The Nation as Identity of Last Resort*

Duk Gyoo Kim†

September 24, 2026

Abstract

National identity is unusual among high-status groups: conditional on formal membership, it has no gatekeeper who can bar a citizen from claiming it on the basis of income. This paper asks what that institutional feature does once group status is dependent on the composition of such a group. I add a composition externality to the standard nation-versus-class framework: the status conferred by identifying with the nation depends on the income composition of *visible national identifiers*, so identifiers whose income lies below current composition status dilute the very standing they seek to borrow. Clubs, professions, and gated communities defend themselves against such dilution by screening on income, while national identity cannot, and I call this income-non-excludability. To give the claim content, I treat the eligibility floor as a counterfactual parameter. A sufficiently stringent floor excludes the entire low-income identifier interval, so the protection is not vacuous; and raising the floor raises equilibrