as the empirical boundary of the public sphere; and linguistic and cultural segmentation across cantons generates within-country variation in the model's cultural primitives, allowing the persuasion architecture to be identified across distinct public spheres. Switzerland is also an informative case: despite its stable, direct-democratic institutions, its long-run increase in affective polarization is among the steepest in the OECD, comparable to the U.S. trend (Boxell et al. (2024)). I study endorsement of universalism as a principle that can justify state action: whether Switzerland should grant foreigners the same opportunities as Swiss citizens or instead give Swiss citizens preferential opportunities. This question is a direct empirical counterpart to the model's object, since it asks respondents to evaluate a candidate rule of public action rather than report a private attitude in the abstract. Universalism—the moral principle of equal treatment across in-group and out-group members, in contrast to context-dependent moral orders in which obligations are stronger toward kin or in-groups (Enke et al. (2022))—is also a substantively important case. It has been characterized as a central feature of WEIRD moral psychology, rooted in the transformations associated with the Western Church that weakened kin-based particularism and expanded impersonal obligations (Henrich (2020)), and can be situated within the expansion of the modern Western moral order rooted in the new theories of Natural Law (Taylor (2004)). Yet Western societies differ sharply in how universalistic principles are institutionalized (Esping-Andersen (1990)), and whether foreigners should receive equal institutional treatment remains one of the most salient contemporary tests of universalism.

The empirical strategy is designed to identify how endorsement responds to variation in the persuasive force of public justification. The SHP provides annual measures of endorsement, allowing fixed-effects designs. I first operationalize the model's notion of channel-specific effective precision by combining canton-year variation in each channel's universalism tilt with predetermined canton-level reach proxies. The peer channel is built from private-sector workers, with reach proxied by association density; the state-agent channel is built from public-sector workers, with reach proxied by the public-employment share. In this exposure design, I restrict the sample to non-employed individuals so that respondents do not mechanically enter the construction of the content-weighted exposure measures, and to canton non-movers so that variation in the discourse environment is not confounded by endogenous relocation across cantons. Conditional on fixed effects, trends, and rich controls, identification comes from within-canton, over-time changes in each channel's universalism tilt, scaled by predetermined reach. These changes capture variation in the informational content circulating through state-agent and peer channels that is not explained by common shocks, canton-specific trajectories, or observed political and demographic conditions. I interpret this variation as shifts in

channel-specific effective precision. The baseline specification estimates a model-disciplined reduced form in which endorsement responds to peer and state-agent effective precision and their interaction, thereby identifying both own-channel persuasiveness and the persuasion architecture. The main identification concern is that these within-canton changes in channel-specific universalism tilt may partly reflect unobserved canton-year shifts in the broader demand for universalism or in the local discourse environment, rather than channel-specific persuasion alone. I therefore complement this design with a second strategy that delivers sharper individual-level exposure shocks while absorbing the canton-year discourse environment through canton-year fixed effects: within-person switches between private and public employment. To address concerns about dynamic self-selection and mission-driven sorting, I focus on private-to-public switches within non-mission industries, excluding health, education, and social care. In these industries, I find no evidence of differential pre-trends in endorsement prior to the switch.

The empirical results deliver the model's reduced-form building blocks. In the baseline content-weighted exposure regressions, endorsement rises with both peer and state-agent effective precision, conditional on individual and year fixed effects, canton-specific trends, and canton-year controls. Because state-agent effective precision retains a sizeable, statistically robust own effect net of peer exposure and the full control set, the two channels constitute distinct sources of influence: the identity of the source matters, which an undifferentiated-discourse view, in which only total exposure counts, would not by itself generate. Both own effects are positive, moreover, ruling out a pure crowd-out interpretation in which exposure to one channel displaces or discounts the other. In magnitude, peer effective precision has the larger marginal effect: a one-standard-deviation increase in peer content-weighted exposure raises endorsement by 5.5 percentage points, evaluated at mean state-agent exposure, while the corresponding increase in state-agent content-weighted exposure raises it by 4.4 percentage points. The interaction between the two channels is negative and economically sizeable. This rules out a purely additive persuasion technology, in which each channel's contribution to endorsement simply sums; some form of diminishing joint returns is present. On its own, however, the sign of this coefficient is not diagnostic of any particular persuasion architecture since it could follow from the concavity of the endorsement link alone, and would be obtained under any technology combining two exposure measures through a concave aggregator, whether or not that aggregator has any merit/implementability structure. The workplace-switch design provides a sharper exposure shock. Because treatment varies at the individual-year level, I can include canton-year fixed effects and hold fixed the local discourse environment. Moving from private to public employment within non-mission industries raises endorsement by roughly 10 percentage points. This larger within-person estimate is consistent with attenuation in the baseline canton-level reach proxies and with the model's prediction that endorsement rises with effective exposure to persuasive public reasons. The switch design also reveals a sharp institutional-tier pattern: the implied net effect of federal public employment relative to private employment is close to zero, while the effect is concentrated among cantonal and communal state agents. This pattern is consistent with the paper's mechanism operating across institutional tiers. Federal office is plausibly the role most strictly conceived as an implementation and coordination device—the most fully Weberian—so that federal agents' public discourse is most specialized in the institutional *how*. Because endorsement is bundle-based, exposure to an h-specialized channel is weakly persuasive on its own, while cantonal and communal agents, whose roles are conceived less strictly, more often bundle the normative *why* with the institutional *how*. Finally, the reverse-switch estimate is small and statistically insignificant. This pattern is consistent with persistence under learning: once beliefs have shifted, removing exposure need not mechanically restore prior beliefs absent offsetting new information.

Adjacent-outcome tests take up the question of which mechanism drives the identified effects. Each design faces a distinct leading alternative. In the baseline content-weighted exposure design,

where respondents are non-workers, the concern is generic ideological diffusion or a broad shift in the local demand for universalism. The estimates are inconsistent with both: exposure moves neither ideology nor trust in government, and leaves redistribution, support for European Union membership, army, and environmental preferences unchanged. Instead, peer exposure significantly *reduces* support for welfare spending—respondents endorse equal institutional treatment as a rule of public action while supporting limits on public provision, a sign-specific pattern that a leftward or pro-social drift would not produce. In the workplace-switch design, two leading alternatives arise: socialization into a public-sector role, and self-interest as a newly public employee. Each carries its own signature, and the estimates are inconsistent with both. Self-interest predicts stronger preferences for a larger public sector; instead, the switch leaves support for welfare spending and redistribution unchanged. Socialization predicts a broad shift toward public-sector-aligned attitudes; instead, the switch generates no evidence of increased institutional trust and, if anything, lowers satisfaction with democracy while raising endorsement of institutional universalism—a pattern difficult to reconcile with imitation or role internalization, and consistent with state agents making visible the gap between universalistic rules and existing democratic performance. These content-specific signatures are what license reading the estimated effects as persuasion through public reasons; the final step is then to identify the architecture of that persuasion—whether merit and implementability arguments travel bundled within channels or are assembled across them.

I then identify the bundle-based persuasion mechanism by leveraging predetermined cultural environments that should shift the persuasion architecture ex ante. Because the argumentative content circulating within each channel is not directly observed, the empirical strategy is not to claim that these environments directly measure merit or implementability. Rather, I use theoretically and historically grounded shifters that correspond to the model's two cultural primitives—the Weberian culture and the grammar of legitimacy–and that the model predicts should move the architecture in *opposite* directions: one toward cross-channel specialization, the other toward within-channel bundling. The sharpest test exploits the asymmetry established in the model: because a positive peer–state-agent interaction is possible only under cross-channel fit, at any exposure level, observing the interaction change sign across cultural environments is evidence that the architecture itself—not merely the level of universalism—varies with the environment.

The first shifter is Weberian culture. Where the surrounding culture expects a strict separation between technocratic administration and normative authority, state agents internalize a legal-rational role conception and carry it into public discourse: producing implementability arguments becomes cheaper for them, and the opposite for merit arguments. In the model, this tilts the state channel's content mix toward the institutional *how* and away from the normative *why*, so that complete justificatory bundles must increasingly be assembled across channels—the cross-channel fit. I proxy this shifter with popular support for state secularity, the cantonal vote share in the 1980 federal popular initiative for the complete separation of church and state, a continuous, population-level measure of demand for role-separation between technocratic administration and normative authority. The campaign record supports this reading of the vote: the initiative committee argued for separation precisely in the register of administrative competence—citing approvingly the constitutionalist Giacometti's formulation that a state proclaiming freedom of belief removes normative questions from the state's competence—while the Federal Council's message recommending rejection defended the churches' public role on the grounds that the state depends on values it cannot itself supply, ones that allow the citizen not only to obey the law but to inwardly affirm it (Eidgenössisches Aktionskomitee für die Trennung von Staat und Kirche (1978), Schweizerischer Bundesrat (1978)). A canton's vote share therefore captures where its electorate stood on precisely the norm–administration separation the model's shifter is meant to index. To isolate this cultural shifter from other dimensions of cultural environment, I estimate the heterogeneity in the content-weighted exposure design, within

German-speaking cantons only. The results match the mechanism's dimension-specific predictions: as secularity support rises, the state-agent channel's own exposure effect weakens, the peer channel's, if anything, weakens as well—consistent with the model's equilibrium two-sided specialization—and the peer–state-agent interaction turns positive—statistically distinguishable from zero, and positive, in the most secular German-speaking cantons. The model shows this sign reversal is possible only under cross-channel fit: no concave persuasion technology without merit–implementability structure can generate a positive interaction at any exposure level. The pattern survives controlling for historical Protestant population share.

The second shifter, exposure to French republican state-building, operates through a different margin of the model: not the cost structure of a single channel, but both channels' incentive to supply persuasive reasons at all. In the republican tradition associated with Rousseauian political thought, legitimate law is not an external command imposed on individuals; it is binding because citizens can understand it as an expression of collective self-legislation, or the general will (Rousseau (1923)). Under the Third Republic, this state-centered grammar was institutionalized by tying the public sector to substantive civic and moral ends (Weber (1976)). Because this grammar ties a rule's bindingness to citizens' own recognition of it as justified, supplying public reasons is more consequential where it prevails: in the model, it raises the common internalization of the legitimacy payoff, which strengthens both channels' bundles together. I proxy exposure to this grammar with French-speaking cantons in the content-weighted exposure design and French-speaking identity in the workplace-switch design[4]. The switch design carries the sharpest evidence on this margin: the public-employment effect is about three times larger for French-speaking than for German-speaking workers—roughly 21 versus 7 percentage points—within the same canton-year environment. Because the design excludes mission-oriented industries, this heterogeneity is identified from exposure to *ordinary* state agents, not from the sector through which the republican grammar was historically institutionalized. Adjacent outcomes sharpen what the larger effect consists of: what moves alongside institutional universalism is only its closest cognate—support for joining the European Union, universalism applied at the supranational level, up about 23 percentage points—while non-cognate policy attitudes are flat and state-attachment measures, if anything, decline. This is the signature of content-specific persuasion; it also cuts against self-interest, since supranational integration would, if anything, dilute the competences of the national administration the switcher has just joined. The content-weighted exposure design contributes the margins the switch design cannot estimate: French-speaking environments strengthen the peer channel's own exposure effect and deepen the peer–state-agent interaction, a pattern robust to allowing the exposure–endorsement mapping to vary with secularity support or with predetermined cantonal ideology. Overall, the results converge toward stronger within-channel bundling in French-speaking environments.

Taken together, these results support a bundle-based persuasion mechanism and the model's view that predetermined cultural environments shift the persuasion architecture itself—in opposite directions along its two primitives—rather than merely the level of universalism.

The mechanism is not specific to Switzerland, but its portability depends on scope conditions. The model suggests three. First, norms must be contestable: open to public questioning rather than treated as inherited background—the condition that also places propaganda regimes, in which coordinated distortion displaces contestable reasons, outside the model's domain. Second, public reasons must circulate, so that individuals can encounter arguments about both merit and implementability; how far the public sphere is *integrated* then determines not whether the mechanism operates but how specialization plays out, since cross-channel assembly requires both

[4]I use language regions as cultural exposure, not as institutional treatment. This usage follows a literature exploiting the Swiss language border to identify cultural differences while holding formal institutions fixed (e.g., Eugster et al. (2011), Eugster and Parchet (2019)).

channels to reach the same citizens. Third, producers of reasons must internalize at least part of the legitimacy payoff, so that supplying persuasive arguments is individually worthwhile; how much they internalize is itself one of the two cultural primitives along which the architecture varies. Where these conditions hold, what shapes persuasion is not institutional design alone but the cultural environment that configures the architecture through which citizens come to see a norm as both desirable and institutionally realizable. External validity therefore depends less on Swiss institutions per se than on whether other settings satisfy the scope conditions. The Swiss setting combines decentralized authority with rich subnational variation, and it is demanding for the argument: it is not a case in which direct democracy mechanically guarantees broad political participation—turnout has historically been modest despite extensive democratic rights (Hosli et al. (2025)). Formal participation opportunities cannot by themselves carry the results: what the paper's designs identify is variation in the public-reason environment in which institutional legitimacy is formed. The institutional-tier pattern noted above makes this concrete. Within the same canton-year, the endorsement effect of entering public employment is concentrated among cantonal and communal agents; the implied effect of federal employment is close to zero. On the reading developed in Section 5—that the role conception attached to federal office is plausibly the most strictly administrative, so that federal discourse supplies the institutional *how* with little normative *why*—this pattern points to a structural feature of the setting rather than a curiosity of the data. Swiss decentralization, combined with a maximally Weberian top tier, describes a society in which the upper reach of public authority is largely withdrawn from normative argument and norms are institutionalized locally: many cantonal public spheres rather than one national one. Legitimacy formation then operates where those spheres are integrated enough for the *why* and the *how* to meet. The framework thus gives the apparent Swiss paradox a precise form: direct democracy raises the stakes of public justification, but it does not determine the architecture through which merit and implementability reach citizens—and where public discourse increasingly separates the two, a country can be institutionally stable at the level where justificatory bundles form and still experience a steep long-run rise in affective polarization (Boxell et al. (2024)).

This paper contributes to the literature on culture and institutions by formalizing how norms become recognized as worthy of institutionalization. Norms are not only inherited cultural background conditions or informal constraints (North (1990), Greif (1994), Tabellini (2008), Bursztyn et al. (2020)); in modern societies, they can claim institutional legitimacy insofar as they can be justified and realized through institutions (Pippin (2008)). This paper defines institutional legitimacy as the recognition of a candidate norm as both normatively desirable and institutionally feasible. Legitimacy is therefore bundle-based: persuasion must answer not only why a norm should hold, but also how it can be administered, enforced, and sustained in practice. The argument therefore moves from persuasion within bounded social spheres to persuasion over norms that claim general public validity. Families, markets, associations, religious communities, and firms all institutionalize norms, but such norms are typically justified within bounded communities. Norms acquire public institutional legitimacy only when they claim general validity and can be defended as rules of public action. The public sphere is the arena in which such claims to general validity are tested (Habermas (1989), Taylor (2004)). This implies a social rather than purely individual account of agency: individuals appear not only as bearers of private preferences, but as participants in public justification over rules whose legitimacy matters both instrumentally, by sustaining voluntary compliance (Tyler (2006)), and thereby conditioning whether similar institutions work across environments (Moscona et al. (2026)), and non-instrumentally, by shaping the norms under which they collectively live (Pippin (2008)). The instrumental channel carries a comparative implication: if performance depends on endorsement, materially similar institutions will perform differently across environments in which endorsement differs, consistent with the evidence in Moscona et al. (2026).

This perspective contrasts with social-capital accounts, which emphasize trust and cooperation as the channels through which civic engagement and associational life sustain democratic performance (Putnam et al. (1993), Putnam (2000)). Here, civic engagement and associational life matter instead as infrastructure for the public sphere: they sustain the free and informal circulation of public reasons. As in Bayesian persuasion models, persuasion operates by changing beliefs rather than preferences (Kamenica and Gentzkow (2011)); in this paper, the relevant beliefs concern the merit and implementability of institutionalizing a candidate norm. Consequently, private preferences over the norm may diverge from support for its institutionalization (Athias and Ventelou (2026)). This distinction also sharpens the paper's contrast with accounts that locate Western distinctiveness in a specific bundle of private values or psychological traits (Henrich (2020)). Western societies may not be distinctive by virtue of sharing a fixed bundle of values, but by their capacity to justify and institutionalize candidate norms. While Acemoglu and Robinson (2012) stress the role of inclusive institutions in long-run prosperity, this paper therefore highlights a prior legitimacy condition for their effectiveness: the continued public justification of the inclusive norms they institutionalize.

This paper highlights persuasion in the public sphere as a channel through which culture can become fluid: public persuasion allows normative claims to be justified, revised, and translated into institutional change. It thereby contributes to the literature on cultural change (see Acemoglu and Robinson (2025) for a review) by formalizing an endogenous process of norm legitimation in the public sphere: the direction and pace of change are determined by the persuasion architecture through which norms acquire institutional legitimacy, with rapid, self-reinforcing shifts arising when cross-channel bundle completion interacts with public-sphere feedback. Where preference-falsification accounts locate rapid change in the social cost of public expression (Kuran (1995)), here it arises from the structure of justification itself. This dynamic structure parallels accounts of economic growth in which sustained change depends on the accumulation of useful knowledge linking new ideas to the means by which they can be made actionable (Mokyr (2010), Mokyr (2016)). The paper identifies an analogous structure for the expansion of the moral order associated with modern natural-law theories (Taylor (2004)): moral expansion requires not only publicly disciplined claims about what is right, but also credible accounts of how such claims can be realized through institutions.

The link between positive persuasion and normative justification connects the paper to welfare economics and social choice, where a central question is how impartial principles can be justified (Harsanyi (1955), Rawls (1971), Rawls (1997)). The state is central in this process because it is the institutional form through which candidate norms can be scaled into general, enforceable, and administratively sustained rules of public action. In this sense, the state is not merely a contract, an apparatus of enforcement and coercion, or a mechanism for aggregating private interests (Besley and Persson (2011), Besley (2020), Tilly (1975), Tilly (1985), Arrow (1951)); it is the institutional form through which publicly justified norms can become actual as rules of public action (Pippin (2008)). The paper also clarifies why direct-democratic authorization is not a sufficient foundation of institutional legitimacy. Authorization makes public justification consequential, but legitimacy requires candidate norms to be recognized as rationally justifiable. This distinction connects to Castoriadis (1996)'s account of autonomy as collective self-institution rather than the mere presence of elections, rights, or constitutional form—democracy as something a society *does*, not something it *has*. The converse case also exists: Luo and Xue (2025) show that liberal principles can diffuse through moral philosophy and generate political consequences even in autocratic settings—public justification operating where authorization is absent, wherever reasons circulate in contestable form. Authorization and justification are thus separate foundations; the public sphere is among the mechanisms disciplining public authority, democratic or not.

Finally, the paper speaks to the literature on bureaucracy by identifying a double-edged role

for bureaucratic organization in cultural change. The economics literature has emphasized how bureaucratic organization shapes state capacity, implementation, and public-sector performance (Besley et al. (2022), Aneja and Xu (2024)). This paper shows that bureaucratic roles also have public-sphere spillovers, and that these spillovers are culturally contingent. The formal obligations of legal–rational office are common to modern bureaucracies (Weber (1922)); the role conception built on them is not (Ekeh (1975), Weber (1976)). Where the surrounding culture expects a strict separation between technocratic administration and normative authority, state agents internalize a role conception that makes implementability arguments cheap and merit arguments costly for them to supply, so that, even outside formal decision-making, their public-sphere discourse shapes whether candidate norms appear institutionally realizable. This connects the paper to models of competition in persuasion, which study how persuasion changes when receivers hear from more than one source (Gentzkow and Kamenica (2017)). Here the relevant sources are not symmetric senders choosing information structures, but socially distinct channels—peers and state agents—whose culturally constituted roles determine which reasons each can cheaply supply, and whose specialization is two-sided in equilibrium: as the state-agent channel exits merit discourse, peers concentrate upon it. In equilibrium this specialization is a division of justificatory labor and, like the division of labour, it is limited by the extent of its market (Smith (1776)): it improves legitimacy formation only where the public sphere is integrated enough for citizens to assemble complete bundles across channels. The welfare implications of bureaucratic organization therefore run through the structure of the public sphere, not through administrative performance alone.

The cultural variation the paper exploits resonates with Taylor (2007)'s account of secular modernity: secularity is not the subtraction of religion from public life, but a transformation of the conditions under which claims about moral order become contestable and require justification. The paper identifies an empirical trace of that transformation in the persuasion architecture itself: where popular support for confining the state's competence to administrable matters is stronger, the state channel withdraws from the normative *why*, and complete justificatory bundles are assembled across channels rather than within them. This mechanism also bears on the politics of bureaucracy blaming. Hinterleitner (2025) documents that bureaucracy is modern democracies' favorite scapegoat, blamed for almost anything in part because it is unlikely to fight back. The framework endogenizes that silence: it is not an intrinsic property of administration but a culturally produced one—the obedience of the paper's epigraph, treated as a variable rather than a constant. Where the Weberian role conception runs deep, state agents lack the register in which normative blame can be answered, and the apparatus absorbs discontent over norms it implements but cannot publicly justify; where that conception is weaker, the model predicts a state-agent channel with a normative voice of its own. The same conditionality nuances a familiar reading of Arendt (1963), on which the administrative role as such extinguishes the office-holder's normative voice. In this paper, the extinction is culturally imposed and internalized rather than intrinsic to administration. State-agent spillovers can therefore discipline legitimacy formation when implementability is tied to a well-articulated normative *why*; but they can also amplify whichever merit claim prevails when the institutional *how* arrives without strengthening the informational discipline of the merit case itself. Heldring (2026) provides the starkest historical illustration: formerly Prussian municipalities—more efficient at providing public goods and collecting taxes, and not differentially anti-Semitic before 1933—were significantly more efficient at deporting German Jews once the regime's content turned murderous. Administrative excellence does not sign its own welfare effect; the prevailing normative claims do. This distinction speaks to a classic tension in accounts of Enlightenment rationality. Horkheimer and Adorno (1944) warn that instrumental reason, detached from substantive ends, can facilitate domination; Habermas (1989) emphasizes, by contrast, the democratic promise of a public sphere in which reasons remain contestable. This paper locates the hinge between the two poles: which obtains is determined

not by rationalized administration as such, but by the persuasion architecture—by whether the institutional *how* meets a publicly disciplined *why*.

The remainder of the paper is structured as follows. Section 2 presents a model of institutional legitimacy formation in the public sphere, in which persuasion operates over a candidate norm for institutional realization, and derives the conditions under which cultural change is smooth or rapid. Section 3 maps the model to empirical counterparts and lays out the identification strategies. Section 4 contains details on the data. I report the baseline results and the workplace-switch evidence in Section 5, and provide mechanism tests in Section 6. Section 7 concludes.

2 A Model of Public-Sphere Persuasion

This section formalizes institutional legitimacy as the recognition of a candidate norm as both justified and realizable through institutions. Conditional on a norm being on the public agenda, I ask how persuasion shapes endorsement of it as a rule of public action. The object of persuasion is therefore not a norm in the abstract, but a norm proposed for institutional realization. Citizens must assess not only whether the norm is desirable, but also whether it can be realized in practice. I capture this by assuming that citizens evaluate candidate norms along two dimensions: *merit*, which captures whether the norm is desirable and may reflect both material and non-material considerations, and *implementability*, which captures whether the norm can be administered, enforced, and sustained in practice. Endorsement is therefore bundle-based. The model does not presume that citizens explicitly debate the primitives (m, h). Public-sphere exchanges typically proceed through concrete cases, policy controversies, and institutional proposals. Such discourse indirectly conveys information about desirability and implementability, and endorsement updates as public reasons accumulate. Formally, (m, h) are sufficient statistics for the informational content of these arguments in the learning problem. A useful illustration is the Zucman proposal for a coordinated minimum taxation standard on ultra-high-net-worth individuals (Zucman (2024)). What is at stake is not merely support for a tax instrument, but endorsement of a norm for institutionalization: that the ultra-rich should not be governed by a different fiscal principle from everyone else. Its persuasive force depends both on merit—equal civic standing, tax fairness, and the material consequences of the proposed fiscal rule—and on implementability—the claim that this principle can now be realized through workable domestic instruments and international administrative coordination. The model formalizes this persuasion problem.

2.1 Model Setup

A continuum of citizens i in a public sphere c at date t considers whether to endorse a norm as a rule of public action. Let $y_{i,c,t} \in \{0, 1\}$ denote the observed public stance, and

$$s_{i,c,t} \;\equiv\; \mathbb{E}[y_{i,c,t} \,|\, \mathcal{F}_{i,c,t}] \in [0, 1)$$

be the endorsement propensity given citizen i's information set $\mathcal{F}_{i,c,t}$, generated by the public arguments to which she is exposed. Public arguments are reasons circulating in publicly contestable form. I define the *public sphere* in cell (c, t) as the set of venues and interactions in which reasons about common affairs are produced and exchanged. Concretely, it includes (i) mass and local media; (ii) open civic forums, such as associations, clubs, and assemblies; and (iii) informal interpersonal communication. In the model, citizens observe only the subset of arguments to which they are exposed; I formalize this below through the information set $\mathcal{F}_{i,c,t}$. I abstract from censorship: any agent can in principle contribute to and access this discourse, though exposure varies across citizens.

Endorsement concerns institutionalization, so two latent dimensions matter for persuasion: a *merit* dimension m, which captures whether the norm is desirable and may reflect material consequences, normative appeal, or both, and an *implementability* dimension h, which captures whether the norm can be administered, enforced, and sustained in practice so that its promised consequences can be realized. Citizen i's information set $\mathcal{F}_{i,c,t}$ induces posterior beliefs about both dimensions. Let

$$M_{i,c,t} \equiv \mathbb{E}[m \mid \mathcal{F}_{i,c,t}], \qquad H_{i,c,t} \equiv \mathbb{E}[h \mid \mathcal{F}_{i,c,t}]$$

denote citizen i's posterior assessments of merit and implementability. I assume that endorsement depends on a latent index

$$\Lambda_{i,c,t} = \ell(M_{i,c,t}, H_{i,c,t}),$$

where $\ell_M > 0$, $\ell_H > 0$, and $\ell_{MH} > 0$. This index represents the posterior value of institutionalizing the norm. Citizen i endorses the norm when this value exceeds an idiosyncratic endorsement threshold. Equivalently, the observed endorsement propensity satisfies

$$s_{i,c,t} = \mathbb{E}[y_{i,c,t} \mid \mathcal{F}_{i,c,t}] = u(\Lambda_{i,c,t}),$$

where $u(\cdot)$ is increasing, concave, and twice differentiable, with the endorsement threshold normalized into u. The complementarity $\ell_{MH} > 0$ captures the bundle-based nature of endorsement: learning that a norm is desirable has a larger effect when citizens also believe it can be implemented, and conversely. Lemma 1 below shows that, under the signal structure of Assumption 4, posterior beliefs depend on exposure only through the cumulative precisions $Q_{k,i,c,t}$; the analysis then works with a tractable specialization of the index ℓ built directly on these sufficient statistics.

Priors are Gaussian and independent:

$$m \sim \mathcal{N}(\mu_m, \sigma_m^2), \qquad h \sim \mathcal{N}(\mu_h, \sigma_h^2).$$

The supply of public reasons is costly. Producing persuasive arguments requires effort because reasons must be formulated in ways that are intelligible, contestable, and informative about either merit or implementability. Merit arguments and implementability arguments draw on different knowledge bases. Merit arguments concern the normative and material desirability of the candidate norm. Implementability arguments concern the institutional *how*: whether the norm can be translated into rule-bound, administratively feasible, enforceable, and sustainable public action. A potential source of asymmetry between producers is institutional role. Legal-rational bureaucracy in the sense of Weber (1922) attaches state agents' authority to office rather than person and makes it valid only within legally defined competences: state agents are accordingly required to justify public action by reference to rules, competences, procedures, and administrative means, within a given normative order. These formal obligations are common to legal-rational office. The role conception built upon them, however, is culturally constituted: where the surrounding culture conceives bureaucratic roles around a strict separation between technocratic administration and normative authority, the office-holder's role is understood—by state agents themselves—as more fully Weberian, and state agents internalize this role conception and carry it into public discourse. Where this conception is absent, state agents participate in public justification on the same cost terms as their fellow citizens.

This motivates distinguishing two aggregate producer channels in the public sphere. The first is state agents, denoted $j = S$: public employees participating in the public sphere, whose role conception may shape the reasons they supply there. The second is peers, denoted $j = P$: other citizens participating in the same public sphere. Each channel summarizes the aggregate content supplied by the relevant set of actors. I do not model coordination among producers: unlike

in families, firms, or organizations, public-sphere discourse is not governed by a single authority coordinating the content or precision of the reasons supplied. Channels therefore choose the content and precision of their arguments independently, both within and across channels. In no case does institutional role confer superior standing: participants have equal justificatory standing, though their informational positions may differ with the surrounding culture. Assumption 1 formalizes the dimension-specific cost structure this generates.

Assumption 1 (Role-based persuasion costs) *Fix a public sphere (c,t) and suppress (c,t) subscripts in producer choices, except for the dependence on a predetermined, public-sphere–level index $W_c \in [0,\overline{W}]$, capturing the Weberian culture: the extent to which the surrounding culture expects a strict separation between technocratic administration and normative authority.*

There exist functions $\delta_h : [0,\overline{W}] \to [0,1)$, $\delta_m : [0,\overline{W}] \to [0,\infty)$, both weakly increasing with $\delta_h(0) = \delta_m(0) = 0$, such that for all $\pi \geq 0$,

$$K_S^{h\prime}(\pi) \;\leq\; \big(1-\delta_h(W_c)\big)\, K_P^{h\prime}(\pi), \qquad\qquad K_S^{m\prime}(\pi) \;\geq\; \big(1+\delta_m(W_c)\big)\, K_P^{m\prime}(\pi).$$

The assumption captures the idea that state agents are not intrinsically more persuasive: both cost asymmetries are role-based, dimension-specific, and culturally contingent, and at $W_c = 0$ the two channels face identical cost structures on both dimensions. The first inequality states a role-based cost advantage in producing implementability reasons that emerges and weakly deepens with Weberian culture. The second inequality captures a role-based *handicap* on the merit dimension that likewise emerges with W_c. The underlying logic is Weberian throughout: where Weberian culture runs deeper, public-sector roles are conceived more fully around legal-rational authority, and the state's claim to a public voice rests exclusively on legally defined competence rather than on any residual claim to normative authority. As W_c rises, this role conception not only cheapens the production of implementability reasons for state agents; it also makes merit assertions costlier: because state agents internalize an authority grounded in legal rationality rather than in normative standing, and carry this role conception with them into the public sphere, assertions about what is desirable fall outside the register of claims their role equips them to produce—in the vocabulary of the paper's epigraphs, the office's stance toward the enacted order is hermeneutic, while merit claims are prescriptive speech. The depth of this internalized role conception is itself shaped by the surrounding Weberian culture, which is why the cost asymmetry varies with W_c rather than being a fixed attribute of public employment.

Assumption 2 (Legitimacy has social value, imperfectly internalized) *Let*

$$S_{c,t} \equiv \int s_{i,c,t}\, di$$

denote aggregate endorsement of the candidate norm in public sphere (c,t). Higher endorsement raises a common payoff $V(S_{c,t})$, with $V'(S) > 0$; for the baseline analysis, I use the normalization $V(S) = BS$, $B > 0$. Each producer channel $j \in \{S,P\}$ internalizes a common fraction $\beta(R_c) \in [0,1]$ of this payoff, where $R_c \in [0,\overline{R}]$ is a predetermined, public-sphere–level index of the extent to which legitimate rules are understood as binding by virtue of collective self-legislation rather than as externally imposed commands, and $\beta(\cdot)$ is weakly increasing and differentiable on $[0,\overline{R}]$.

Public endorsement is valuable because it can improve institutional performance through higher voluntary compliance and lower enforcement burdens, and because individuals may attach non-instrumental value to living under rules they recognize as publicly justified. Each channel is modeled as a representative producer of public reasons: its precision choices affect the information received by

a positive mass of citizens through reach, and therefore aggregate endorsement $S_{c,t}$. The internalized fraction $\beta(R_c)$ generates a common wedge whenever $\beta(R_c) < 1$; it is a reduced-form measure of producers' internalization of instrumental and non-instrumental legitimacy payoffs, given whatever extrinsic or intrinsic motives make the supply of public reasons individually worthwhile. The model therefore abstracts from within-channel collective-action problems; unlike Assumption 1, $\beta(R_c)$ enters symmetrically across channels. Cultural environments can shift $\beta(R_c)$, as formalized in Corollary 2 below.

For each channel j, individual exposure within (c,t) decomposes as

$$E^j_{i,c,t} = E^j_{c,t} + \eta^j_{i,c,t}, \qquad \mathbb{E}[\eta^j_{i,c,t} \mid c,t] = 0,$$

where $E^j_{c,t}$ is aggregate reach in public sphere (c,t) for channel j, and $\eta^j_{i,c,t}$ is idiosyncratic reach. I assume that $\eta^j_{i,c,t}$ has support such that $E^j_{i,c,t} \geq 0$ almost surely. Citizens condition on their realized exposure.

Assumption 3 (Within-period predetermined reach) *Within (c,t), channel reach $E^j_{c,t}$, for $j \in \{S,P\}$, is determined by many micro-broadcasts and information-environment constraints, such as communication technology and media-market structure. Any single citizen or individual sender is negligible, so aggregate reach $E^j_{c,t}$ is taken as predetermined within the period.*

Channel reach $E^j_{c,t}$ is therefore predetermined within the period. By contrast, each channel's precision choice is non-atomistic: changing π^k_j shifts the informativeness of the signals received by the positive mass of citizens exposed through reach, and therefore affects aggregate endorsement.

Assumption 4 (Exposure–signal aggregation) *For each channel $j \in \{S,P\}$ and dimension $k \in \{m,h\}$, citizen i observes an aggregated signal*

$$\tilde{x}^k_{j,i,c,t} = \begin{cases} m + \tilde{\varepsilon}^m_{j,i,c,t}, & k = m, \\ h + \tilde{\varepsilon}^h_{j,i,c,t}, & k = h. \end{cases}$$

If $\pi^k_j E^j_{i,c,t} > 0$,

$$\tilde{\varepsilon}^k_{j,i,c,t} \sim \mathcal{N}\Big(0,\ (\pi^k_j E^j_{i,c,t})^{-1}\Big).$$

If $\pi^k_j E^j_{i,c,t} = 0$, no informative signal on (j,k) is observed; equivalently, the posterior contribution of that channel-dimension pair is zero. Signals are independent across j and k conditional on (m,h).

Citizen i's information set in (c,t) is

$$\mathcal{F}_{i,c,t} \;=\; \sigma\Big(E^S_{i,c,t},\, E^P_{i,c,t},\, \tilde{x}^m_{S,i,c,t},\, \tilde{x}^h_{S,i,c,t},\, \tilde{x}^m_{P,i,c,t},\, \tilde{x}^h_{P,i,c,t}\Big),$$

where inactive channel-dimension signals are treated as null observations.

Microfoundations. The mean-zero signal structure reflects the public and contestable nature of the arguments. Public-sphere claims are citable, answerable, and open to counter-argument; with many micro-senders and heterogeneous views, aggregation washes out idiosyncratic distortions, leaving no systematic drift in either dimension[5]. The precision formula can be microfounded in

[5]The mean-zero structure marks the model's domain boundary with respect to propaganda. Propaganda differs from persuasion here in two respects: it is produced by a coordinated sender whose distortion does not wash out in aggregation, and it operates by shifting posterior means or suppressing counter-argument rather than by raising the precision of contestable signals. With rational receivers who understand the bias technology, known bias is de-biased and residual uncertainty about it only lowers effective precision, so within this framework propaganda enters as precision destruction rather than as high-powered persuasion; genuine manipulation of posterior means requires naïve receivers, outside the model.

two equivalent ways. First, citizen i receives $E^j_{i,c,t}$ independent repetitions of a basic Gaussian signal from channel j on dimension k, and averaging yields cumulative precision $\pi^k_j E^j_{i,c,t}$. Second, information arrives as a continuous-time flow with rate π^k_j over an exposure window of length $E^j_{i,c,t}$, again yielding cumulative precision $\pi^k_j E^j_{i,c,t}$. Thus, exposure and content precision are complements in the production of posterior precision. Individual endorsement depends on realized exposure $E^j_{i,c,t}$; producers' payoffs depend on its aggregate through $S_{c,t}$, and I characterize this mapping in Section 2.2.

2.2 Analysis

2.2.1 Static analysis

Posterior with exposure-scaled precision. Define prior precisions $\tau_m \equiv \sigma_m^{-2}$, $\tau_h \equiv \sigma_h^{-2}$, and the total effective precision on dimension $k \in \{m, h\}$,

$$Q_{k,i,c,t} \equiv \sum_{j \in \{S,P\}} \pi^k_j \, E^j_{i,c,t}.$$

Reach scales informativeness linearly: each additional unit of exposure to channel j adds π^k_j units of precision on dimension k.

Lemma 1
Under the Gaussian priors and Assumption 4, posteriors decouple by dimension and are Gaussian with precision $\tau_k + Q_{k,i,c,t}$. In particular,

$$m \mid \{\tilde{x}^m_{j,i,c,t}\}_j \sim \mathcal{N}\left(\frac{\tau_m \mu_m + \sum_j (\pi^m_j E^j_{i,c,t})\, \tilde{x}^m_{j,i,c,t}}{\tau_m + \sum_j \pi^m_j E^j_{i,c,t}},\ (\tau_m + \sum_j \pi^m_j E^j_{i,c,t})^{-1} \right),$$

$$h \mid \{\tilde{x}^h_{j,i,c,t}\}_j \sim \mathcal{N}\left(\frac{\tau_h \mu_h + \sum_j (\pi^h_j E^j_{i,c,t})\, \tilde{x}^h_{j,i,c,t}}{\tau_h + \sum_j \pi^h_j E^j_{i,c,t}},\ (\tau_h + \sum_j \pi^h_j E^j_{i,c,t})^{-1} \right),$$

where channel-dimension pairs with $\pi^k_j E^j_{i,c,t} = 0$ are omitted from the relevant sums.

Proof 1 (Sketch) *Conditional on each dimension k, signals are independent Gaussians with precisions $\pi^k_j E^j_{i,c,t}$. Gaussian conjugacy implies that precisions add and that the posterior mean is a precision-weighted average of the prior mean and observed signals. Independence across dimensions implies that the posteriors for m and h decouple.*

From posteriors to the bundle index. Lemma 1 establishes that cumulative precision $Q_{k,i,c,t}$ is a sufficient statistic for what citizen i has learned on dimension k: posteriors tighten with Q_k regardless of how exposure is distributed across channels. For the remainder of the analysis, I adopt a tractable specialization of the endorsement index built directly on these sufficient statistics: the probability that channel j delivers a complete justificatory bundle is $\Lambda_j = \left(1 - e^{-Q^m_j}\right)\left(1 - e^{-Q^h_j}\right)$, where $1 - e^{-Q}$ is the probability that cumulative informativeness Q has delivered at least one decisive argument on the relevant dimension—the Poisson-arrival reading of the same exposure-scaled precision—and the product structure implements the complementarity $\ell_{MH} > 0$ in its starkest form: a channel persuades only when both the normative *why* and the institutional *how* arrive. This specialization preserves the two features of Lemma 1 that drive the results: informativeness is increasing and concave in exposure, and merit and implementability learning enter jointly rather than additively.

Benchmark: within-channel bundling (no cross-channel fit). Fixing (i, c, t) and suppressing subscripts, write $Q_j^k \equiv \pi_j^k E^j \geq 0$ for the channel-dimension components of the cumulative precisions in Lemma 1. In the benchmark, endorsement requires a matched pair of reasons (m, h) assembled *within the same channel*: the probability that channel j delivers a complete bundle is the index $\Lambda_j(Q_j^m, Q_j^h) = (1 - e^{-Q_j^m})(1 - e^{-Q_j^h})$ introduced above. Let $\Lambda \equiv \Lambda_S + \Lambda_P$ denote the total complete-bundle index, and let endorsement be

$$s(E^S, E^P) = u(\Lambda),$$

where u satisfies the maintained properties of Section 2.1, with $u : \mathbb{R}_+ \to [0, 1)$ and the normalization $u'(0) = 1$; a canonical case is $u(\Lambda) = 1 - e^{-\Lambda}$.

Proposition 1 (Within-channel bundling generates substitutability across channels) *Under the benchmark above, endorsement is weakly increasing in each exposure and submodular:*

$$\frac{\partial s}{\partial E^S} \geq 0, \qquad \frac{\partial s}{\partial E^P} \geq 0, \qquad \frac{\partial^2 s}{\partial E^S \, \partial E^P} \leq 0.$$

Proof 2 (Sketch) *Since $\Lambda = \Lambda_S(E^S) + \Lambda_P(E^P)$, we have $\partial s / \partial E^j = u'(\Lambda)\, \Lambda_j'(E^j) \geq 0$. Moreover,*

$$\frac{\partial^2 s}{\partial E^S \, \partial E^P} = u''(\Lambda)\, \Lambda_S'(E^S)\, \Lambda_P'(E^P) \;\leq\; 0,$$

because $u'' \leq 0$ and $\Lambda_j' \geq 0$, with $\Lambda_j' > 0$ whenever $\pi_j^m > 0, \pi_j^h > 0$, and $E^j > 0$.

Interpretation. Monotonicity ($u' > 0$) captures that additional reach increases the arrival of complete argument bundles and hence endorsement. Concavity ($u'' \leq 0$) reflects diminishing returns from saturation, limited attention, and audience overlap. In the benchmark, $\Lambda = \Lambda_S(E^S) + \Lambda_P(E^P)$ is additively separable across channels, so the cross–partial is negative only through the curvature of u: $s_{SP} = u''(\Lambda)\Lambda_S'(E^S)\Lambda_P'(E^P) \leq 0$. This is therefore a prediction about how complete persuasive bundles combine at the endorsement stage, not about exposure mechanics per se.

Producer choice in the benchmark. In the benchmark, complete argument bundles are assembled within channels. I therefore model precision choices $\{\pi_j^m, \pi_j^h\}_{j \in \{S,P\}}$ as determined by within-channel optimization. The purpose of this step is to show how reach, internalized legitimacy payoffs, and role-based marginal costs map into equilibrium precision choices, and hence into the completeness of each channel's argumentative bundle. In the subsequent specialization extension, I first take these content mixes as given in order to isolate how cross-channel fit changes the sign of the exposure cross-partial; Remark 2.2.1 then lets the peer channel's choices respond.

Lemma 2 (Benchmark producer optimality conditions)
Fix a public sphere (c, t) and suppress (c, t) subscripts, except for the dependence on the predetermined indices W_c and R_c. Channel $j \in \{S, P\}$ has aggregate reach $E^j \geq 0$ and chooses dimension-specific precisions $\pi_j^m, \pi_j^h \geq 0$, facing the marginal-cost schedules $K_j^{k\prime}(\cdot)$ of Assumption 1, which for $j = S$ depend on W_c; I write $K_j^k(\pi_j^k; W_c)$ to make this dependence explicit, with $K_P^k(\pi; W_c) = K_P^k(\pi)$ for all W_c. Under within-channel bundling, the complete-bundle index supplied by channel j is

$$\Lambda_j(\pi_j^m, \pi_j^h; E^j) \;=\; (1 - e^{-\pi_j^m E^j})(1 - e^{-\pi_j^h E^j}).$$

Let

$$\Lambda = \Lambda_j + \Lambda_{-j}$$

denote the total complete-bundle index, taking the other channel's precision choices as given. Channel j chooses (π_j^m, π_j^h) to maximize its internalized legitimacy payoff net of precision costs:

$$\max_{\pi_j^m, \pi_j^h \geq 0} \; \beta(R_c)\, B\, u(\Lambda) \; - \sum_{k \in \{m,h\}} K_j^k(\pi_j^k; W_c), \qquad j \in \{S, P\}.$$

For an interior solution, the first-order condition for dimension $k \in \{m, h\}$ is

$$\beta(R_c)\, B\, u'(\Lambda)\, \frac{\partial \Lambda_j}{\partial \pi_j^k} \;=\; K_j^{k\prime}(\pi_j^k; W_c).$$

Equivalently,

$$\beta(R_c)\, B\, u'(\Lambda)\, E^j e^{-\pi_j^m E^j} \left(1 - e^{-\pi_j^h E^j}\right) \;=\; K_j^{m\prime}(\pi_j^m; W_c),$$

and

$$\beta(R_c)\, B\, u'(\Lambda)\, E^j e^{-\pi_j^h E^j} \left(1 - e^{-\pi_j^m E^j}\right) \;=\; K_j^{h\prime}(\pi_j^h; W_c).$$

Interpretation. The optimality conditions show that precision investment equates the marginal legitimacy benefit of making a channel's bundle more complete with the marginal cost of producing precision. The marginal return to precision is increasing in the internalized legitimacy payoff $\beta(R_c)B$ and depends on reach E^j, which determines how many citizens receive the channel's arguments and how quickly the channel approaches saturation. The conditions also display within-channel complementarity: the marginal return to merit precision is higher when implementability precision is higher, and conversely. Lemma 2 shows how Assumption 1 shifts equilibrium precision choices. At $W_c = 0$, the two channels face identical cost structures and choose identical content mixes for given reach. As Weberian culture deepens, state agents' marginal cost of implementability precision falls while their marginal cost of merit precision rises, and their precision choice tilts accordingly toward the institutional *how*—and away from the normative *why*. I use this result as the supply-side foundation for the specialization extension below. The point is not that state agents are intrinsically more persuasive, but that the internalized legal-rational role conception, where the surrounding culture sustains it, reshapes which reasons are cheap for them to produce.

Corollary 1 (Weberian culture and the persuasion architecture)
Let $(\pi_S^m(W_c), \pi_S^h(W_c))$ denote the state-agent channel's equilibrium precision choices characterized in Lemma 2, under Assumption 1, holding the peer channel's precision choices (π_P^m, π_P^h) and reach (E^S, E^P) fixed. Then:

(i) $\pi_S^m(W_c)$ is weakly decreasing in W_c, and $\pi_S^h(W_c)$ is weakly increasing in W_c;

(ii) the state channel's within-channel bundle $\Lambda_S(W_c) = \left(1 - e^{-\pi_S^m(W_c)E^S}\right)\left(1 - e^{-\pi_S^h(W_c)E^S}\right)$ is weakly decreasing in W_c, and $\Lambda_S(W_c) \to 0$ as $W_c \to \overline{W}$ whenever $\pi_S^m(\overline{W}) = 0$;

(iii) the quantity $\kappa(W_c) \equiv \pi_S^m(W_c)\, \pi_P^h + \pi_S^h(W_c)\, \pi_P^m$ satisfies $\kappa(W_c) \to \pi_S^h(\overline{W})\, \pi_P^m > 0$ as $W_c \to \overline{W}$, whenever $\pi_S^h(\overline{W}) > 0$ and $\pi_P^m > 0$.

Proof 3 (Proof sketch) *Part (i) follows from monotone comparative statics on the first-order conditions of Lemma 2. Raising W_c shifts the marginal-cost bound $(1 + \delta_m(W_c))K_P^{m\prime}(\pi)$ up and $(1 - \delta_h(W_c))K_P^{h\prime}(\pi)$ down, uniformly in π; since $\partial\Lambda_S/\partial\pi_S^m$ and $\partial\Lambda_S/\partial\pi_S^h$ are each decreasing in the corresponding argument (from the bundle technology's own curvature), while the marginal-cost schedules are weakly increasing (from the convexity of K_j^k), the standard single-crossing argument applies term by term, giving $d\pi_S^m/dW_c \leq 0$ and $d\pi_S^h/dW_c \geq 0$. Part (ii) follows by direct substitution into Λ_S. Part (iii) is immediate from the definition of κ once $\pi_S^m \to 0$, since the first term vanishes and the second remains bounded away from zero by assumption.*

Interpretation. Corollary 1 formalizes the double-edged role of W_c taken up in the normative discussion of Section 2.2.3. A Weberian culture that expects a stricter separation between technocratic administration and normative authority does not merely sharpen the state's comparative advantage in implementability: it simultaneously erodes the state-agent channel's capacity to complete a within-channel bundle on its own, since a complete bundle requires both factors of Λ_S to be bounded away from zero. Where, then, does the state channel's persuasive contribution go? The specialization extension below supplies the answer: Proposition 2 identifies κ—the quantity tracked in part (iii)—as the parameter governing whether complete bundles can instead be assembled *across* channels, and part (iii) shows that Weberian culture pushes precisely this parameter up. As W_c rises, the state-agent channel's own contribution to endorsement therefore increasingly runs through the cross-channel term $\Lambda_\times$ introduced below rather than through Λ_S. The corollary yields three jointly testable, dimension-specific predictions that a generic concave persuasion technology with no role for W_c does not generate: the state-agent channel's own exposure effect should weaken as W_c rises; since Assumption 1 leaves the peer channel's cost structure independent of W_c, the peer channel's own exposure effect should be unaffected by W_c through the cost channel; and the peer×state-agent exposure interaction should move from substitutes toward complements as W_c rises. Throughout, the corollary holds the peer channel's precision choices fixed, isolating the supply-side shock to the state-agent channel; Remark 2.2.1 below relaxes this and shows that, under the specialization extension, the peer channel's equilibrium response reinforces rather than offsets the shift toward cross-channel fit.

Corollary 2 (Grammar of legitimacy and within-channel completeness) *Let $\left(\pi_j^m(R_c), \pi_j^h(R_c)\right)_{j\in\{S,P\}}$ denote the equilibrium precision choices characterized in Lemma 2 under the internalization parameter $\beta(R_c)$ of Assumption 2, holding reach (E^S, E^P) and Assumption 1's cost primitives fixed. Then:*

(i) $\pi_j^m(R_c)$ and $\pi_j^h(R_c)$ are weakly increasing in R_c for both $j \in \{S, P\}$;

(ii) consequently, both channels' within-channel bundles, $\Lambda_S(R_c)$ and $\Lambda_P(R_c)$, are weakly increasing in R_c.

Proof 4 (Proof sketch) *Part (i) follows directly from the first-order conditions of Lemma 2: raising $\beta(R_c)$ scales the marginal legitimacy benefit $\beta(R_c)\, B\, u'(\Lambda)$ appearing on the left-hand side of every precision choice, for both channels and both dimensions, since neither Assumption 1 nor the reach terms E^j depend on R_c; standard comparative statics on a uniformly rescaled marginal benefit against a fixed, convex marginal cost then give a weakly higher optimal precision on every margin. Part (ii) follows by direct substitution into Λ_S and Λ_P.*

Interpretation. Because raising $\beta(R_c)$ scales precision without altering the additively separable bundle structure $\Lambda = \Lambda_S + \Lambda_P$, Proposition 1 continues to apply at every value of R_c: the peer×state-agent interaction remains negative throughout, regardless of how strongly cultural exposure raises both channels' internalization. What is not pinned down by parts (i)–(ii) alone is whether the interaction becomes more negative as R_c rises. Part (ii) is a statement about the *level* of each channel's bundle at a given exposure; the interaction instead depends on the *slope* $\Lambda_j'(E^j)$, the marginal return to an additional unit of exposure. Since $\Lambda_S(E^S) = (1 - e^{-\pi_S^m E^S})(1 - e^{-\pi_S^h E^S})$ is S-shaped in exposure, raising precision shifts this curve upward at every point without determining whether a given public sphere sits on its steepening or its flattening segment: public spheres whose exposure was previously insufficient to complete the bundle see the marginal return to exposure rise as precision increases, while public spheres already close to a complete bundle see it fall. Whether the interaction deepens with R_c is therefore an empirical question.

Extension: cross-channel fit under specialization. When channels specialize, citizens may assemble a matched pair across channels: a merit reason from one channel and an implementability reason from the other. I capture this possibility by adding a cross-channel bundle index

$$\Lambda_\times \;\equiv\; \left(1-e^{-Q_S^m}\right)\left(1-e^{-Q_P^h}\right) \;+\; \left(1-e^{-Q_S^h}\right)\left(1-e^{-Q_P^m}\right),$$

and setting $s(E^S,E^P)=u(\Lambda_S+\Lambda_P+\Lambda_\times)$. This extension isolates cross-channel fit: exposure to one channel can raise the marginal persuasive value of exposure to the other when each channel supplies a different component of the justificatory bundle.

Proposition 2 (Cross-channel fit governs complementarity)
Fix (i,c,t) and suppress subscripts. Let $f(x)\equiv 1-e^{-x}$, and let $\Lambda\equiv\Lambda_S+\Lambda_P+\Lambda_\times$ with the indices defined above; endorsement is $s(E^S,E^P)=u(\Lambda)$ under the maintained properties of Section 2.1. Define

$$\kappa \;\equiv\; \pi_S^m\pi_P^h+\pi_S^h\pi_P^m.$$

Then:

(i) (Global necessity.) *If $\kappa=0$, then $\Lambda_\times\equiv 0$ for all $(E^S,E^P)\in\mathbb{R}^2_+$, and*

$$\frac{\partial^2 s}{\partial E^S\,\partial E^P} \;\le\; 0 \qquad \text{for all } (E^S,E^P)\in\mathbb{R}^2_+;$$

equivalently, $\partial^2 s/\partial E^S\partial E^P>0$ at any exposure level implies $\kappa>0$.

(ii) (Local sufficiency.) *If $\kappa>0$, then*

$$\left.\frac{\partial^2 s}{\partial E^S\,\partial E^P}\right|_{E^S=E^P=0} = u'(0)\,\kappa \;>\; 0,$$

and by continuity there exists $\varepsilon>0$ such that $\partial^2 s/\partial E^S\partial E^P>0$ for all $(E^S,E^P)\in[0,\varepsilon]^2$.

Proof 5 (Sketch) *Differentiating $s=u(\Lambda_S+\Lambda_P+\Lambda_\times)$ twice by the chain rule, at any exposure level,*

$$s_{SP} \;=\; u''(\Lambda)\left[\Lambda_S'+\Lambda_{\times,S}\right]\left[\Lambda_P'+\Lambda_{\times,P}\right] \;+\; u'(\Lambda)\,\Lambda_{\times,SP},$$

where $\Lambda_{\times,S}\equiv\partial\Lambda_\times/\partial E^S$ and $\Lambda_{\times,SP}\equiv\partial^2\Lambda_\times/\partial E^S\partial E^P$. Since $f'(x)=e^{-x}$,

$$\Lambda_{\times,SP} \;=\; \pi_S^m\pi_P^h\,e^{-Q_S^m-Q_P^h} \;+\; \pi_S^h\pi_P^m\,e^{-Q_S^h-Q_P^m} \;\ge\; 0,$$

with equality for all (E^S,E^P) if and only if $\kappa=0$, in which case $\Lambda_\times\equiv 0$ identically. Because $u''\le 0$ and the two bracketed marginal effects are non-negative, the first term of s_{SP} is non-positive everywhere; when $\kappa=0$ the second term vanishes identically as well, giving part (i). Part (ii) follows by evaluating at $E^S=E^P=0$, where $Q_j^k=0$, $f(0)=0$, $f'(0)=1$, so Λ_S and Λ_P contribute no cross-partial and $\Lambda_{\times,SP}|_0=\kappa$; by the chain rule $s_{SP}|_0=u'(0)\kappa$, and continuity of s_{SP} yields positivity on a neighborhood when $\kappa>0$.

Interpretation. The benchmark assumes that citizens obtain a complete case for endorsement—a persuasive *why* (m) and a persuasive *how* (h)—from within the same channel. Specialization breaks this: if one channel is relatively better at supplying implementability and the other at supplying normative justification, complete bundles may be assembled *across* channels. The term $\Lambda_\times$ captures precisely this possibility: $f(Q_S^h)f(Q_P^m)$ is the probability of encountering an implementability

argument via the state-agent channel *and* a merit argument via the peer channel, while $f(Q_S^m)f(Q_P^h)$ captures the opposite pairing. The parameter $\kappa \equiv \pi_S^m \pi_P^h + \pi_S^h \pi_P^m$ summarizes the scope for this cross-channel assembly.

Part (i) of the proposition gives the result its empirical force: a positive peer $\times$ state-agent interaction, observed at any level of exposure, is impossible under pure within-channel bundling, since $\kappa = 0$ forces the interaction to be non-positive everywhere. Observing a positive interaction is therefore diagnostic of cross-channel fit regardless of where in the exposure distribution it is estimated. The converse is weaker, and deliberately so: part (ii) guarantees only that $\kappa > 0$ generates complementarity *locally*, near the origin. Away from the origin, saturation can drive the first, u''-weighted term of s_{SP} to dominate the second, so that a specialized architecture with $\kappa > 0$ may still exhibit a negative interaction at high exposure. The sign of the interaction therefore reveals the architecture asymmetrically: a positive interaction is unconditional evidence of cross-channel fit, while a negative interaction is the generic case under either architecture and does not, on its own, establish within-channel bundling.

Specialization is therefore a force multiplier, not a universal obstacle to legitimacy formation. It raises precision along each channel's comparative-advantage dimension: peer exposure may strongly increase perceived merit, while state-agent exposure may strongly increase perceived implementability. But because endorsement is bundle-based, one-dimensional precision is not sufficient on its own. Its effect depends on the structure of exposure. In integrated public spheres, where citizens encounter both channels, specialization can accelerate legitimacy formation because audiences receive the full justificatory bundle across channels. In fragmented public spheres, by contrast, specialization can impede consensus because citizens may receive only one amplified argument without the complementary reason needed for endorsement.

Together, the benchmark and the specialization extension generate the paper's central empirical asymmetry. When channels are internally complete ($\kappa = 0$), peer and state-agent exposures substitute for one another at every exposure level, because each channel can independently deliver a persuasive (m, h) bundle. When channels specialize ($\kappa > 0$), exposure to one channel can raise the value of exposure to the other, at least at low exposure, because citizens assemble the justificatory bundle across channels. A positive exposure cross-partial, wherever it is found, is therefore unconditional evidence of cross-channel fit; a negative cross-partial is consistent with, but does not by itself establish, within-channel bundling.

Remark (Equilibrium two-sided specialization) The comparative statics of Corollary 1 hold the peer channel's precision choices fixed. Under the specialization extension, they respond. In the extended technology, the state-agent channel's precisions enter the peer channel's marginal products through $\Lambda_\times$: the return to peer merit precision rises with π_S^h, while the return to peer implementability precision rises with π_S^m. Holding the common marginal-value term $u'(\Lambda)$ fixed, an increase in W_c—which raises π_S^h and lowers π_S^m by Corollary 1(i)—therefore raises the peer channel's optimal merit precision and lowers its optimal implementability precision: peers specialize in the normative *why* as the state specializes in the institutional *how*. Two implications follow. First, the cross-channel fit parameter κ rises by more than the direct effect alone, since $\pi_S^h \pi_P^m$ grows on both factors. Second, the peer channel's within-channel bundle Λ_P falls whenever the decline in π_P^h dominates, so both channels' own-exposure effects can weaken with W_c even as their interaction turns positive. With a strictly concave endorsement link, all four first-order conditions share the multiplier $u'(\Lambda)$, whose response to W_c is ambiguous in sign; the dimension-specific substitution effects above are the first-order forces, and the signs stated here obtain whenever neither the common-multiplier feedback nor the peer channel's own within-channel complementarity dominates

them—for instance, when u is locally close to linear and the shock transmits chiefly through the cross-channel pairing terms. Appendix A.1 states the formal conditions.

2.2.2 Dynamics: Feedback and cultural change

The previous results characterize within-period persuasion. I now allow endorsement to feed back into future reach through public-sphere propagation, generating norm dynamics.

Let aggregate endorsement at date t be $S_t \in [0,1]$. Public-sphere propagation implies that reach in the next period is increasing in current endorsement:

$$E^j_{t+1} \;=\; E^j(S_t), \qquad j \in \{S,P\}.$$

Given reach (E^S_{t+1}, E^P_{t+1}), the total bundle index is $\Lambda_{t+1} \equiv \Lambda(E^S_{t+1}, E^P_{t+1})$, and aggregate endorsement evolves according to

$$S_{t+1} \;=\; u(\Lambda_{t+1}) \;\equiv\; \Phi(S_t), \qquad \Phi(S) \;=\; u\Big(\Lambda\big(E^S(S), E^P(S)\big)\Big).$$

This dynamic formulation differs from preference-falsification accounts of rapid norm change. In such accounts, individuals change their public stance when the perceived social acceptability of a position shifts, generating threshold cascades ((Kuran (1991), Kuran (1995), Bursztyn et al. (2020))). A reduced-form representation is

$$y_{i,t} = \mathbf{1}\{\, p_i + \mu_t - c_i(S_{t-1}) + \varepsilon_{i,t} \geq 0 \,\}, \qquad S_t = \mathbb{E}[y_{i,t}],$$

where p_i is the private stance, $c_i(\cdot)$ decreases in perceived support so that expressing support becomes cheaper as perceived support rises, and μ_t is an acceptability shock. In this class of models, rapid change operates through the social cost of public expression. By contrast, the mechanism here operates through persuasion: reach determines which arguments citizens encounter, producer roles shape the precision and content of those arguments, and bundling determines whether endorsement changes smoothly or accelerates through cross-channel completion.

Assumption 5 (Exposure maps) *For $j \in \{S,P\}$, $E^j : [0,1] \to [0,\infty)$ is twice differentiable and weakly increasing, with ($E^{j\,\prime} \geq 0$) and bounded derivatives. Exposure may saturate, so $E^{j\,\prime}$ may decline at high endorsement levels.*

Lemma 3 (Curvature decomposition) *For $\Phi(S) = u(\Lambda(S))$,*

$$\Phi'(S) = u'(\Lambda)\,\Lambda'(S), \qquad \Phi''(S) = u''(\Lambda)\,[\Lambda'(S)]^2 \;+\; u'(\Lambda)\,\Lambda''(S).$$

In the specialization extension, $\Lambda''(S)$ contains the strictly positive cross-fit term

$$2\,E^{S\,\prime}(S)\,E^{P\,\prime}(S)\left[\pi^m_S \pi^h_P\, e^{-Q^m_S(S)-Q^h_P(S)} \;+\; \pi^h_S \pi^m_P\, e^{-Q^h_S(S)-Q^m_P(S)}\right],$$

which is absent in the benchmark without cross-channel fit.

Proof 6 (Sketch) *The first display follows from the chain rule. For the second, write $\Lambda_\times(S) = f(Q^m_S(S))f(Q^h_P(S)) + f(Q^h_S(S))f(Q^m_P(S))$ with $f(x) = 1 - e^{-x}$. Differentiating twice yields, for each off-diagonal pair, a cross term $2\,f'(Q^\ell_S)\,f'(Q^{\ell'}_P)\,Q^{\ell\prime}_S(S)\,Q^{\ell'\prime}_P(S)$, with $f'(x) = e^{-x}$ and $Q^{k\prime}_j(S) = \pi^k_j E^{j\,\prime}(S)$. Summing over the two off-diagonal pairs gives the stated expression.*

Interpretation. The update map has two sources of curvature. The first is the concavity of u, which captures saturation in endorsement and pushes toward gradual adjustment. The second is the curvature of the bundle index $\Lambda(S)$. Cross-channel fit adds a positive curvature term because exposure to one channel raises the marginal value of exposure to the other: when one channel supplies the normative *why* and the other supplies the institutional *how*, public-sphere propagation can accelerate endorsement. Specialization can therefore generate nonlinear cultural change. At low or intermediate endorsement levels, cross-channel completion may make the update map locally convex; at high endorsement levels, saturation in exposure and endorsement can restore concavity. This provides the mechanism through which persuasion can generate both smooth cultural change, when bundling is internal to channels or feedback is weak, and rapid cultural shifts, when cross-channel completion interacts with sufficiently strong feedback.

The two cultural primitives interact in these dynamics. Because $\beta(R_c)$ scales all four first-order conditions, a grammar of legitimacy that makes public justification more consequential intensifies whatever content pattern the Weberian margin has produced: under deep specialization, higher R_c means more implementability from the state and more merit from peers, and the cross-fit curvature term of Lemma 3 scales accordingly. A public sphere high on both margins is therefore the model's highest-variance configuration, not its most favorable one: where reach is integrated and merit discourse contested, it produces the fastest legitimate norm formation; where reach is fragmented and merit discourse weakly disciplined, it produces the fastest amplification of poorly justified claims. The two historical illustrations below occupy opposite regions of this configuration space; they are not tests of the model, but mappings from its primitives to well-known cases.

Interwar Germany: Weberian culture under fragmented public sphere. Interwar Germany can be interpreted as an environment in which receptiveness to particularist and exclusionary narratives was already nontrivial, consistent with the long-run intellectual currents emphasized by Kohn (1950) and Mosse (1964), together with a conception of the bureaucratic role in strictly legal-rational terms (Arendt (1963), Heldring (2026)). In the model, the latter corresponds to high W_c: state agents' public discourse specializes in implementability, raising π_S^h relative to π_S^m. The former supplies the weakly disciplined merit discourse on which, per Section 2.2.3, such specialization operates. When this institutional *how* is paired with moral content supplied through peer discourse, cross-channel fit increases. Public-sphere reach, moreover, was fragmented rather than broadly integrative: despite dense associational life, civic engagement occurred largely within segmented subpublics, limiting the common arenas in which complete justificatory bundles could form (Berman (1997)). Mass persuasion technologies, including radio propaganda (Adena et al. (2015)), can be interpreted as raising effective reach in this fragmented environment. The model then suggests a mechanism through which preexisting particularist merit claims, bureaucratic implementability discourse, segmented public-sphere structure, and a late increase in reach can jointly generate rapid endorsement shifts.

The French Revolution: grammar of legitimacy under integrated public sphere. A contrasting illustration is late–1780s France. Peer discourse already bundled the normative *why* and the institutional *how*: claims about general interest, equality, and representation circulated together with concrete institutional proposals. Contributions such as Condorcet (1785) and Sieyès (1789) illustrate this coupling. In the model, this corresponds to substantial within-channel bundling in the peer channel, with high π_P^m and π_P^h, so that peer-delivered complete bundles can be generated without relying primarily on cross-channel matching. The pamphlet and newspaper boom, together with salons, clubs, and other forms of associational life, increased peer reach E^P by expanding

the frequency and scale of public argument exchange (Habermas (1989), Darnton (1982), Popkin (1990)). As the Estates–General convened and parliamentary procedure emerged through the National and Constituent Assemblies, public endorsement became tied to formal political choice[6]. In the model, this is a shift in the grammar of legitimacy: as rules come to be understood as collectively self-legislated, $\beta(R_c)$ rises, because producing and relaying public reasons becomes more consequential for institutional outcomes. On this interpretation, revolutionary acceleration reflects broad increases in reach and already-strong within-peer bundling, reinforced by the internalization mechanism of Corollary 2.

2.2.3 Welfare: amplification risk and guardrails

This subsection is normative discussion: it draws welfare implications from the model's structure and is not tested empirically in this paper. The positive analysis highlights a welfare trade-off. Because endorsement is bundle-based—it requires both a persuasive *why* (m) and a persuasive *how* (h)—precision investments can improve the informational basis of legitimacy. But under strong role specialization, the state-agent channel supplies mainly implementability arguments while contributing little merit discourse. In that case, implementability can become a content-agnostic amplifier: it increases the institutional realizability of whichever merit claim prevails, rather than disciplining the merit claim itself. The model's two cultural primitives separate two corresponding wedges. The first is an internalization wedge, governed by $\beta(R_c)$, that determines the overall supply of public reasons. The second is a composition wedge, governed by W_c, that determines whether implementability precision completes a well-disciplined merit case or amplifies a weakly contested one.

Planner benchmark. Fix a public sphere (c, t) and suppress subscripts, except for the dependence on the predetermined indices W_c and R_c. In the decentralized benchmark, channel $j \in \{S, P\}$ chooses (π_j^m, π_j^h) to maximize

$$\max_{\pi_j^m, \pi_j^h \geq 0} \; \beta(R_c)\, B\, u(\Lambda) \; - \sum_{k \in \{m,h\}} K_j^k(\pi_j^k; W_c),$$

as in Lemma 2. A utilitarian planner instead internalizes the full legitimacy payoff and chooses all precision levels:

$$\max_{\{\pi_j^m, \pi_j^h\}_{j \in \{S,P\}}} \; B\, u(\Lambda) \; - \sum_{j \in \{S,P\}} \sum_{k \in \{m,h\}} K_j^k(\pi_j^k; W_c).$$

The decentralized problem therefore contains a standard under-internalization wedge: whenever $\beta(R_c) < 1$, the marginal private return to producing public reasons is below the marginal social return, uniformly across channels and dimensions. Because $\beta(\cdot)$ is increasing in R_c, cultural environments in which legitimate rules are understood as collectively self-legislated close part of this wedge on their own; policies or institutions that raise internalization further, or that reduce the cost of producing precise and contestable reasons, increase the supply of legitimacy-relevant information. Note that this wedge is symmetric: it scales the supply of merit and implementability reasons alike, and therefore affects the *volume* of public justification, not its composition.

[6]The *cahiers de doléances* were lists of complaints and reform proposals compiled across France in early 1789 by local assemblies of the three estates to instruct their deputies to the Estates–General.

Amplification risk. The composition wedge is different in kind, and the positive analysis shows it is not a hypothetical configuration. By Corollary 1, a deep Weberian culture tilts the state channel's equilibrium content mix toward

$$\pi_S^h \gg \pi_S^m \simeq 0,$$

and by Remark 2.2.1, the peer channel's equilibrium response reinforces the resulting specialization rather than offsetting it. State-agent discourse then supplies the institutional *how* without supplying much of the normative *why*. In the specialization extension, this creates cross-channel fit when paired with peer-side merit arguments, and Proposition 2 shows that a positive exposure interaction can arise only in this configuration.

The welfare reading of this configuration is a trade-off between production efficiency and assembly fragility. Specialization is a division of justificatory labor: it allocates each dimension to the channel that produces it cheaply—implementability to state agents, merit to peers—so that, for given total effort, more precision is supplied. But it simultaneously changes the assembly technology from robust to fragile. Under within-channel bundling, a single channel reaching a citizen can complete the case: Λ_S alone suffices. Under specialization, completion is multiplicative across channels: $\Lambda_\times$ requires both channels to reach the same citizen. The division of justificatory labor is therefore limited by the integration of the public sphere[7]: specialization improves legitimacy formation only where reach is sufficiently integrated for cross-channel assembly. It is, moreover, content-agnostic on the merit side: implementability arguments amplify whatever merit discourse the peer channel supplies. When peer-side merit discourse is well contested and precise, the complementarity improves legitimacy formation; when it is weakly disciplined, implementability precision instead scales up a one-sided or poorly justified merit claim. The risk is compounded by the equilibrium response in Remark 2.2.1: as peers withdraw from implementability argumentation, no channel remains that pairs the institutional *how* with an internally disciplined *why*, so the quality of the aggregate justificatory bundle rests entirely on the contestability of peer-side merit discourse.

The two margins therefore interact: raising internalization—the R_c margin—is unambiguously wedge-closing, but where the W_c margin has produced deep specialization, it also raises the stakes of the composition wedge, scaling whichever configuration—completion or amplification—the architecture delivers.

Guardrails. The amplification risk is greatest when three conditions coincide: (i) the Weberian culture is deep, so state-agent discourse is strongly specialized in implementability; (ii) reach is low or fragmented, so complete bundles are scarce and the local-complementarity region of Proposition 2 is the relevant one; and (iii) merit discourse is weakly contested. The corresponding guardrails map directly into the model's primitives. First, composition guardrails sustain π_S^m against extreme h-only specialization in the state-agent channel—for instance, institutional mandates or professional norms that require public authorities to state the normative grounds of proposed rules, not only their administrative feasibility. Second, merit-discipline guardrails raise the precision and contestability of peer-side merit discourse, through media quality, verification capacity, civic education, and deliberative infrastructure; under strong specialization these operate on the only channel still supplying the normative *why*. Third, public-sphere integration guardrails increase cross-cutting reach and common fora, reducing the low-exposure fragmentation under which cross-channel amplification is most powerful.

[7]The parallel is to Smith (1776): the division of labour is limited by the extent of the market. Here, the division of justificatory labor is limited by the integration of the public sphere.

The scale of the public sphere. The guardrails above take the boundary of the public sphere as given. A final normative question concerns that boundary itself. Some norms are intrinsically scale-dependent. Climate rules, migration regimes, and other forms of cross-border coordination require implementability arguments that refer to institutions beyond the local arena, so their justificatory bundles cannot be completed within a local public sphere. In the model's terms, a segmented public sphere structurally under-legitimates such norms: each local arena can supply the normative merit, but the institutional implementability has no arena in which to meet it. The results in Section 6.2 are consistent with this mechanism: among the most strongly persuaded respondents, endorsement of universalism is associated with support for supranational integration. The countervailing cost of an integrated sphere is narrow: it arises only where the material consequences of a uniform rule genuinely differ across places, so that a single norm completes fewer bundles than locally tailored ones. For norms whose implementability depends on supra-local institutions, legitimacy may therefore depend less on the volume of justification any single democracy generates than on whether an integrated public sphere exists at the scale at which their implementability must be argued. Digital communication lowers the cost of reach, but reach alone need not produce integration. The same technologies that connect can segment, and the dynamic analysis above carries a warning from the interwar case: rising reach in a fragmented, weakly disciplined public sphere amplifies prevailing claims rather than legitimating norms. The relevant constraint is the existence of shared languages and epistemic standards under which merit claims remain contestable, and implementability claims intelligible, across places. Mass schooling helped build such arenas at national scale (Weber (1976)). Whether analogous arenas can emerge at the scale required by contemporary policy problems remains an open question.

For reference, Appendix Table A1 reports the main notation used throughout the model.

3 From Model to Empirics: Measurement and Identification

This section maps the model to empirical counterparts and lays out the identification strategy. The empirical analysis targets the model's static persuasion predictions: whether peer and state-agent effective precision—the product of exposure and content precision—increases endorsement of institutional universalism, and whether their interaction reveals within-channel bundling or cross-channel fit.

3.1 Empirical counterparts

The model's object is norm-institutionalization: the arena in which the normative *why* of a candidate norm meets the institutional *how* of its realization defines the relevant public sphere. In Switzerland, this arena is cantonal. Cantons retain broad sovereignty over domains not delegated to the Confederation, and Swiss federalism allocates public tasks according to subsidiarity. Even local government is a cantonal competence. The institutional realization of candidate norms—their translation into rules, administration, and enforcement—therefore occurs at the cantonal tier. Accordingly, I use the canton-year as the empirical boundary of the public sphere: if merit and implementability arguments are linked in public discourse at all, the cantonal level is where that link can form.

Consider a canton-year public sphere (c, t). The dependent variable is a binary indicator of endorsement of institutional universalism. Let $Y_{i,c,t} = 1$ if individual i supports equal institutional treatment, including for foreigners, and $Y_{i,c,t} = 0$ if the individual favors preferential opportunities

for Swiss citizens. Thus, the outcome measures endorsement of universalism as a rule of public action rather than a purely private distributive preference.

In the model, persuasion operates through two objects: exposure to a channel $j \in \{S, P\}$, denoted E^j, and the content carried by that channel. Empirically, I construct content-weighted exposure measures

$$X^j_{c,t} \equiv E^j_c \cdot p^j_{c,t}, \qquad j \in \{S, P\}, \tag{1}$$

where E^j_c is a predetermined reach proxy for channel j in canton c, and $p^j_{c,t} \in [0, 1]$ is the channel's universalist tilt in (c, t), measured as the share of workers in channel j who endorse institutional universalism. I proxy the state-agent channel with public-sector employees (communal, cantonal, or federal) and the peer channel with private-sector employees. I compute $p^j_{c,t}$ outside the outcome sample. The outcome sample consists of non-workers, so the channel-level tilt captures the universalist content to which respondents are exposed in their canton-year public sphere, rather than mechanically reflecting their own responses. The model distinguishes the normative *why* and the institutional *how*, but the data do not separately observe merit and implementability arguments. The empirical object $p^j_{c,t}$ proxies the compositional attitude environment of each channel, which the model interprets as informative about argumentative content.

To proxy reach, I use predetermined canton-level stock variables and infrastructure:

$$E^S_c \approx \text{Public-employment share}^{(\text{baseline})}_c, \qquad E^P_c \approx \text{Association density}^{(\text{baseline})}_c. \tag{2}$$

These proxies capture stable differences in opportunities for exposure to state agents and peers: the local public-employment share captures potential contact with state agents, while association density captures the density of civic interaction and peer diffusion. Appendix Table A2 reports the exact construction.

In the model, exposure is individual-specific: $E^j_{i,c,t}$ scales the precision with which citizen i receives reasons from channel j in public sphere (c, t). Empirically, I observe only predetermined canton-level proxies E^j_c. A natural mapping is

$$E^j_{i,c,t} = E^j_c + \eta^j_{i,c,t}, \qquad \mathbb{E}[\eta^j_{i,c,t} \mid c, t] = 0,$$

where $\eta^j_{i,c,t}$ captures within-canton heterogeneity in contacts and participation. Estimating the exposure specification with E^j_c replaces the individual content-weighted exposure $E^j_{i,c,t}p^j_{c,t}$ by the proxy $E^j_c p^j_{c,t}$. Under a classical proxy condition—that $\eta^j_{i,c,t}$ is mean-zero and orthogonal to the structural error conditional on fixed effects and controls—this substitution introduces measurement error that attenuates the estimated exposure effects. If instead $\eta^j_{i,c,t}$ reflects endogenous contact choice, the proxy may also induce selection bias. The baseline design therefore focuses on non-workers who do not move across cantons, so that variation in content-weighted exposure comes from predetermined canton-level reach interacted with canton-year changes in channel content, rather than from individuals' own mobility or workplace contact choices.

3.2 Baseline specification

I estimate

$$Y_{i,c,t} = \theta_S X^S_{c,t} + \theta_P X^P_{c,t} + \theta_{SP}\, X^S_{c,t} X^P_{c,t} + \mathbf{Z}'_{c,t}\Gamma + \alpha_i + \lambda_t + \delta_c \cdot t + \varepsilon_{i,c,t}, \tag{3}$$

where α_i are individual fixed effects, λ_t are year fixed effects, and $\delta_c \cdot t$ are canton-specific linear trends. The vector $\mathbf{Z}_{c,t}$ includes time-varying canton-level controls that may jointly affect workers'

channel-specific universalist tilt and non-workers' endorsement of institutional universalism, such as electoral timing, unemployment, immigration flows, cantonal political ideology, and other canton-year covariates. I include individual time-varying covariates in robustness checks. I demean $X^S_{c,t}$ and $X^P_{c,t}$ in the estimation sample before constructing the interaction, so that θ_P and θ_S capture the marginal effect of one channel's content-weighted exposure evaluated at the mean of the other channel.

Interpretation and link to the model. Equation (3) is a model-disciplined reduced form. The regressors are constructed to match the model's exposure-scaling logic: reach is predetermined at the canton level, while channel content varies over canton-years. The object

$$X^j_{c,t} = E^j_c p^j_{c,t}$$

therefore proxies effective exposure to universalism-relevant content in channel j. The coefficients θ_S and θ_P capture how endorsement responds to marginal increases in state-agent and peer content-weighted exposure. The interaction coefficient θ_{SP} captures whether the two channels are substitutes or complements in generating persuasive (m, h) bundles. In the benchmark with within-channel bundling, each channel can independently supply a persuasive (m, h) bundle. Concavity in endorsement then implies substitutability across channels: the cross-partial is weakly negative absent cross-channel pairing. In the specialization extension, by contrast, citizens can assemble matched (m, h) reasons *across* channels. Cross-channel fit then generates local complementarity at low reach (Proposition 2). The sign of θ_{SP} is therefore informative about the persuasion architecture: a negative interaction is consistent with within-channel bundling, while a positive interaction is consistent with cross-channel fit.

In the model, persuasion depends on effective precision $Q^k_j = \pi^k_j E^j$ (Lemma 1). Empirically, I observe predetermined canton-level reach proxies E^j_c and channel-specific universalist tilt $p^j_{c,t}$, but not the separate merit and implementability components of precision. The specification therefore uses $X^j_{c,t} = E^j_c p^j_{c,t}$ as the empirical counterpart of effective exposure to universalism-relevant public reasons. The interaction $X^S_{c,t} X^P_{c,t}$ asks whether increases in exposure to both channels jointly raise endorsement more or less than additively, as predicted by the model's distinction between within-channel bundling and cross-channel fit.

Identification. The baseline design isolates persuasion within a fixed local public sphere. Because $X^j_{c,t}$ is defined at the canton-year level, canton-by-year fixed effects would absorb the identifying variation. I therefore use individual fixed effects, year fixed effects, canton-specific linear trends, and rich canton-year controls, and restrict the baseline sample to canton non-movers. This restriction avoids endogenous relocation across cantons and ensures that within-individual variation in content-weighted exposure comes from changes in channel-specific tilt $p^j_{c,t}$ within a predetermined canton-level reach environment, rather than from individuals moving into different exposure environments.

The identifying variation in the baseline specification comes from within-individual changes in

$$X^j_{c,t} = E^j_c p^j_{c,t},$$

induced by canton-year changes in channel-specific tilt $p^j_{c,t}$, conditional on individual fixed effects, year fixed effects, canton trends, and canton-year covariates. The baseline exposure design is therefore best interpreted as identifying the model's public-sphere persuasion architecture: whether peer and state-agent content-weighted exposures increase endorsement, and whether their interaction is consistent with within-channel bundling or cross-channel fit.

This design does not, by itself, fully rule out residual canton-year shocks correlated with both workers' universalist tilt and non-workers' endorsement. For this reason, I complement it with a workplace-switch design, which delivers within-person shifts in direct exposure to state agents while holding canton-year discourse conditions fixed. The switch design provides the sharper causal test of the effect of state-agent exposure on endorsement.

3.3 Workplace exposure shocks

The baseline exposure design identifies the public-sphere persuasion architecture at the canton-year level. To obtain sharper causal evidence on direct exposure to state agents, I exploit switches between private and public employment. In the model, the state-agent channel is characterized by a content and precision mix (π_S^m, π_S^h), empirically proxied at the canton-year level by $p_{c,t}^S$, measured outside the outcome sample. A worker's sector switch does not mechanically change this canton-year content. What changes is the worker's access to state-agent contacts through the workplace. Entering public employment raises the intensity of day-to-day interactions with public employees, and hence the individual exposure component $E_{i,c,t}^S$; leaving public employment reduces it.

Baseline switcher specification. I estimate the following fixed-effects regression, where identification comes from individuals who switch sector:

$$Y_{i,c,t} = \eta\, State_{it} + \psi\, Fed_{it} + \alpha_i + \delta_{c,t} + \varepsilon_{i,c,t}. \tag{4}$$

The unit of observation is an individual i in canton c and year t. The variable $State_{it}$ equals one if individual i works in the public sector—communal, cantonal, or federal—and zero if the individual works in the private sector, including private nonprofit employment. The variable Fed_{it} equals one if the individual works in the federal administration and zero otherwise. Thus, η captures the effect of moving into non-federal public employment relative to private employment, while ψ captures the incremental effect of federal employment relative to other public employment. The total effect of federal employment relative to private employment is therefore $\eta + \psi$. Individual fixed effects α_i absorb time-invariant selection into sector. Canton-by-year fixed effects $\delta_{c,t}$ absorb the local public-sphere environment and all canton-year shocks common to workers in the same canton and year. Standard errors are clustered at the individual level.

The key advantage of Equation (4) relative to the baseline exposure specification is the inclusion of canton-by-year fixed effects. In the baseline design, the treatment variables $X_{c,t}^j$ vary only at the canton-year level, so canton-by-year fixed effects would absorb the identifying variation. In the workplace-switch design, treatment varies at the individual-year level through $State_{it}$. I can therefore include $\delta_{c,t}$, which holds fixed the entire canton-year public sphere: the local discourse environment, worker universalist tilt $p_{c,t}^j$, local political shocks, labor-market conditions, immigration shocks, and any other canton-year factor common to individuals in the same canton and year. Identification comes from within-person changes before and after a sector switch, benchmarked against stayers observed in the same canton-year environment. Thus, the estimate isolates the effect of direct workplace access to state-agent contacts, rather than variation in canton-year discourse conditions.

By construction, a sector switch changes the composition of daily professional interactions. It generates a discrete shift in contact intensity with public employees relative to residential exposure alone. This workplace-based exposure shock is therefore less subject to attenuation from canton-level reach proxies than the baseline content-weighted exposure specification.

Dynamic self-selection and service-type sorting. Equation (4) addresses time-invariant selection into sector, but a remaining concern is dynamic self-selection: individuals may become

more favorable to institutional universalism before entering public employment. Prior work shows that civic-minded agents may sort into caring services—health, education, and social care—that are delivered by both public and private providers (Francois (2003), Gregg et al. (2011)). This implies that selection on endorsement is likely to operate primarily at the service-type level rather than at the institutional-sector level.

I probe this possibility by comparing pre-switch endorsement among future switchers and stayers while individuals are still employed in the private sector. Among private-sector workers, I estimate

$$Y_{i,c,t}^{Priv} = \phi Switch_i^{pub} + \delta_{c,t} + \mathbf{Z}'_{i,t}\gamma + u_{i,t}, \tag{5}$$

where $Y_{i,c,t}^{Priv}$ is endorsement measured while individual i is employed in the private sector. The variable $Switch_i^{pub}$ equals one if individual i switches from private to the public employment at any point in the future. Canton-by-year fixed effects $\delta_{c,t}$ compare eventual switchers to stayers within the same canton-year. The vector Z_{it} includes individual covariates: gender, number of children, age, religion, culture, college education, marital status, health status, and urban residence. I also repeat the exercise excluding federal employment, allowing selection patterns to differ across institutional tiers.

To sharpen the service-type implication, I partition jobs into HES sectors—health, education, and social care—and non-HES sectors. I then re-estimate Equation (5) separately in each private-sector subsample using destination-specific indicators: $Switch_i^{pub,HES}$, equal to one if individual i later switches from private to public HES employment, and $Switch_i^{pub,non\text{-}HES}$, equal to one if individual i later switches from private to public non-HES employment. Appendix Figure A3 shows that pre-switch differences are confined to service type: future switchers into public HES exhibit higher baseline endorsement, whereas future switchers into public non-HES do not. Accordingly, the causal switch design focuses on private-to-public transitions within non-HES industries.

Finally, I compare baseline characteristics of private non-HES workers who subsequently move to public non-HES with those who start in public non-HES and remain there (Appendix Table A5). Switchers are slightly younger and less likely to hold a college degree. In Switzerland, comparable public jobs pay a premium—on average about 12%, and up to 29% at the lower end of the pay scale (Portmann et al. (2024))—so pecuniary considerations plausibly contribute to switching within non-HES occupations.

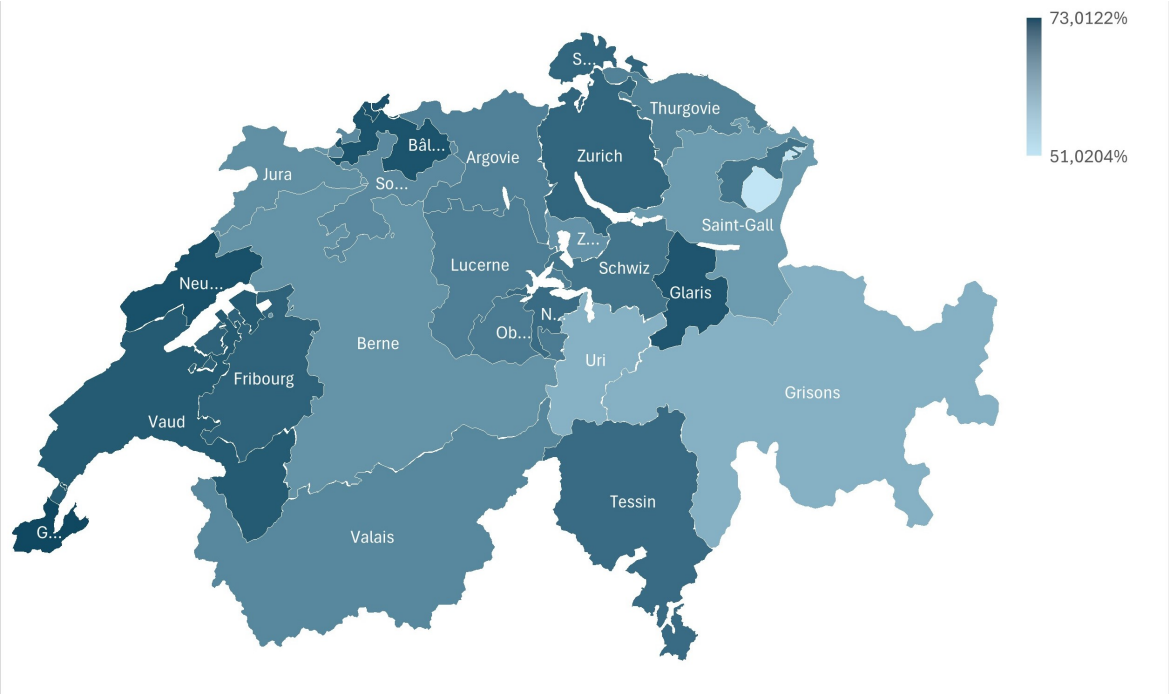
4 Data

Swiss Household Panel (SHP). The main individual-level data come from the Swiss Household Panel (SHP), a longitudinal survey conducted annually since 1999 that follows a random sample of approximately 5,000 households, covering over 12,000 individuals residing in Switzerland. The estimation window is 1999–2003, the five waves of the SHP's original sample: the 2004 wave introduced both a large refreshment sample (SHP_II), changing the composition of the respondent pool from which the channel content measures $p_{c,t}^j$ are constructed, and a reformulation of the employment-type instrument on which the workplace-switch classification relies. I therefore hold the window common across designs, over which both treatment constructions are internally consistent. The panel is geocoded at the municipality of residence, which I aggregate to the canton-year level to construct the empirical counterparts to local public-sphere environments. For the content-weighted exposure design, I restrict the sample to Swiss nationals aged 18 and older who are not working and canton non-movers, the restrictions motivated in Section 3. I focus on Swiss nationals because the institutional-universalism item explicitly asks respondents to trade off opportunities for Swiss citizens versus foreigners. The baseline sample contains $N = 6,920$ individuals and 19,243 person-year

observations (Appendix Table A3); conditioning on item response and on individuals observed at least twice—as the individual fixed effects require—the baseline estimation sample comprises 8,426 person-year observations.

Institutional universalism. The outcome measures endorsement of institutional universalism: support for equal institutional treatment of socially distant others rather than preferential treatment of the national in-group. The SHP asks respondents: *'Are you in favour of Switzerland offering foreigners the same opportunities as those offered to Swiss citizens, or in favour of Switzerland offering Swiss citizens better opportunities?"* Responses are: (i) 'in favour of equality of opportunities," (ii) 'neither," and (iii) 'in favour of better opportunities for Swiss citizens." I define $Y_{i,c,t} = 1$ for respondents choosing equality of opportunities and $Y_{i,c,t} = 0$ for respondents choosing better opportunities for Swiss citizens. Respondents answering "neither" are excluded from the baseline measure. This category accounts for less than 10% of responses and plausibly pools heterogeneous positions, including ambivalence, non-attitudes, and social-desirability responding. Mean endorsement in the full Swiss-adult sample is 0.657; equivalently, 34.3% prefer preferential opportunities for Swiss citizens. In the baseline estimation sample, mean endorsement is 0.616, reflecting the additional sample restrictions. Figure 1 documents sizeable cross-canton heterogeneity in endorsement, ranging from roughly 51% to 73% across cantons.

Figure 1: Endorsement of institutional universalism averaged at the state (cantonal) level



Note: The map displays the average share of Swiss individuals aged 18+ in each canton who endorse the universalism institutional norm. This corresponds to the cantonal average of the variable Y_{ict}.

Individual-level correlates. Appendix Figure A1 reports OLS associations between $Y_{i,c,t}$ and individual covariates, including education, ideology, language, urbanicity, family structure, religiosity,

and income. Endorsement is higher among college-educated, left-leaning, urban, higher-income, and French-speaking respondents, and lower among more religious respondents and those with more children. These correlations are descriptive and characterize the sample. The main specifications rely on within-individual variation. Importantly for the individual fixed-effects design, endorsement exhibits substantial within-person variation: about one quarter of respondents switch their stance at least once over 1999–2003 (Appendix Table A3).

Institutional sectors. The SHP records respondents' institutional sector through the question: *"Are you employed by a private company or a state organization?"* I define $State_{it} = 1$ if respondent i reports being employed by a state organization—communal, cantonal, or federal—and $State_{it} = 0$ if employed by a private organization, including non-profit employment. I exclude respondents employed by international organizations, which are not part of the domestic state apparatus. I separately code federal employment using $Fed_{it} = 1$ for confederation-level employment and $Fed_{it} = 0$ otherwise. The SHP also reports industry affiliation. I partition jobs into HES industries—health, education, and social care—and non-HES industries using the economic-activity classification. This split distinguishes mission-oriented public-service industries from other activities and is used both for descriptive comparisons and to sharpen interpretation in the workplace-switch design.

Appendix Table A4 reports descriptive endorsement rates by institutional sector. Public-sector employees are more likely to endorse institutional universalism than private-sector employees, 74% versus 66%. This gap is present both within HES and within non-HES industries, suggesting that it is not confined to traditionally mission-oriented fields. Additional descriptives also reveal heterogeneity within the public sector: endorsement is higher among cantonal and communal employees than among federal employees, 78% versus 67%, respectively. Cantonal and communal employees exercise public authority within subnational institutional environments, whereas federal employees occupy roles within the national administration. I therefore allow federal employment to have a distinct effect in the workplace-switch design. Appendix Figure A2 plots endorsement trends separately for public- and private-sector employees. The two series follow distinct trajectories and are only weakly correlated at the canton-year level ($\rho = 0.32$). This descriptive pattern supports treating the two channels separately in the construction of the canton-year content environments, p^S_{ct} and p^P_{ct}.

Switch-design sample. For the workplace-exposure design, I restrict the sample to employees aged 20–54, to focus on voluntary job-to-job mobility, and exclude the self-employed. Over 1999–2003, this worker panel contains 3,777 individuals and 9,849 person-year observations. I further partition jobs into HES and non-HES industries and emphasize switches within non-HES sectors. Among non-HES workers, 8.6% switch institutional sector during the observation window. Switches occur in both directions in roughly equal numbers, and no individual switches more than once. These mobility patterns are consistent with the relative permeability of Swiss public employment: recruitment is typically open, without competitive entrance examinations, and formal civil-service status was abolished in all but two cantons during the 1990s (Audier and Bacache-Beauvallet (2007)).

Appendix Table A2 provides definitions, construction details, and data sources for all variables used throughout the empirical analysis, while Appendix Table A3 reports their summary statistics.

5 Empirical Results

5.1 Content-weighted exposure estimates

Table 1 reports estimates of the baseline content-weighted exposure specification in Equation (3). The sample consists of non-working Swiss respondents who do not move across cantons during the estimation window[8]. This design estimates the model's reduced-form building blocks: whether peer and state-agent content-weighted exposures each raise endorsement—that is, whether the two channels constitute distinct sources of influence rather than a single undifferentiated discourse environment—and how their contributions combine at the margin.

Because $X^P_{c,t}$ and $X^S_{c,t}$ are centered in the estimation sample, the main-effect coefficients θ_P and θ_S can be read as marginal effects evaluated at the mean exposure of the other channel. Differentiating Equation (3) yields

$$\frac{\partial Y_{i,c,t}}{\partial X^P_{c,t}} = \theta_P + \theta_{SP} X^S_{c,t}, \qquad \frac{\partial Y_{i,c,t}}{\partial X^S_{c,t}} = \theta_S + \theta_{SP} X^P_{c,t}. \tag{6}$$

Thus, θ_{SP} captures how the marginal effect of one channel varies with exposure to the other channel. The two signs are not equally informative. By Proposition 2, a positive interaction is possible only under cross-channel fit: no concave persuasion technology without merit–implementability structure can generate it at any exposure level. A negative interaction is instead the generic case: it arises under within-channel bundling (Proposition 1), but it would equally arise under any technology combining two exposure measures through a concave aggregator. The sign of θ_{SP} therefore carries asymmetric information about the persuasion architecture.

All specifications include individual fixed effects and year fixed effects. Because the baseline sample is restricted to canton non-movers, individual fixed effects absorb time-invariant canton differences. Column (1) presents a parsimonious specification without time-varying canton controls or canton trends. Column (2) adds contemporaneous canton-year controls to absorb observable canton-year factors that could jointly affect workers' channel-specific universalist tilt and non-workers' endorsement, such as electoral timing, unemployment, immigration flows, cantonal political ideology, and other canton-year covariates. Column (3) is the preferred specification: it further adds canton-specific linear trends, absorbing gradual canton-level changes in political culture, demographic composition, or other unobservables that could correlate with trends in $p^j_{c,t}$. Column (4) adds time-varying individual controls as a robustness check. These controls are introduced to verify that the estimates are not driven by observable life-cycle changes not absorbed by individual fixed effects. Column (5) replaces contemporaneous canton controls with their one-period lags, while keeping the electoral calendar contemporaneous, to address the concern that contemporaneous canton covariates may themselves respond to shocks that also shift the public-sphere discourse environment.

Across specifications, peer and state-agent content-weighted exposures are positively associated with endorsement of institutional universalism. Since $Y_{i,c,t}$ is binary and Equation (3) is estimated as a linear probability model, coefficients can be read in percentage points. In the preferred specification, Column (3), a one-standard-deviation increase in peer content-weighted exposure, $\mathrm{sd}(X^P_{c,t}) = 0.165$, increases endorsement by about 5.5 percentage points, evaluated at the mean state-agent exposure. A one-standard-deviation increase in state-agent content-weighted exposure, $\mathrm{sd}(X^S_{c,t}) = 1.351$, increases endorsement by about 4.4 percentage points, evaluated at the mean peer exposure. Relative to a baseline endorsement rate of about 0.62, these correspond to increases of

[8] As described in Section 3.1, channel exposures are constructed as $X^j_{c,t} = E^j_c \, p^j_{c,t}$, using predetermined canton-level reach proxies E^j_c and canton–year variation in the channel discourse environment $p^j_{c,t}$, computed outside the outcome sample. Results are robust to including canton movers (Appendix A.2 Table A6).

roughly 8.9% and 7.1%, respectively. The positive own effects carry two implications. First, because state-agent effective precision retains a sizeable, statistically robust effect net of peer exposure and the full control set, the two channels constitute distinct sources of influence: the identity of the source matters, which an undifferentiated-discourse view, in which only total exposure counts, would not by itself generate. Second, both own effects are positive, which rules out a pure crowd-out interpretation in which exposure to one channel displaces or discounts the other. However, because the baseline design relies on canton-level reach proxies, these estimates should be interpreted as evidence on the two channels' distinct exposure–endorsement gradients, rather than as the sharpest causal estimate of direct state-agent exposure. The workplace-switch design below provides that complementary causal test by exploiting individual-level changes in access to state agents while holding the canton-year discourse environment fixed.

The interaction term is negative across specifications. In the preferred specification, a one-standard-deviation increase in state-agent exposure reduces the effect of a one-standard-deviation increase in peer exposure by about 1.6 percentage points. Equivalently, the peer effect falls from 5.5 percentage points at mean state-agent exposure to about 3.9 percentage points when state-agent exposure is one standard deviation above its mean. The same calculation applies symmetrically: a one-standard-deviation increase in peer exposure reduces the effect of a one-standard-deviation increase in state-agent exposure from about 4.4 to 2.8 percentage points. Thus, the interaction is not only statistically negative but economically meaningful: the marginal effect of each channel is substantially smaller when the other channel is already strong. This substitution pattern rules out a purely additive persuasion technology, in which each channel's contribution to endorsement simply sums: some form of diminishing joint returns is present, and the marginal persuasive value of each channel is substantially smaller when the other is already strong. It is consistent with the benchmark bundling model of Proposition 1. On its own, however, the negative sign is not diagnostic of that architecture: as noted above, it would obtain under any concave aggregator, whether or not the underlying technology has merit–implementability structure. Section 6 provides the discriminating evidence, exploiting the asymmetry established in Proposition 2: a positive interaction, wherever observed, is possible only under cross-channel fit.

The estimates are stable when adding time-varying individual controls in Column (4), supporting the view that the baseline variation is not driven by observable life-cycle changes. The content-weighted exposure coefficients also remain qualitatively similar when contemporaneous canton controls are replaced by lagged canton controls in Column (5), and the negative interaction becomes larger in magnitude. This suggests that the level effects and substitution pattern are not artifacts of the particular timing of the canton controls.

5.2 Workplace-switch estimates

Table 2 estimates Equation (4) in a linear probability model with individual fixed effects and canton-by-year fixed effects. This is the key difference from the baseline exposure design. Because the treatment varies at the individual-year level through sector switches, I can include $\delta_{c,t}$, which absorbs the entire canton-year public-sphere environment: channel-specific discourse $p^S_{c,t}$ and $p^P_{c,t}$, local political shocks, labor-market conditions, immigration shocks, campaigns, and any other canton-year factor common to individuals in the same canton and year. Identification therefore comes from within-person changes before and after a sector switch, benchmarked against stayers observed in the same canton-year environment.

Column (1) uses all observed switches between private and public employment. The coefficient on $State_{it}$, $\hat{\eta}$, is positive and statistically significant: entering public employment increases the probability of endorsing institutional universalism. Since canton-by-year fixed effects hold fixed the

Table 1: Exposure to channels and endorsement of institutional universalism

Y_{ict}	(1)	(2)	(3)	(4)	(5)
	Endorsement of institutional universalism				
Peer c.w. exposure (X^P_{ct})	0.239***	0.231**	0.333***	0.331***	0.300**
	(0.074)	(0.085)	(0.092)	(0.094)	(0.123)
State-agent c.w. exposure (X^S_{ct})	0.025*	0.027**	0.033**	0.033**	0.041**
	(0.012)	(0.011)	(0.013)	(0.012)	(0.015)
Peer × state-agent ($X^P_{ct} \times X^S_{ct}$)	-0.040*	-0.035*	-0.070**	-0.069**	-0.108***
	(0.022)	(0.019)	(0.027)	(0.027)	(0.035)
Observations	8,426	8,349	8,349	8,341	8,349
Individual fixed effects	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes
Contemporaneous canton controls	No	Yes	Yes	Yes	No
Canton-specific linear trends	No	No	Yes	Yes	Yes
Time-varying individual controls	No	No	No	Yes	No
Lagged canton controls					Yes

Notes: The dependent variable Y_{ict} is a binary indicator equal to one if respondent i endorses institutional universalism in canton c and year t. Sample: Swiss respondents aged 18+; non-workers; canton non-movers. All specifications estimate equation (3) by OLS with individual, canton, and year fixed effects; standard errors are clustered at the canton level. The peer and state-agent content-weighted (c.w.) exposures, X^P_{ct} and X^S_{ct}, are mean-centered in the estimation sample so that the main-effect coefficients are evaluated at the mean of the other channel's exposure. *Contemporaneous canton controls* include: (i) the weighted-average left–right ideology of the cantonal executive (constructed by mapping parties' cabinet shares onto a 1–4 scale using party positions from Vatter et al. (2024)); (ii) foreign population share; (iii) number of referenda; (iv) unemployment rate; and (v) taxable income per capita (cantonal average). *Lagged canton controls* replace (ii)–(v) by their one-year lag (while keeping the electoral calendar contemporaneous). *Time-varying individual controls* include marital-status indicators, number of children, and an indicator for health limitations in daily activities. Column (3) is the preferred specification; Columns (4)–(5) provide robustness to adding individual controls and to using lagged canton covariates. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

local discourse environment, this effect is not driven by changes in the canton-year universalist tilt of public employees. It reflects a change in individual access to state-agent contacts through the workplace. Columns (2)–(4) progressively tighten the identifying comparison. Column (2) restricts attention to switches that occur within the same service type, HES versus non-HES, addressing selection into caring-service occupations that may operate in both sectors. Column (3) focuses on private-to-public switches. Column (4) further restricts to private-to-public switches within non-HES jobs, which is the preferred causal estimate because pre-switch comparisons show no differential baseline endorsement for future switchers into public non-HES employment. Across Columns (1)–(4), the point estimate increases as switches into public HES jobs are excluded. This is consistent with the pre-switch evidence: future public-HES switchers are already universalist before entering the public sector. Including them therefore attenuates the estimated workplace-exposure effect.

The preferred estimate in Column (4) implies that moving from private to public employment within non-HES industries raises the probability of endorsing institutional universalism by about 10 percentage points. Relative to a mean endorsement rate of 0.65 among private non-HES workers, this corresponds to an increase of about 15%. Appendix Figure A4 shows that the estimate is not driven by any single canton. In Column (3) of Table 1, a one-standard-deviation increase in peer content-weighted exposure raises endorsement by about 5.5 percentage points, while a one-standard-deviation increase in state-agent content-weighted exposure raises endorsement by about 4.4 percentage points, both evaluated at the mean of the other channel's exposure. The switch estimate of 9.9 percentage points is 2.25 times the one-standard-deviation state-agent exposure effect. This larger within-person estimate is consistent with attenuation in the baseline canton-level reach proxies and with the model's prediction that endorsement rises with effective exposure to persuasive public reasons $E^S_{i,c,t}$.

The coefficient on Fed_{it} is negative, implying that the effect of switching into federal public employment is smaller than the effect of switching into cantonal or communal public employment. Quantitatively, the implied net effect of federal employment relative to private employment, $\eta + \psi$, is close to zero. This sharp institutional-tier pattern is consistent with the paper's mechanism operating across institutional tiers: the role conception attached to federal office is plausibly the most fully Weberian—conceived more strictly as an implementation and coordination device than cantonal or communal roles. Federal agents internalize this role conception, and their public discourse is correspondingly specialized in the institutional *how*. Because endorsement is bundle-based, exposure to federal agents is weakly persuasive on its own. By contrast, cantonal and communal agents, whose roles are conceived less strictly, more often bundle the normative *why* with the institutional *how*, and exposure to them is correspondingly more persuasive.

Column (5) probes whether the estimated effect of state-agent exposure is instead driven by observable features of public jobs. I augment the preferred private-to-public non-HES specification with job-attribute controls: log income, perceived workplace atmosphere, and job security. The public-employment estimate remains similar in magnitude, while the added job attributes are close to zero. This suggests that the switch effect is not simply picking up contemporaneous changes in pay, job security, or reported job quality.

Column (6) considers the reverse transition, from public to private employment, within non-HES industries. The estimated change upon moving to the private sector is small and imprecise. Given the smaller switcher sample in this column, I treat the absence of a detectable reversal as suggestive rather than definitive. This pattern is however consistent with persistence under learning: if workplace exposure shifts beliefs, the posterior need not immediately revert when exposure falls, absent offsetting new information.

Table 2: Shift in exposure to state agents and endorsement of institutional universalism

Y_{ict}	(1) All switches	(2) All switches within non-HES & HES	(3) Private to public within non-HES & HES	(4) Private to public within non-HES	(5) Private to public within non-HES	(6) Public to private within non-HES	(7) Private to public within non-HES
$State_{it}$	0.046**	0.053**	0.084**	0.099**	0.103*		0.067
	(0.020)	(0.027)	(0.034)	(0.049)	(0.056)		(0.055)
Fed_{it}	-0.051*	-0.048	-0.114**	-0.122*	-0.133*		-0.144**
	(0.027)	(0.035)	(0.055)	(0.069)	(0.078)		(0.069)
Ln yearly income					0.002		
					(0.025)		
Job atmosphere					-0.003		
					(0.006)		
Job security					0.004		
					(0.011)		
Private sector						-0.027	
						(0.043)	
$State_{it} \times French_i$							0.127**
							(0.063)
Observations	8,672	4,709	6,343	4,531	3,906	1,084	4,531
Number of individuals	3,590	2,222	2,794	1,960	1,805	441	1,960

Notes: The dependent variable is an indicator equal to one if respondent i endorses institutional universalism in year t. $State_{it}$ equals one when i is employed in the public sector (communal/cantonal or federal) and zero when employed in the private sector (for-profit or non-profit). Fed_{it} equals one when i is employed by the federal administration. The coefficient on $State_{it}$ therefore compares non-federal public employment to private employment; Fed_{it} is the incremental effect of federal (relative to non-federal) public employment. All specifications include individual fixed effects and canton×year fixed effects. Standard errors are clustered at the individual level. Column (5) augments the specification of Column (4) with job-attribute controls. Column (6) restricts to public-to-private switchers within non-HES and uses a private-employment indicator; the reported coefficient therefore captures the within-person change upon moving from public to private. *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.

6 Mechanisms: Persuasion and Its Architecture

The previous section established the reduced-form facts: peer and state-agent exposures constitute distinct sources of influence, both raise endorsement, their interaction is negative on average, and a private-to-public switch raises endorsement by about 10 percentage points within the same canton-year environment. I now examine the mechanism which drives the identified effects. Each design faces its own leading alternatives to persuasion: in the content-weighted exposure design, where respondents are non-workers, generic ideological diffusion or a broad shift in the local demand for universalism; in the workplace-switch design, socialization into a public-sector role and self-interest as a newly public employee. Section 6.1 addresses this question with adjacent-outcome tests. Each alternative carries a distinct signature across cognate political attitudes, and the estimates are inconsistent with all of them, while displaying content-specific patterns consistent with argument-based updating—the combination that licenses reading the estimated effects as persuasion through public reasons. The second question is which architecture governs that persuasion: whether merit and implementability reasons travel bundled within channels or are assembled across them. Section 6.2 addresses it with the model's two culturally grounded shifters—the Weberian culture and the grammar of legitimacy—which the model predicts should move the architecture in opposite directions. The sharpest test exploits the asymmetry of Proposition 2. A positive peer–state-agent interaction, at any exposure level, is possible only under cross-channel fit. Hence, observing the interaction change sign across cultural environments does more than document heterogeneous treatment effects: it reveals that the persuasion architecture itself varies with the environment.

6.1 Adjacent-outcome mechanism tests

I re-estimate the baseline exposure and workplace-switch specifications on a set of adjacent societal and political attitudes: welfare spending, redistribution through taxation of high incomes, support for a strong Swiss army, support for environmental protection over economic growth, self-reported left–right ideology, satisfaction with democracy, trust in the federal government, and support for European Union membership (Appendix Table A2 reports exact variable definitions). Each outcome targets an alternative mechanism. Ideology, and the standard policy positions—redistribution, army, environment—carry the signature of generic ideological diffusion or a broad shift in the local demand for universalism: if exposure proxies a canton drifting left or becoming generically more open, these should move together with endorsement. Welfare spending and redistribution additionally carry the signature of self-interest in the switch design: a newly public employee's material stake is in a larger public sector. Trust in the federal government and satisfaction with democracy carry the signature of socialization and institutional loyalty. Support for European Union membership plays a distinct role: it is the outcome closest in content to institutional universalism itself—equal treatment extended beyond the national in-group—so its behavior is informative about the *breadth* of the persuasive content: a narrowly targeted persuasion should move the target norm without moving even its nearest cognate at moderate doses, with spillover expected first where the persuasive effect is strongest. Because eight outcomes are tested per design, I place the diagnostic weight on the joint pattern—null estimates precisely where each alternative predicts effects—rather than on the statistical significance of any single coefficient, and I read the two marginally significant results below with corresponding caution.

Adjacent outcomes: Content-weighted exposure design. Table 3 re-estimates Equation (3) using the above-mentioned adjacent political and societal attitudes as dependent variables. Across outcomes, the coefficients on peer exposure, state-agent exposure, and their interaction are generally

small and statistically indistinguishable from zero. In particular, content-weighted exposure does not systematically predict left–right ideology, support for a strong army, environmental attitudes, satisfaction with democracy, or trust in the federal government. This pattern weighs against the interpretation that the baseline results reflect broad ideological realignment, generic political salience, or a generalized increase in institutional trust. Support for European Union membership—universalism's nearest cognate—remains unmoved, consistent with narrowly targeted persuasive content rather than a broad opening of attitudes.

Two exceptions are diagnostically important. First, peer exposure significantly reduces support for welfare spending. This result is counterintuitive under a simple pro-social, egalitarian, or left-wing preference-shift interpretation: the same peer exposure that predicts stronger endorsement of institutional universalism predicts lower support for expansive public provision. The pattern is instead consistent with the model's mechanism. Peer discourse appears to make equal institutional treatment compatible with a less expansive view of public provision. In the model's terms, the peer channel does not merely increase exposure to universalist agents; it carries a particular argumentative content about how universalism should be institutionally realized.

Second, the peer–state-agent interaction is positive for satisfaction with democracy, although this result is only marginally significant. The pattern is nevertheless consistent with the bundle-based mechanism: when peer and state-agent discourse jointly circulate, democratic institutions may appear more compatible with the realization of universalistic norms. Importantly, this effect does not extend to trust in the federal government, which suggests that the mechanism is not a generalized increase in institutional trust.

Table 3: Exposure to channels and adjacent outcomes

	(1) Welfare	(2) Redistri.	(3) Env. vs growth	(4) Army	(5) Ideology	(6) Join EU	(7) Trust fed. gov.	(8) Sat. Dem.
X^P_{ct}	-0.275**	0.184	0.071	0.162	0.147	0.006	0.294	0.330
	(0.105)	(0.161)	(0.145)	(0.128)	(0.297)	(0.097)	(0.382)	(0.403)
X^S_{ct}	-0.005	0.007	0.008	-0.008	-0.010	0.005	-0.026	-0.003
	(0.016)	(0.019)	(0.014)	(0.019)	(0.046)	(0.014)	(0.036)	(0.044)
$X^P_{ct} \times X^S_{ct}$	0.069	-0.016	0.009	-0.038	0.084	-0.009	0.018	0.196*
	(0.044)	(0.037)	(0.050)	(0.046)	(0.113)	(0.026)	(0.095)	(0.098)
Observations	9,494	9,604	9,707	9,732	8,528	8,497	9,819	9,781

Notes: Mechanism tests are estimated on the same outcome sample as the baseline specification: Swiss respondents aged 18+; non-workers; canton non-movers. The peer and state-agent content-weighted (c.w.) exposures X^P_{ct} and X^S_{ct} are mean-centered in the estimation sample so that main effects are evaluated at the mean of the other channel's exposure. OLS models with individual, canton, and year fixed effects, canton-specific linear trends, time-varying canton controls, and clustering at the canton level, as in column (3) of Table 1. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Adjacent outcomes: Workplace-switch design. Table 4 re-estimates the preferred workplace-switch specification on the same adjacent outcomes as in Table 3. Across welfare spending, redistribution, environmental preferences, support for a strong army, political ideology, support for European Union membership, and trust in the federal government, the estimated effects of switching into public employment are small and statistically indistinguishable from zero. The welfare and redistribution nulls do specific work here: self-interest as a newly public employee predicts stronger preferences for a larger, better-funded public sector, and neither coefficient moves. The switch estimate is likewise robust to controlling directly for pay, job security, and workplace atmosphere (Table 2, Column 5), closing the compensation side of the same alternative.

The only systematic movement is for satisfaction with democracy. Switching into non-federal

public employment is associated with a lower level of satisfaction with democracy, although the coefficient is only marginally significant. It is consistent with a content-specific mechanism in which exposure to state agents increases endorsement of institutional universalism while also making respondents more aware of the gap between universalistic rules and existing democratic performance. In other words, state-agent exposure does not appear to induce generalized institutional loyalty. If anything, it combines stronger endorsement of equal institutional treatment with a more critical assessment of how democratic institutions currently perform.

Table 4: Shift in exposure to state agents and adjacent outcomes

	(1) Welfare	(2) Redistri.	(3) Env. vs growth	(4) Army	(5) Ideology	(6) Join EU	(7) Trust fed. gov.	(8) Sat. Dem.
$State_{it}$	-0.105	0.029	-0.036	-0.049	-0.091	0.078	-0.300	-0.380*
	(0.107)	(0.122)	(0.080)	(0.115)	(0.209)	(0.077)	(0.263)	(0.226)
Fed_{it}	0.167	-0.052	0.183	0.127	0.117	-0.003	0.234	0.191
	(0.171)	(0.153)	(0.116)	(0.139)	(0.337)	(0.089)	(0.326)	(0.311)
Observations	4,980	5,036	2,988	5,031	4,451	4,568	5,091	5,046
Number of individuals	2,030	2,040	1,559	2,036	1,898	1,951	2,052	2,039

Notes: The sample is restricted to private-to-public switchers and stayers in non-HES industries, as in the preferred switch specification. All specifications include individual fixed effects and canton-by-year fixed effects. Standard errors are clustered at the individual level. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

6.2 Persuasion architecture mechanism tests

Weberian culture. The first persuasion architecture shifter is Weberian culture: the extent to which the surrounding culture expects a strict separation between technocratic administration and normative authority. Stronger Weberian culture sharpens state agents' comparative-cost advantage in implementability while raising their relative cost of supplying merit arguments under the cost structure of Assumption 1. Hence, as W_c increases, $\pi_S^m(W_c)$ falls and $\pi_S^h(W_c)$ rises (Corollary 1). Because a complete within-channel bundle requires both factors of Λ_S to be bounded away from zero, Weberian culture weakens the state-agent channel's capacity to complete the endorsement bundle on its own and shifts its contribution toward the cross-channel term $\Lambda_\times$. The empirical implication is therefore a sign prediction on the peer–state-agent interaction: in more Weberian environments, the interaction should become less negative and may turn positive. By Proposition 2, a positive interaction at any exposure level is possible only under cross-channel fit. Thus, observing a sign reversal is evidence that persuasion shifts from bundled within-channel reasons to reasons assembled across peer and state-agent channels.

I proxy W_c with popular support for state secularity, the cantonal vote share in the 1980 federal popular initiative for the complete separation of church and state. This provides a continuous, population-level measure of demand for role separation between technocratic administration and normative authority. The initiative committee's own campaign literature makes this connection to administrative rationalization explicit. In its October 1978 submission to the National Council's preparatory commission, the committee approvingly cites constitutional scholar Zaccaria Giacometti's argument that a state proclaiming freedom of religion positions itself apart from questions of faith precisely in order to remove them from the state's competence (Eidgenössisches Aktionskomitee für die Trennung von Staat und Kirche (1978)). The same brief also justifies separation on administrative grounds, arguing that continued state financing of churches is not market-conforming and violates the principle of subsidiarity that should govern public administration. This is the cultural logic W_c

is meant to capture: not opposition to religion as such, but a view of the state as an institution whose legitimate competence is confined to administrable, rule-bound matters, with questions of norms and belief relegated outside the domain of state authority. The Federal Council's message recommending rejection confirms that the initiative was understood along this axis. Its defense of existing church–state arrangements emphasized the churches' role in transmitting foundational values and enabling citizens not merely to obey the law, but to affirm it inwardly (Schweizerischer Bundesrat (1978, §532.2)). The vote therefore separates electorates more willing to confine the state's legitimate competence to administrable matters from those more willing to link state authority to normative grounds. In the model, this cultural axis matters because it shapes the role conception internalized by state agents: where state authority is understood as confined to administrable, rule-bound matters, state agents specialize in implementability arguments, while merit arguments become relatively costlier for them to supply.

To isolate W_c from the second cultural margin examined below, I restrict the estimation sample to German-speaking cantons. This restriction does more than avoid language confounding. Where the two cultural primitives co-vary—Geneva, at once Calvin's and Rousseau's city, is plausibly high on both—their predicted effects superpose: the grammar of legitimacy strengthens both channels' effort while Weberian culture tilts the state channel's content toward implementability, and the model then makes no sharp prediction for the sign of the interaction. Within German-speaking cantons, variation in secularity support isolates the composition margin alone. I also verify that the mechanism shifter does not mechanically generate differential trends in the content-weighted exposure measures themselves. Specifically, I estimate $X^j_{ct} = a_j + \varphi_j W_c \times t + \delta_c + \lambda_t + \nu^j_{ct}$ for $j \in \{S, P\}$. The coefficient (φ_j) is small and statistically indistinguishable from zero. Thus, after absorbing canton and year fixed effects, W_c does not predict differential trends in peer or state-agent exposure. The heterogeneity estimates therefore compare exposure–endorsement slopes over similar exposure dynamics, rather than mechanically reflecting different exposure paths across more and less secular cantons. I estimate

$$Y_{i,c,t} = (\theta_P + \beta_{P,W} W_c) X^P_{c,t} + (\theta_S + \beta_{S,W} W_c) X^S_{c,t} + (\theta_{SP} + \beta_{SP,W} W_c) X^P_{c,t} X^S_{c,t} + Z'_{c,t}\Gamma + \alpha_i + \lambda_t + \delta_c \cdot t + \varepsilon_{i,c,t}, \tag{7}$$

on the German-speaking-canton subsample, with the fixed effects, trends, and canton-year controls of the baseline specification.

Table 5 reports estimates of θ_{SP} and $\beta_{SP,W}$—the peer×state-agent interaction and its slope in W_c. Column (2) adds interactions with historical Protestant population share to address the concern that W_c proxies a longer-standing Reformed/Catholic cultural divide within German-speaking Switzerland rather than the Weberian role-separation mechanism in Corollary 1. The peer–state-agent interaction is negative at low levels of secularity support, while its slope in W_c is positive and significant in both specifications. Adding the Protestant-share interactions does not attenuate the architecture result; if anything, it strengthens it. The estimated slope of the peer–state-agent interaction in W_c more than doubles, from 1.433 to 3.100, and becomes more precisely estimated. This pattern is difficult to reconcile with the view that the secularity proxy merely captures a Protestant cultural background. Instead, the estimates are consistent with the mechanism. As support for state secularity rises, the state-agent channel's own exposure effect weakens, while the peer–state-agent interaction increases. The peer channel's own effect, if anything, also weakens—insignificantly without controls, significantly once the Protestant-share interactions are included. This is not an anomaly for the model but consistent with its equilibrium extension: under cross-channel fit, the state channel's specialization enters the peer channel's marginal products, inducing peers to reallocate precision toward merit and away from implementability (Remark 2.2.1). The peer channel's own within-channel bundle can therefore weaken with W_c even as the cross-

channel interaction strengthens—precisely the joint pattern in Table 5. As a leave-one-canton-out diagnostic, I re-estimate the specification excluding each German-speaking canton in turn. Appendix Figure A5 shows that the positive slope of the peer–state-agent interaction in W_c is not driven by any single canton. The coefficient remains positive in every leave-one-canton-out sample and stays close to the baseline magnitude.

Figure 2 shows that the implied peer–state-agent interaction turns positive in the most secular German-speaking cantons. It plots the implied interaction, $-0.354 + 1.433\,W_c$, together with a 90% confidence band computed by the delta method from the full coefficient covariance matrix of Column (1). Column (1) is the baseline plotted for transparency, while Column (2) produces an even steeper sign-reversal pattern. The interaction is statistically negative for $W_c \lesssim 0.19$ and statistically positive for $W_c \gtrsim 0.31$; the point estimate crosses zero at $W^\star \approx 0.25$. By Proposition 2, this sign reversal reveals a change in persuasion architecture: a positive peer–state-agent interaction is possible only when endorsement is assembled through cross-channel merit–implementability fit. No concave persuasion technology without merit–implementability structure can generate a positive interaction at any exposure level.

Table 5: Support for state secularity and the persuasion architecture: German-Speaking Cantons

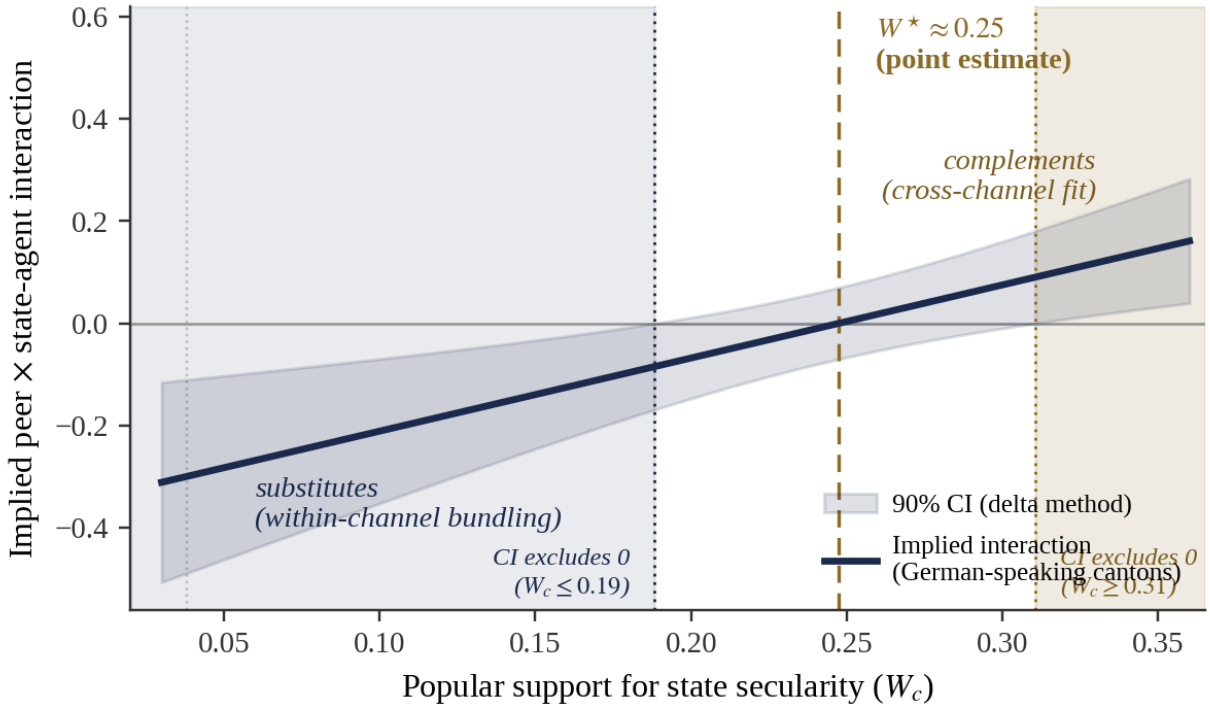
	(1)	(2)
Y_{ict}	Endorsement of institutional universalism	
X_{ct}^P	0.577*	0.582*
	(0.321)	(0.280)
X_{ct}^S	0.127*	0.139**
	(0.071)	(0.060)
$X_{ct}^P \times X_{ct}^S$	-0.354**	-0.414***
	(0.125)	(0.127)
$X_{ct}^P \times W_c$	-2.836	-3.969**
	(1.808)	(1.481)
$X_{ct}^S \times W_c$	-0.535*	-1.102***
	(0.274)	(0.212)
$X_{ct}^P \times X_{ct}^S \times W_c$	1.433**	3.100***
	(0.487)	(0.677)
Observations	4,878	4,878
Protestant share interactions control	No	Yes

Notes: Sample restricted to German-speaking cantons. W_c is the cantonal vote share in the 1980 federal popular initiative on complete church–state separation. Column (2) allows the exposure–endorsement mapping to vary with historical canton-level Protestant population share by interacting it with peer exposure, state-agent exposure, and their interaction. OLS models with individual, canton, and year fixed effects, canton-specific linear trends, time-varying canton controls, and clustering at the canton level, as in column (3) of Table 1. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Grammar of legitimacy. The second architecture shifter is the grammar of legitimacy. When public justification is understood as part of what makes the rule binding, incentives to supply persuasive reasons rise, strengthening both channels' bundles (Corollary 2). The empirical implication is therefore that each channel's own exposure effect rises. Proposition 1 guarantees that the interaction remains negative throughout, since raising $\beta(R_c)$ scales precision without altering the additively separable bundle structure. Whether the interaction becomes *more* negative as R_c rises, however, depends on the slope of each channel's own bundle in exposure.

I proxy R_c with a binary French-exposure indicator. In the republican tradition associated with Rousseau, legitimate law is not an external command imposed on individuals; it is binding because

Figure 2: Interaction as a function of secularism support (German-speaking cantons)



Notes: German-speaking-canton subsample only (18 cantons; the equation is estimated with individual, canton, and year fixed effects, canton-specific linear trends, and canton-year controls, with standard errors clustered by canton). The Figure plots the implied peer × state-agent interaction, $-0.354 + 1.433\, W_c$ (Column 1 of Table 5), with a 90% confidence band computed by the delta method from the full coefficient covariance matrix. The dotted vertical lines mark where this band excludes zero ($W_c \lesssim 0.19$ and $W_c \gtrsim 0.31$); $W^\star$ marks where the point estimate alone crosses zero.

citizens can understand it as an expression of collective self-legislation, or the general will (Rousseau (1923)). This tradition gives a distinctive role to public authority: the state is not merely an administrative instrument for implementing independently given norms, but a political form through which collectively authored norms are given binding legal form. Under the Third Republic, this state-centered grammar was institutionalized by tying the state to substantive civic and moral ends (Weber (1976))[9]. This is not a claim that the relevant mechanism operates through contemporaneous differences in formal institutions across language regions. Rather, the claim is that French-speaking environments are more exposed to a republican logic of legitimacy in which public authority is understood as the vehicle for realizing collectively justified norms. The workplace-switch design further sharpens this interpretation. Because it includes canton-by-year fixed effects, it absorbs formal institutions, local political conditions, and discourse environments common to all individuals in a canton-year. Moreover, the heterogeneity is estimated using individual French-speaking identity interacted with the public-employment switch. The design therefore asks whether the same increase in exposure to state agents has a different persuasive effect for French-speaking individuals, holding fixed the contemporaneous canton-year environment. Because the workplace-switch design excludes mission-oriented industries—health, education, and social services—its French heterogeneity is identified from exposure to *ordinary* state agents, not from the sector through which the republican grammar was historically institutionalized. A larger switch effect among French speakers therefore cannot reflect exposure to a special class of merit-supplying state agents; it indicates that the grammar operates as a feature of the receiving culture, raising the persuasive yield of exposure to the ordinary state channel. The content-weighted exposure design contributes the complementary object: it is the only design in which both channels' exposures enter, and therefore the only one that can estimate how the peer–state-agent interaction varies with R_c—the margin on which Corollary 2's architecture implications bear. In the workplace-switch design, exposure to the grammar is measured at the individual level, $R_i = 1$ for French-speaking individuals and 0 for German-speaking individuals. In the content-weighted exposure design, $R_c = 1$ for French-speaking cantons and 0 for German-speaking cantons[10].

In the workplace-switch design, where exposure to state agents varies directly at the individual-year level, the public-employment effect is about three times larger for French-speaking workers than for German-speaking workers (Table 2, Column (7)), roughly 21 percentage points compared with 7 percentage points. Interestingly, the adjacent-outcome estimates for the French subsample show that the switch moves neither ideology, redistribution and welfare preferences, army or environmental attitudes, nor satisfaction with democracy, and trust in the federal government. What does move, alongside institutional universalism, is its closest cognate: support for joining the European Union—universalism applied at the supranational level, equal treatment and open institutions extended beyond the national in-group. Among French-speaking switchers, the switch raises support for EU membership by 23 percentage points, on a subsample in which the within-person switch coefficient is significant for no other outcome (Table 6). The joint pattern is a content signature: the two universalism-cognate outcomes move together—domestic equal treatment by about 20 percentage points, supranational openness by 23—and both are roughly three times larger than in the full

[9]The Third Republic's vehicle for this institutionalization was mass schooling: the republican schoolteacher—the Republic's "hussard noir"—was, in the model's terms, a state agent whose role conception licensed the supply of merit arguments, the inversion of the Weberian conception (Weber (1976)).

[10]The heterogeneity exercise focuses on the French–German contrast rather than a French versus non-French comparison. The mechanism concerns exposure to a French republican logic of state legitimacy; German-speaking Switzerland therefore provides the cleanest benchmark within the same federal state, as well as the largest comparison group. Pooling Italian- and Romansh-speaking regions with German-speaking Switzerland would create a heterogeneous residual category. Cantonal language shares are strongly bimodal: cantons classified as French- or German-speaking are predominantly monolingual, and I exclude the bilingual cantons from the estimation sample.

switcher sample, while the seven non-cognate outcomes are flat in both samples. Content-specific persuasion predicts precisely this shape.

Table 6: French-speaking individuals: workplace switch and adjacent outcomes

	(1) Welfare	(2) Redistri.	(3) Env. vs growth	(4) Army	(5) Ideology	(6) Trust fed. gov.	(7) Sat. dem.	(8) Join EU
$State_{it}$	0.150	-0.207	-0.014	-0.152	0.221	-0.622	0.241	0.226**
	(0.225)	(0.165)	(0.074)	(0.186)	(0.348)	(0.532)	(0.445)	(0.089)
Fed_{it}	-0.351	0.220	0.166	0.482**	0.394	0.687	-0.430	-0.106
	(0.281)	(0.201)	(0.163)	(0.242)	(0.556)	(0.630)	(0.605)	(0.102)
Observations	1,516	1,516	1,030	1,504	1,233	1,544	1,531	1,408
Number of idpers	571	571	473	567	503	575	572	556

Notes: Sample: French-speaking private-to-public switchers and stayers, non-HES. All specifications include individual fixed effects and canton×year fixed effects. Standard errors are clustered at the individual level. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

In the content-weighted exposure design, French-speaking environments strengthen the marginal effect of peer exposure and shift the peer–state-agent interaction toward greater substitution (Table 7, Column (1)). The benchmark rows make the contrast concrete: the interaction is small and insignificant in German-speaking cantons and significantly more negative in French-speaking ones—the substitution pattern is concentrated where the republican grammar prevails. The state-agent heterogeneity is smaller in this design, and two non-exclusive explanations are available. The first is attenuation: the aggregate canton-level reach proxy measures individual exposure with error, compressing heterogeneity estimates; the switch design, where exposure varies at the individual level, measures this margin directly and finds it about three times larger. The second is confounding from region-correlated moderators: any predetermined canton characteristic that both differs across language regions and shifts the exposure–endorsement mapping loads onto the French interactions when omitted. Consistent with the second, the state-agent interaction rises monotonically as these region-correlated interaction sets are added, and becomes positive and significant once the mapping is allowed to vary with predetermined cantonal ideology (Table 7, Column (3)): once regional differences are absorbed, the exposure design shows both channels' own effects strengthening with R_c, corroborating Corollary 2. Across these specifications, the French-speaking interaction with peer exposure remains positive and the triple interaction remains negative and significant (Columns (2)–(3)). Furthermore, when controlling for support for secularity, the French triple interaction is larger in magnitude (Column (2)), consistent with the two shifters moving the architecture in opposite directions. Independent evidence supports reading the residual French–German contrast as cultural rather than as unobserved canton heterogeneity. Exploiting discontinuities at the Swiss language border—where formal institutions are held fixed and gradually varying unobservables are balanced by design—Athias and Wicht (2025) document that French-speaking municipalities both choose in-house public provision more often than adjacent German-speaking municipalities and operate it relatively more efficiently. This joint pattern is, moreover, the one the grammar-of-legitimacy mechanism predicts: where public authority is understood as the vehicle for realizing collectively justified norms, public provision is both more readily endorsed and more effective—an instance of materially similar institutions performing differently across cultural environments (Moscona et al. (2026)), here through the endorsement channel of Assumption 2. By Proposition 1, the interaction remains negative at every value of R_c; whether it becomes *more* negative as R_c rises is not pinned down by the model, since it depends on where the estimation sample's cantons sit on the S-shaped bundle-completion curve. The estimated deepening is therefore consistent with, rather than uniquely implied by, Corollary 2: it indicates that these cantons lie on the accelerating segment of the curve

over the observed exposure range.

Table 7: Exposure to the French grammar of legitimacy: heterogeneity

Y_{ict}	(1)	(2)	(3)
	Endorsement of institutional universalism		
X^P_{ct}	0.203	0.064	0.154
	(0.125)	(0.162)	(0.136)
X^S_{ct}	0.017	0.011	0.021
	(0.012)	(0.014)	(0.019)
$X^P_{ct} \times X^S_{ct}$	-0.037	-0.044	-0.035
	(0.031)	(0.037)	(0.078)
$X^P_{ct} \times R_c$	0.399*	0.520*	0.407**
	(0.221)	(0.285)	(0.187)
$X^S_{ct} \times R_c$	0.018	0.034	0.042*
	(0.019)	(0.029)	(0.022)
$X^P_{ct} \times X^S_{ct} \times R_c$	-0.120*	-0.183***	-0.189***
	(0.059)	(0.058)	(0.054)
Secularity support interactions	No	Yes	No
Predetermined ideology interactions	No	No	Yes
Observations	6,586	6,586	6,576

Notes: R_c is an indicator equal to one for French-speaking cantons. Column (1) reports the baseline French-exposure heterogeneity specification. Column (2) allows the exposure–endorsement mapping to vary with canton-level support for state secularity by interacting referendum support for church–state separation with peer exposure, state-agent exposure, and their interaction. Column (3) analogously allows the exposure–endorsement mapping to vary with cantonal ideology measured in the initial sample period. Secularity support and predetermined cantonal ideology are demeaned in the estimation sample of Column (1); main effects in Columns (2)–(3) are therefore evaluated at the sample mean of the added moderator. All specifications include the fixed effects and controls from the baseline specification. Standard errors are clustered at the canton level. $^*p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

The two designs therefore point the same way: the within-channel bundling pattern is concentrated in French-speaking environments, consistent with the republican-grammar mechanism. This interpretation also helps qualify the descriptive pattern reported in Section 4: endorsement of institutional universalism is higher among French-speaking respondents. The French-exposure results should not be read as merely reflecting a higher baseline attachment to universalistic rules. They indicate that exposure itself maps differently into endorsement in French-speaking environments. This is more consistent with a republican grammar in which public authority is judged by its capacity to realize collectively justified norms. Such a grammar can make universalistic rules more persuasive while also making existing institutions more exposed to criticism when they fall short of those rules, a tension visible in both designs' state-attachment estimates, and of modern public self-justification (Pippin (1999)).

7 Conclusion

Institutions are legitimate only if the norms they realize are recognized as rationally justifiable. This paper studies how such recognition is formed. Citizens must encounter public reasons about both the normative merit of a candidate norm and the institutional means through which it can be realized. Whether the two kinds of reason travel together, however, is not fixed. Two features of the cultural environment discipline the persuasion architecture. Where the surrounding culture is Weberian, meaning that it expects a strict separation between technocratic administration and normative authority, state agents internalize a legal-rational role conception that makes implementability arguments cheap and merit arguments costly for them to supply. Where the prevailing grammar

of legitimacy ties a rule's bindingness to citizens' own recognition of it as justified, both channels' incentives to supply persuasive reasons rise. The equilibrium architecture is asymmetric in what it can produce: under within-channel bundling, peer and state-agent exposures are substitutes at every exposure level, while a positive interaction is possible only when complete justificatory bundles are assembled across channels.

The empirical evidence delivers the model's building blocks and its architecture predictions. Content-weighted exposure to peers and to state agents each raises endorsement of institutional universalism, establishing the two channels as distinct sources of influence. Their interaction is negative on average, and, within non-mission industries, moving from private to public employment raises endorsement by about ten percentage points within the same canton-year environment. This switch effect is concentrated among cantonal and communal rather than federal state agents, suggesting that the same role-conception mechanism operates across institutional tiers. Adjacent-outcome tests license reading these effects as persuasion: the estimates are inconsistent with generic ideological diffusion, socialization into a public-sector role, and self-interest as a public employee, while displaying content-specific signatures. Peer exposure makes equal treatment compatible with limits on public provision, and state-agent exposure raises endorsement while sharpening critical evaluations of democratic performance. The architecture tests then show that the two cultural margins move the persuasion architecture in opposite directions. Where popular support for state secularity is stronger, within German-speaking cantons, the state-agent channel's own effect weakens and the peer–state-agent interaction reverses sign, turning significantly positive in the most secular cantons, a reversal that no concave persuasion technology without merit–implementability structure can generate. Where exposure to the French republican grammar of legitimacy is stronger, both channels' effects strengthen, the interaction deepens, and the workplace-switch effect is about three times larger, with movement concentrated in universalism-cognate content.

The broader implication concerns where legitimacy is formed. Institutions shape endorsement by changing incentives, constraints, and formal rights; but the recognition of a norm as rationally justifiable is formed through public justification, in the public-reason environment, and the architecture of that environment—who supplies the normative *why*, who the institutional *how*, and whether the two meet—is disciplined by the cultural primitives this paper identifies rather than by institutional design alone. This is why democratic authorization is not a sufficient foundation: authorization makes public justification consequential, but the architecture through which merit and implementability reach citizens lies beyond its reach. That architecture carries its own normative stakes. A deep Weberian culture produces, in equilibrium, a two-sided specialization in which the state-agent channel supplies the institutional *how* and the normative *why* is concentrated in peer discourse. Specialization is not itself a pathology—it is a division of justificatory labor—but it removes the redundancy of within-channel bundling: the quality of legitimacy formation then rests on a single margin, the contestability and integration of peer-side merit discourse, and where that margin fails, implementability precision amplifies whichever merit claims prevail.

Culture, on this account, is not only an inherited stock of values or a set of stable preferences: it is the environment that configures the persuasion architecture, and it is what that architecture, over time, transforms. Cultural change is thus an endogenous process of norm legitimation, whose pace and direction reflect an architecture shaped by the same cultural environment it changes. Whether such an architecture can exist at the scale that contemporary norms require is a first-order question.

References

Acemoglu, Daron and James A. Robinson (2012) *Why Nations Fail: The Origins of Power, Prosperity, and Poverty*, New York: Crown Publishing Group.

——— (2025) "Culture, Institutions, and Social Equilibria: A Framework," *Journal of Economic Literature*, 63 (2), 637–92, 10.1257/jel.20241680.

Adena, Maja, Ruben Enikolopov, Maria Petrova, Veronica Santarosa, and Ekaterina Zhuravskaya (2015) "Radio and the Rise of The Nazis in Prewar Germany," *The Quarterly Journal of Economics*, 130 (4), 1885–1939, 10.1093/qje/qjv030.

Aneja, Abhay and Guo Xu (2024) "Strengthening State Capacity: Civil Service Reform and Public Sector Performance during the Gilded Age," *American Economic Review*, 114 (8), 2352–87, 10.1257/aer.20230019.

Arendt, Hannah (1963) *Eichmann in Jerusalem: A Report on the Banality of Evil*, New York: The Viking Press.

Arrow, K. (1951) *Social Choice and Individual Values*: Cowles Foundation.

Athias, Laure and Bruno Ventelou (2026) "The Heart Has Reasons That Reason Cannot Understand: Private vs Institutional Universalism," working paper.

Athias, Laure and Pascal Wicht (2025) "Make or buy for public services: Culture matters for efficiency considerations," *The Quarterly Review of Economics and Finance*, 99, 101899, https://doi.org/10.1016/j.qref.2024.101899.

Audier, F. and M. Bacache-Beauvallet (2007) "Emploi dans la fonction publique et fonctions "d'intérêt public": Que nous apprennent les comparaisons internationales ?" *Revue de l'OFCE*, 103, 323–350, 10.3917/reof.103.0323.

Berman, Sheri (1997) "Civil Society and the Collapse of the Weimar Republic," *World Politics*, 49 (3), 401–429, 10.1353/wp.1997.0008.

Besley, Timothy (2020) "State Capacity, Reciprocity, and the Social Contract," *Econometrica*, 88 (4), 1307–1335, https://doi.org/10.3982/ECTA16863.

Besley, Timothy, Robin Burgess, Adnan Khan, and Guo Xu (2022) "Bureaucracy and Development," *Annual Review of Economics*, 14 (Volume 14, 2022), 397–424, https://doi.org/10.1146/annurev-economics-080521-011950.

Besley, Timothy and Torsten Persson (2011) *Pillars of Prosperity*, Princeton: Princeton University Press, doi:10.1515/9781400840526.

——— (2019) "Democratic Values and Institutions," *American Economic Review: Insights*, 1 (1), 59–76, 10.1257/aeri.20180248.

Boxell, Levi, Matthew Gentzkow, and Jesse M. Shapiro (2024) "Cross-Country Trends in Affective Polarization," *The Review of Economics and Statistics*, 106 (2), 557–565, 10.1162/rest_a_01160.

Bursztyn, Leonardo, Georgy Egorov, and Stefano Fiorin (2020) "From Extreme to Mainstream: The Erosion of Social Norms," *American Economic Review*, 110 (11), 3522–48, 10.1257/aer.20171175.

Castoriadis, Cornelius (1987) *The Imaginary Institution of Society*, Cambridge, Massachusetts: The MIT Press.

——— (1996) *La montée de l'insignifiance: Les carrefours du labyrinthe IV*, Paris: Éditions du Seuil.

Condorcet, Marie-Jean-Antoine-Nicolas de Caritat (1785) *Essay on the Application of Analysis to the Probability of Majority Decisions (Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix)*: Paris.

Darnton, Robert (1982) *The Literary Underground of the Old Regime*: Harvard University Press.

Eidgenössisches Aktionskomitee für die Trennung von Staat und Kirche (1978) "Zur Trennung von Staat und Kirche in der Schweiz: Stellungnahme des Initiativkomitees zuhanden der vorberatenden Kommission des Nationalrates,"Technical report, Eidgenössisches Aktionskomitee für die Trennung von Staat und Kirche, Bern, Campaign document, Swissvotes vote no. 299.

Ekeh, Peter P. (1975) "Colonialism and the Two Publics in Africa: A Theoretical Statement," *Comparative Studies in Society and History*, 17 (1), 91–112.

Enke, Benjamin, Ricardo Rodriguez-Padilla, and Florian Zimmermann (2022) "Moral Universalism: Measurement and Economic Relevance," *Management Science*, 68 (5), 3175–3973, https://doi.org/10.1287/mnsc.2021.4086.

Esping-Andersen, Gosta (1990) *The three worlds of welfare capitalism*, Princeton (NJ): Princeton University Press.

Eugster, Beatrix, Rafael Lalive, Andreas Steinhauer, and Josef Zweimüller (2011) "The Demand for Social Insurance: Does Culture Matter?" *The Economic Journal*, 121 (556), F413–F448, https://doi.org/10.1111/j.1468-0297.2011.02479.x.

Eugster, Beatrix and Raphaël Parchet (2019) "Culture and Taxes," *Journal of Political Economy*, 127 (1), 296–337, 10.1086/700760.

Francois, Patrick (2003) "Not-for-Profit Provision of Public Services," *The Economic Journal*, 113 (486), C53–C61.

Gentzkow, Matthew and Emir Kamenica (2017) "Competition in Persuasion," *The Review of Economic Studies*, 84 (1 (298)), 300–322.

Gratton, Gabriele and Barton E Lee (2024) "Liberty, Security, and Accountability: The Rise and Fall of Illiberal Democracies," *The Review of Economic Studies*, 91 (1), 340–371, 10.1093/restud/rdad030.

Gregg, Paul, Paul Grout, Anita Ratcliffe, Sarah Smith, and Frank Windmeijer (2011) "How important is pro-social behaviour in the delivery of public services?" *Journal of Public Economics*, 95 (7), 758–766.

Greif, Avner (1994) "Cultural Beliefs and the Organization of Society: A Historical and Theoretical Reflection on Collectivist and Individualist Societies," *Journal of Political Economy*, 102 (5), 912–950.

Habermas, Jürgen (1990) *Moral Consciousness and Communicative Action*, Cambridge, MA: MIT Press.

Habermas, Jürgen (1989) *The Structural Transformation of the Public Sphere: An Inquiry into a Category of Bourgeois Society*: MIT Press.

Harsanyi, John C. (1955) "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility," *Journal of Political Economy*, 63 (4), 309–321.

Heldring, Leander (2026) "The cost of state-building: Evidence from Germany," working paper.

Henrich, Joseph (2020) *The WEIRDest People in the World: How the West Became Psychologically Peculiar and Particularly Prosperous*, New York: Farrar, Strauss, and Giroux.

Hinterleitner, Markus (2025) *Blaming Bureaucracy: Reckoning with a Problematic Political Activity*: Oxford University Press, 10.1093/9780198980186.001.0001.

Horkheimer, Max and Theodor W. Adorno (1944) *Dialektik der Aufklärung: Philosophische Fragmente*, Amsterdam: Querido Verlag.

Hosli, Madeleine O., Luiza Martins Santos, and Henry Boeree (2025) "A political-economic analysis of Swiss referendums 1848 to 2022: turnout, acceptance rates and the double-majority threshold," *Constitutional Political Economy*, 36 (2), 245–270, 10.1007/s10602-025-09468-1.

Kamenica, Emir and Matthew Gentzkow (2011) "Bayesian Persuasion," *American Economic Review*, 101 (6), 2590–2615, 10.1257/aer.101.6.2590.

Kohn, Hans (1950) "Romanticism and the Rise of German Nationalism," *The Review of Politics*, 12 (4), 443–472.

Kuran, Timur (1991) "The East European Revolution of 1989: Is it Surprising that We Were Surprised?" *The American Economic Review*, 81 (2), 121–125.

——— (1995) *Private Truths, Public Lies*: Harvard University Press.

Luo, Xizi and Melanie Meng Xue (2025) "Enlightenment Under Autocracy: The Origins of Liberalism in China," working paper.

Malinowski, Bronislaw (1922) *Argonauts Of The Western Pacific: An Account of Native Enterprise and Adventure in the Archipelagoes of Melanesian New Guinea*, London: Routledge & Kegan Paul.

Mokyr, Joel (2010) *The Enlightened Economy*: Yale University Press.

——— (2016) *A Culture of Growth*: Princeton University Press.

Moscona, Jacob, Nathan Nunn, and James A Robinson (2026) "Searching for Fish in Trees? Economic Development when Context Matters," Working Paper 34810, National Bureau of Economic Research, 10.3386/w34810.

Mosse, George L. (1964) *The Crisis of German Ideology: Intellectual Origins of the Third Reich*, New York: Grosset and Dunlap.

North, D. C. (1990) *Institutions, Institutional Change and Economic Performance*, Cambridge: Cambridge UniversityPress.

Pippin, Robert B. (1999) *Modernism as a Philosophical Problem: On the Dissatisfactions of European High Culture*: Wiley-Blackwell.

——— (2008) *Hegel's Practical Philosophy: Rational Agency as Ethical Life*, Cambridge: Cambridge University Press.

Popkin, Jeremy D. (1990) *News and Politics in the Age of Revolution: Jean Louis Carrier and the Révolutions de Paris, 1789–1797*: Cornell University Press.

Portmann, Marco, Frederik Blümel, and Christoph A. Schaltegger (2024) "Lohnprämie in den öffentlichen Verwaltungen: Neue Analyse 2024," Policy Papers 18, IWP.

Putnam, Robert D. (2000) *Bowling Alone:The Collapse and Revival of American Community*, New York: Simon & Schuster.

Putnam, Robert D., Robert Leonardi, and Raffaella Y. Nonetti (1993) *Making Democracy Work: Civic Traditions in Modern Italy*: Princeton University Press.

Rawls, John (1971) *A Theory of Justice*, Cambridge, MA: Harvard University Press.

——— (1997) "The Idea of Public Reason Revisited," *University of Chicago Law Review*, 64, 765–807.

Rousseau, Jean-Jacques (1923) *The Social Contract and Discourses*, (G. D. H. Cole, Trans.). London and Toronto: J. M. Dent & Sons, Original work published 1762.

Sahlins, Marshall (1976) *Culture and Practical Reason*, Chicago: The University of Chicago Press.

——— (2000) *Culture in Practice: Selected Essays*, New York: Zone Books.

Schweizerischer Bundesrat (1978) "Botschaft über die Volksinitiative betreffend die vollständige Trennung von Staat und Kirche,"Technical report, Schweizerische Eidgenossenschaft, Bern, Bundesblatt 1978 II 665–698.

Sieyès, Emmanuel-Joseph (1789) *Qu'est-ce que le Tiers-État?*: Paris.

Smith, Adam (1776) *An Inquiry into the Nature and Causes of the Wealth of Nations*, London: W. Strahan and T. Cadell.

Tabellini, Guido (2008) "Institutions and Culture," *Journal of the European Economic Association*, 6 (2-3), 255–294, 10.1162/JEEA.2008.6.2-3.255.

Taylor, Charles (2004) *Modern Social Imaginaries*, Durham and London: Duke University Press.

——— (2007) *A Secular Age*, Cambridge, Massachusetts, and London, England: The Belknap Press of Harvard University Press.

Tilly, C. (1975) *Reflections on the History of European State-Making*, Princeton: Princeton University Press.

——— (1985) *War Making and State Making as Organized Crime*, Cambridge: Cambridge University.

Tyler, Tom R. (2006) *Why People Obey the Law*, Princeton, New Jersey: Princeton University Press.

Vatter, Adrian, Tobias Arnold, Alexander Arens et al. (2024) "Patterns of Democracy in the Swiss Cantons, 1979–2023," Dataset. University of Bern, Institute of Political Science.

Weber, Eugen (1976) *Peasants into Frenchmen: The Modernization of Rural France, 1870-1914*, Stanford, Calif.: Stanford University Press.

Weber, Max (1922) *Economy and Society*, 4th ed. Tübingen.

Zucman, Gabriel (2024) "A Blueprint for a Coordinated Minimum Effective Taxation Standard for Ultra-High-Net-Worth Individuals,"Technical report, commissioned by the Brazilian G20 Presidency.

A Appendix

A.1 Proof of the comparative statics in Remark 2.2.1

Fix (c,t) and suppress subscripts. Under the specialization extension, the peer channel's objective is

$$\max_{\pi_P^m,\pi_P^h\geq 0}\ \beta(R_c)\,B\,u(\Lambda)-K_P^m(\pi_P^m)-K_P^h(\pi_P^h),\qquad \Lambda=\Lambda_S+\Lambda_P+\Lambda_\times,$$

with $\Lambda_\times=f(Q_S^m)f(Q_P^h)+f(Q_S^h)f(Q_P^m)$, $f(x)=1-e^{-x}$, $Q_j^k=\pi_j^k E^j$. Interior first-order conditions:

$$G^m\equiv\beta B\,u'(\Lambda)\,E^P e^{-Q_P^m}\left[f(Q_P^h)+f(Q_S^h)\right]-K_P^{m\prime}(\pi_P^m)=0, \tag{8}$$

$$G^h\equiv\beta B\,u'(\Lambda)\,E^P e^{-Q_P^h}\left[f(Q_P^m)+f(Q_S^m)\right]-K_P^{h\prime}(\pi_P^h)=0. \tag{9}$$

Treat $(\pi_S^m(W_c),\pi_S^h(W_c))$ as given functions of W_c with $d\pi_S^h/dW_c\geq 0$, $d\pi_S^m/dW_c\leq 0$ (Corollary 1(i)). Totally differentiating (8)–(9) in W_c gives

$$J\begin{pmatrix}d\pi_P^m/dW_c\\ d\pi_P^h/dW_c\end{pmatrix}=-\begin{pmatrix}\partial G^m/\partial W_c\\ \partial G^h/\partial W_c\end{pmatrix},\qquad J\equiv\begin{pmatrix}\partial G^m/\partial\pi_P^m & \partial G^m/\partial\pi_P^h\\ \partial G^h/\partial\pi_P^m & \partial G^h/\partial\pi_P^h\end{pmatrix}.$$

Direct effects. Differentiating the bracketed terms:

$$\frac{\partial G^m}{\partial W_c}\Big|_{u'\text{ fixed}}=\beta B\,u'\,E^P e^{-Q_P^m}\,f'(Q_S^h)\,E^S\,\frac{d\pi_S^h}{dW_c}\ \geq\ 0,\qquad \frac{\partial G^h}{\partial W_c}\Big|_{u'\text{ fixed}}=\beta B\,u'\,E^P e^{-Q_P^h}\,f'(Q_S^m)\,E^S\,\frac{d\pi_S^m}{dW_c}\ \leq\ 0,$$

since $f'>0$. *Feedback effect.* The full derivatives add the common-multiplier term

$$\beta B\,u''(\Lambda)\,\frac{d\Lambda}{dW_c}\times\begin{cases}E^P e^{-Q_P^m}[f(Q_P^h)+f(Q_S^h)]\\ E^P e^{-Q_P^h}[f(Q_P^m)+f(Q_S^m)],\end{cases}$$

whose sign is that of $-\,d\Lambda/dW_c$, ambiguous because $d\Lambda_S/dW_c\leq 0$ (Corollary 1(ii)) while the two terms of $d\Lambda_\times/dW_c$ move oppositely.

Signing the system. At an interior optimum the second-order conditions require J to be negative semidefinite; I assume they hold strictly—guaranteed, for instance, by sufficiently convex costs $K_P^{k\prime\prime}$—so that $\det J>0$ and the diagonal entries are negative. The diagonal entries are negative for any $K_P^{k\prime\prime}\geq 0$, since the own-second derivatives of $\Lambda_P+\Lambda_\times$ in each π_P^k are negative. The off-diagonal entries carry the within-channel complementarity term $u'\,\partial^2\Lambda_P/\partial\pi_P^m\partial\pi_P^h>0$ and the feedback term $u''\,(\partial\Lambda/\partial\pi_P^m)(\partial\Lambda/\partial\pi_P^h)<0$. A sufficient condition for the stated signs is that the direct effects dominate the feedback and complementarity terms; in particular, if u is locally linear ($u''=0$ in a neighborhood of the optimum), J is diagonal up to the complementarity term, and Cramer's rule gives

$$\operatorname{sign}\frac{d\pi_P^m}{dW_c}=\operatorname{sign}\frac{\partial G^m}{\partial W_c}\geq 0,\qquad \operatorname{sign}\frac{d\pi_P^h}{dW_c}=\operatorname{sign}\frac{\partial G^h}{\partial W_c}\leq 0,$$

provided the within-channel complementarity term does not reverse either inequality, for which a sufficient condition is $|u''|\,\big|\frac{d\Lambda}{dW_c}\big|$ and $u'\,\partial^2\Lambda_P/\partial\pi_P^m\partial\pi_P^h$ small relative to the direct effects, i.e., the shock transmits chiefly through the cross-channel pairing terms. Under these conditions, peers reallocate precision toward merit and away from implementability as W_c rises, as stated in Remark 2.2.1. ■

A.2 Figures and Tables

Table A1: Model notation

Object	Definition / interpretation
Indices and sets	
i	Citizen (agent).
(c,t)	Public-sphere context c at date t.
$j \in \{S,P\}$	Communication channel: S = state-agent channel; P = peer channel.
$k \in \{m,h\}$	Content dimension: m = merit/"why"; h = implementation/"how".
Latent states, priors, and signals	
m,h	Latent "merit" and "implementation" states relevant for persuasion.
μ_k, σ_k^2	Prior mean and variance for dimension $k \in \{m,h\}$.
$\tau_k \equiv \sigma_k^{-2}$	Prior precision for dimension k.
$\tilde{x}^k_{j,i,c,t}$	Signal received by i from channel j about dimension k in (c,t).
$\mathcal{F}_{i,c,t}$	Citizen i's information set in (c,t): $\mathcal{F}_{i,c,t} = \sigma\big(E^S_{i,c,t}, E^P_{i,c,t}, \tilde{x}^m_{S,i,c,t}, \tilde{x}^h_{S,i,c,t}, \tilde{x}^m_{P,i,c,t}, \tilde{x}^h_{P,i,c,t}\big)$.
$\sigma(\cdot)$	Sigma-algebra generated by the listed random variables (information available to the agent).
Exposure / reach	
$E^j_{i,c,t}$	Realized exposure of citizen i to channel j in (c,t) (nonnegative).
$E^j_{c,t}$	Aggregate reach (average exposure intensity) of channel j in public sphere (c,t).
$\eta^j_{i,c,t}$	Idiosyncratic reach component: $E^j_{i,c,t} = E^j_{c,t} + \eta^j_{i,c,t}$, $\mathbb{E}[\eta^j_{i,c,t} \mid c,t] = 0$.
Precision and effective precision	
e^k_j	Producer effort by channel j to generate precision on dimension k.
$v^k_j(\cdot)$	Precision technology: $\pi^k_j = v^k_j(e^k_j)$ (increasing, concave).
π^k_j	Precision supplied by channel j on dimension k.
Q^k_j	Effective precision / intensity from channel j on dimension k: $Q^k_j \equiv \pi^k_j E^j$ (suppressing indices when convenient).
$Q_{k,i,c,t}$	Total effective precision for citizen i on dimension k in (c,t): $Q_{k,i,c,t} \equiv \sum_{j\in\{S,P\}} \pi^k_j E^j_{i,c,t}$.
Bundling and endorsement	
$f(x)$	Poisson "arrival" map: $f(x) \equiv 1 - e^{-x}$.
Λ_j	Within-channel bundle intensity: $\Lambda_j \equiv f(Q^m_j)\, f(Q^h_j)$.
$\Lambda_\times$	Cross-channel fit (extension): $\Lambda_\times \equiv f(Q^m_S) f(Q^h_P) + f(Q^h_S) f(Q^m_P)$.
Λ	Total bundle intensity: $\Lambda \equiv \Lambda_S + \Lambda_P$ (benchmark) or $\Lambda \equiv \Lambda_S + \Lambda_P + \Lambda_\times$ (extension).
s	Endorsement probability / share: $s = u(\Lambda)$.
$u(\cdot)$	Endorsement link, increasing and concave ($u' > 0$, $u'' \le 0$); canonical case $u(\Lambda) = 1 - e^{-\Lambda}$.
κ	Off-diagonal "fit" index (local complementarity at the origin): $\kappa \equiv \pi^m_S \pi^h_P + \pi^h_S \pi^m_P$.
Producer payoffs and costs	
$C^k_j(\cdot)$	Cost of effort e^k_j for channel j on dimension k (convex).
B	Social benefit scale from endorsement/legitimacy ($B > 0$).
$\beta(R_c)$	Internalization weight: channels internalize fraction $\beta(R_c) \in (0,1]$ of B, where R_c is the predetermined, public-sphere level, grammar of legitimacy.
$\delta_h(W_c)$,$\delta_m(W_c)$	Comparative-advantage wedges, where W_c is the predetermined, public-sphere level, Weberian culture W_c (Assumption 1).
Dynamics (when used)	
S_t	Aggregate endorsement at date t (e.g., population average of s).
$E^j(S)$	Exposure map: reach as a function of current endorsement S (public-sphere propagation).
$\Phi(S)$	Update map: $\Phi(S) \equiv u(\Lambda(S))$ so $S_{t+1} = \Phi(S_t)$.

Table A2: Variable Descriptions and Sources

Table	Variable	Definition	Coding	Source	Time period
1,2,5,6	Institutional universalism	Are you in favour of Switzerland offering foreigners the same opportunities as those offered to Swiss citizens, or in favour of Switzerland offering Swiss citizens better opportunities?	0 = better opportunities for Swiss citizens and 1 = equal opportunities	Swiss Household Panel	1999–2003
3,4	Welfare	Are you in favour of a diminution or in favour of an increase of the Confederation social spendings?	1 = diminution of social spending, 2 = no change and 3 = more social spending	Swiss Household Panel	1999–2003
3,4	Redistri.	Are you in favour of an increase or in favour of a decrease of the tax on high incomes?	1 = decrease, 2 = no change and 3 = increase	Swiss Household Panel	1999–2003
3,4	Army	Are you in favour of Switzerland having a strong army or for Switzerland not having an army?	1 = no army, 2 = neither and 3 = strong army	Swiss Household Panel	1999–2003
3,4	Env. vs growth	Are you in favour of Switzerland being more concerned with protection of the environment than with economic growth, or in favour of Switzerland being more concerned with economic growth than with protection of the environment?	1 = more concerned with economic growth, 2 = neither and 3 = more concerned with environment protection	Swiss Household Panel	1999–2003
3,4	Join EU	Are you in favour of Switzerland joining the European Union or are you in favour of Switzerland staying outside of the European Union?	0 = staying outside of EU and 1 = joining the EU	Swiss Household Panel	1999–2003
3,4	Ideology	When they talk about politics, people mention left and right. Personally, where do you position yourself?	0 = 'left' and 10 = 'right'	Swiss Household Panel	1999–2003
3,4	Sat. Dem.	Overall, how satisfied are you with the way in which democracy works in our country?	0 = 'not at all satisfied' and 10 ='completely satisfied	Swiss Household Panel	1999–2003
3,4	Trust fed. gov.	How much confidence do you have in the Federal Government?	0 = 'no confidence' and 10 = 'full confidence	Swiss Household Panel	1999–2003
1,3,5,6	Average ideological position of the cantonal executive	Computed as the mean of seat-level party positions within the executive council. Each seat is coded on a four-point left–right scale ranging from 1 (far left) to 4 (far right); the index is obtained by summing these seat codes and dividing by the number of executive seats, so that higher values indicate a more right-leaning government.	1 = far left to 4 = far right	Vatter et al. (2024)	1999–2003
1,3,5,6	Foreign pop share	Share of the cantonal resident population holding foreign citizenship, measured as the number of non-citizens divided by the total resident population.	Percentage	State Secretariat for Migration	1999–2003

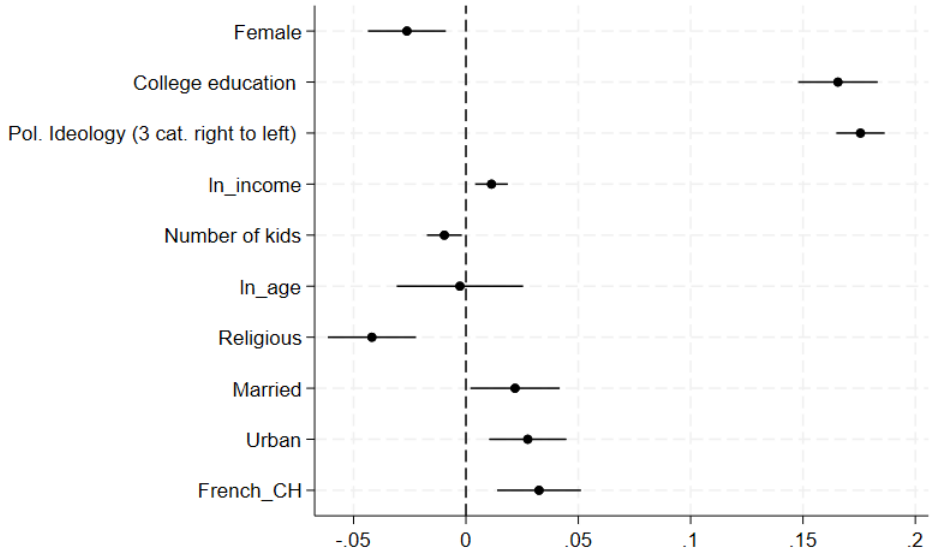
Table	Variable	Definition	Coding	Source	Time period
1,3,5,6	Referendum	Number of referenda held in canton.	Count	Vatter et al. (2024)	1999–2003
1,3,5,6	Unemployment rate	Cantonal unemployment rate in percent, measured as the share of unemployed individuals in the active population.	Percentage	State Secretariat for Economic Affairs	1999–2003
1,3,5,6	Taxable income per capita	Per-capita taxable income base, measured as the total mass of income subject to the personal income tax divided by the cantonal resident population.	Swiss francs per capita	Federal Finance Administration	1999–2003
5,6	Support for state secularity	Canton-level vote shares in the 1980 federal popular initiative “concerning the complete separation of Church and State", submitted to popular vote on 2 March 1980.	0 to 1	Swiss Federal Chancellery	1980
1,3,5,6	Public-employment share as a proxy for E_c^S used in X_{ct}^S	Percentage of a canton’s employed labor force working in the public sector measured as the number of full-time equivalent.	Percentage	Federal Statistical Office	2001
1,3,5,6	Association density as a proxy for E_c^P used in X_{ct}^P	Cantonal number of registered associations per 1,000 inhabitants.	Number per 1000 inhabitants	Federal Statistical Office	1995
1,3,5,6	$p_{c,t}^P$ used in X_{ct}^P	Cantonal share of workers in the private sector who endorse institutional universalism	0 to 1	Swiss Household Panel	1999–2003
1,3,5,6	$p_{c,t}^S$ used in X_{ct}^S	Cantonal share of workers in the public sector who endorse institutional universalism	0 to 1	Swiss Household Panel	1999–2003

Table A3: Summary Statistics

	Mean	Std. dev.	Min	Max	Between SD	Within SD
Panel A. Individual-level time-varying variables						
Institutional universalism	0.616	0.486	0.000	1.000	0.440	0.250
Peer content-weighted exposure (X_{ct}^{P})	0.000	0.165	-0.608	0.642	0.156	0.053
State-agent content-weighted exposure (X_{ct}^{S})	0.000	1.351	-4.633	5.850	1.317	0.339
Welfare	2.247	0.714	1.000	3.000	0.601	0.403
Redistribution	2.591	0.641	1.000	3.000	0.546	0.362
Army	2.189	0.794	1.000	3.000	0.709	0.369
Env. vs growth	2.295	0.712	1.000	3.000	0.590	0.417
Join EU	0.529	0.499	0.000	1.000	0.463	0.191
Ideology	4.959	2.115	0.000	10.000	1.903	0.925
Sat. w. Democracy	5.939	2.006	0.000	10.000	1.749	1.060
Trust in fed. gov.	5.804	2.244	0.000	10.000	1.987	1.092
Panel B. Cantonal-level time-varying variables						
Average ideology	2.710	0.347	1.800	3.571	–	–
Foreign pop share	0.194	0.054	0.074	0.334	–	–
Nb. of Referenda	3.767	3.432	0.000	24.000	–	–
Unemployment rate	2.368	1.077	0.300	6.500	–	–
Taxable inc. pc	32'555.02	6'591.835	22'634.02	75'379.49	–	–
Panel C. Cantonal-level pre-determined variables						
French-speaking canton	0.236	0.425	0.000	1.000	–	–
Secularity support	0.215	0.058	0.038	0.352	–	–
Public-employment share	6.161	1.532	3.500	11.100	–	–
Association density	0.899	0.190	0.444	1.316	–	–
Observations (person-years)						19,243
Individuals (idpers)						6,920
Years						1999–2003

Notes. The sample is restricted to non-working Swiss respondents aged 18+; canton non-movers. Unit of observation is the individual-year. Between SD is the standard deviation of individual means. Within SD is the standard deviation of deviations from individual means.

Figure A1: Correlates of endorsement of the universalism institutional norm



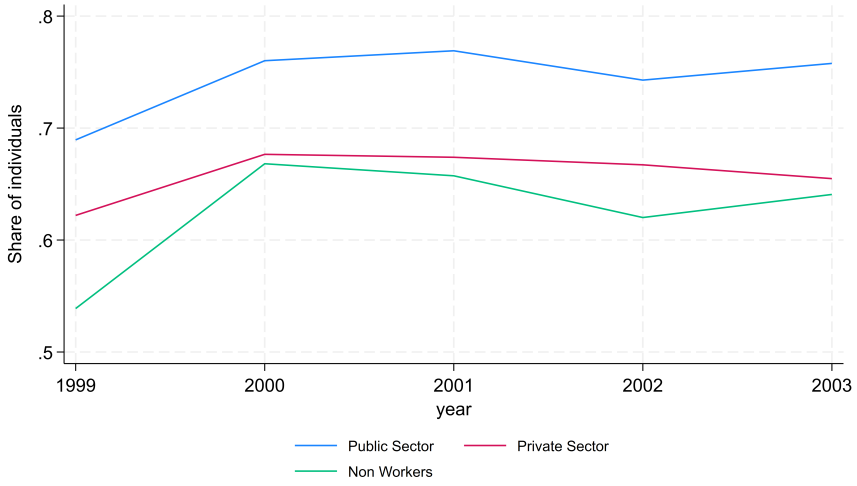
Note: This figure presents OLS Estimates from regressions of the Y_{ict} variable on all sociodemographic correlates, with SE clustered at the individual level, along with 95% confidence intervals.

Table A4: Institutional universalism endorsement by institutional sector

	Public sector		Private sector		
	(1) Mean	N	(2) Mean	N	(3) Difference
All workers	0.742	4,250	0.657	7,473	0.085***
	(0.438)		(0.475)		(0.009)
Within HES industries	0.800	2,112	0.769	924	0.030*
	(0.400)		(0.421)		(0.016)
Within non-HES industries	0.666	1,569	0.641	5,265	0.026*
	(0.472)		(0.480)		(0.014)

Note: The table reports means of Y_{ict}, an indicator for endorsement of institutional universalism. The sample includes Swiss respondents aged 18+ working in the public or private sector. Industry (HES/non-HES) is observed only for a subset of employed respondents; breakdown columns therefore exclude observations with missing industry and need not sum to the overall sample. HES denotes health, education, and social care industries; non-HES denotes all other industries. Column (3) reports mean differences with robust standard errors in parentheses. Standard deviations are in parentheses below means in Columns (1) and (2). $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Figure A2: Yearly share of individuals who endorse the universalism social norm, by group



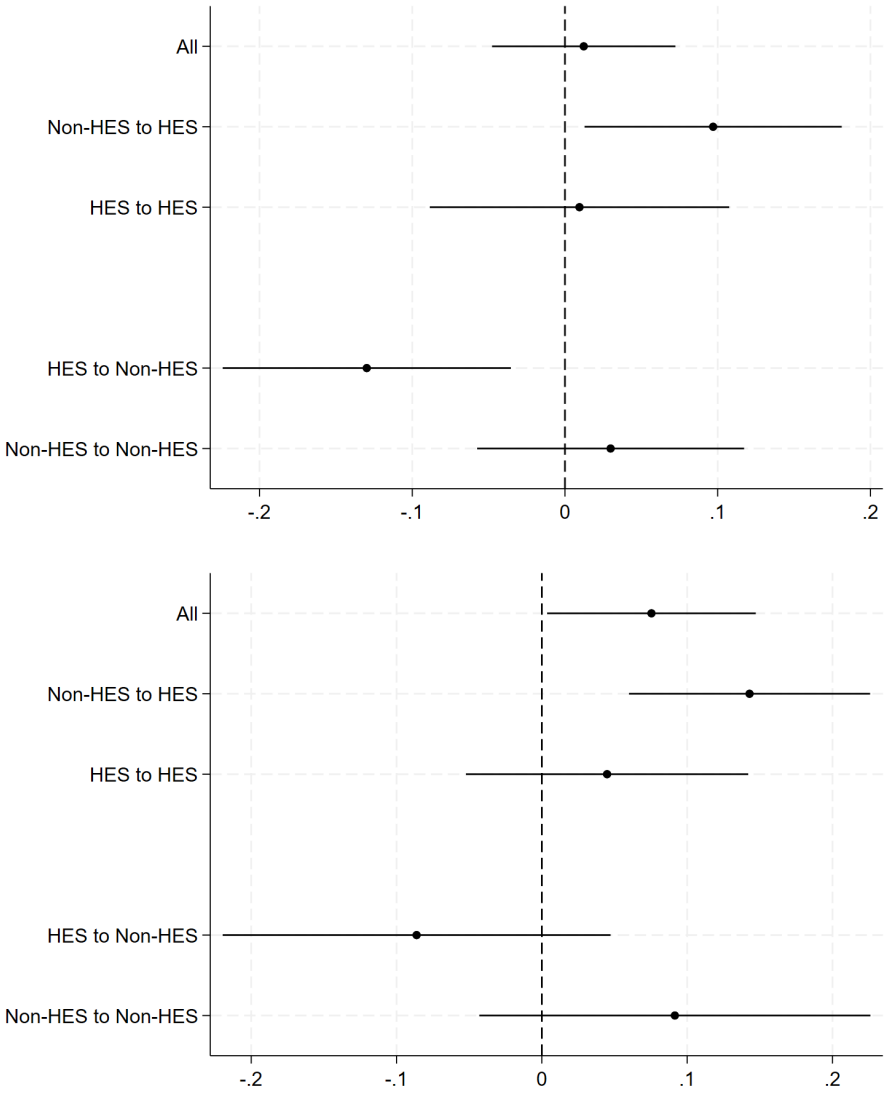
Note: The figure shows the share of Swiss individuals aged 18 and over who endorse institutional universalism over time, measured as the yearly mean of Y_{ict}, separately for non workers, state employees, and private sector workers.

Table A5: Switchers to *vs.* stayers in the public non-HES sector

Variable	(1) Switchers	(2) Stayers	(3) Difference
Share of female	0.429	0.351	0.077
	(0.497)	(0.478)	(0.054)
Share of married	0.623	0.676	-0.053
	(0.487)	(0.469)	(0.053)
Number of children	1.388	1.584	-0.197
	(1.289)	(1.310)	(0.141)
Age	38.714	40.512	-1.798*
	(9.467)	(8.666)	(1.021)
Share of Educ.: college level	0.143	0.242	-0.099**
	(0.352)	(0.428)	(0.040)
Share of living in urban area	0.612	0.574	0.038
	(0.490)	(0.495)	(0.053)
Ln yearly income	10.924	11.051	-0.127*
	(0.612)	(0.602)	(0.070)
State Universalism	0.679	0.729	-0.050
	(0.470)	(0.445)	(0.056)
Ideology	4.273	4.540	-0.268
	(2.263)	(2.107)	(0.258)

Notes: The table displays socio-demographic characteristics of switchers from the private into the public sector and stayers in the public sector within non-HES industries. The last column presents the difference in the mean value of each variable between the two sectors. Columns (1) and (2) standard deviation in parentheses. Columns (3) robust standard errors in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Figure A3: Self-selection based on endorsement of institutional universalism across industries



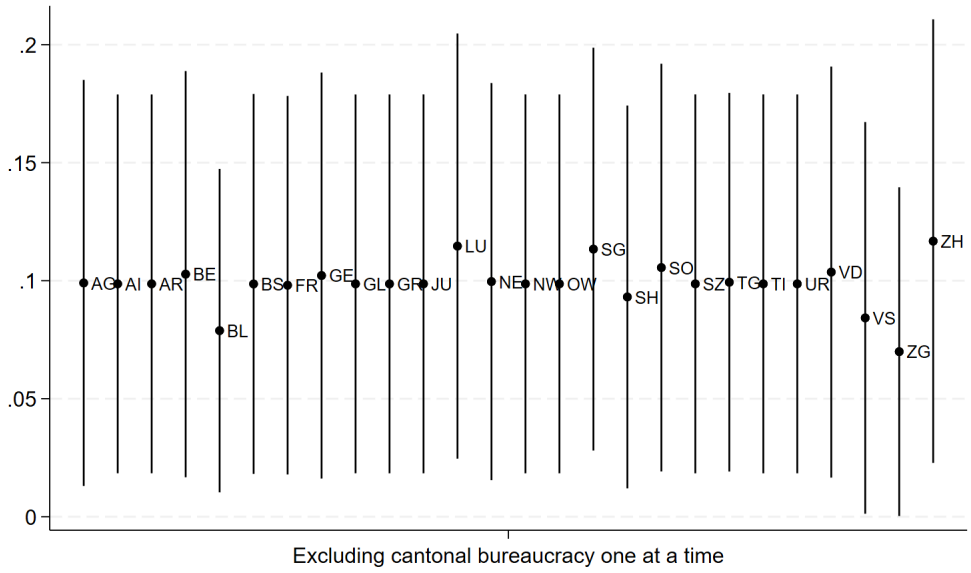
Notes: The figure summarizes regression coefficients of the switching indicators, capturing systematic differences in universalism norm endorsement between those who stay and those who switch at some future point. The above figure considers all switches while the figure below excludes switches at the federal level. See the text for further details. The 90% confidence intervals are based on standard errors clustered at the individual level.

Table A6: Exposure to deliberation channels and endorsement of institutional universalism: Full sample (including movers)

	(1)	(2)	(3)	(4)	(5)
Y_{ict}	Endorsement of institutional universalism				
Peer content-weighted exposure (X^P_{ct})	0.224**	0.225**	0.311***	0.307**	0.285*
	(0.090)	(0.099)	(0.106)	(0.109)	(0.156)
State-agent content-weighted exposure (X^S_{ct})	0.021	0.023*	0.029**	0.029**	0.035**
	(0.015)	(0.012)	(0.013)	(0.013)	(0.017)
Peer × state-agent interaction ($X^P_{ct} \times X^S_{ct}$)	-0.025	-0.021	-0.052*	-0.050*	-0.087**
	(0.025)	(0.020)	(0.028)	(0.028)	(0.037)
Observations	9,041	8,957	8,957	8,949	8,957
Individual fixed effects	Yes	Yes	Yes	Yes	Yes
Canton fixed effects	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes
Contemporaneous canton controls	No	Yes	Yes	Yes	No
Canton-specific linear trends	No	No	Yes	Yes	Yes
Time-varying individual controls	No	No	No	Yes	No
Lagged canton controls	No	No	No	No	Yes

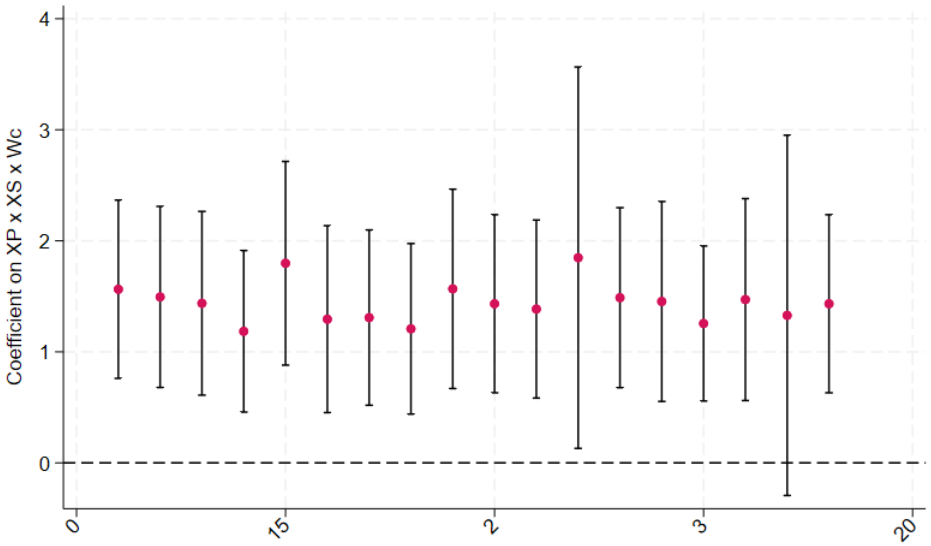
Notes: The dependent variable Y_{ict} is a binary indicator equal to one if respondent i endorses institutional universalism in canton c and year t. The sample is restricted to non-working Swiss respondents aged 18+. All specifications estimate equation (3) by OLS with individual, canton, and year fixed effects; standard errors are clustered at the canton level. The peer and state-agent content-weighted exposures, X^P_{ct} and X^S_{ct}, are mean-centered in the estimation sample so that the main-effect coefficients are evaluated at the mean of the other channel's content-weighted exposure. *Contemporaneous canton controls* include: (i) the weighted-average left–right ideology of the cantonal executive (constructed by mapping parties' cabinet shares onto a 1–4 scale using party positions from Vatter et al. (2024)); (ii) foreign population share; (iii) number of referenda; (iv) unemployment rate; and (v) taxable income per capita (cantonal average). *Lagged canton controls* replace (ii)–(v) by their one-year lag (while keeping the electoral calendar contemporaneous). *Time-varying individual controls* include marital-status indicators, number of children, and an indicator for health limitations in daily activities. Column (3) is the preferred specification; Columns (4)–(5) provide robustness to adding individual controls and to using lagged canton covariates. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Figure A4: Robustness of public employment in the non-HES sector to excluding individual cantons



Notes: The figure reports point estimates of the state effect (specification from Column (4) in Table 2), considering switches within non-HES occupations, dropping one canton at a time. Reporting 90% confidence intervals. Standard errors clustered at the individual level.

Figure A5: Robustness of the secularity interaction to excluding individual cantons



Notes: The figure reports leave-one-canton-out estimates of the coefficient on $X^P_{ct} \times X^S_{ct} \times W_c$, obtained by re-estimating the specification in Column (1) of Table 5 after excluding each German-speaking canton in turn. Dots are point estimates and vertical bars are 90% confidence intervals based on standard errors clustered at the canton level. The dashed horizontal line marks zero.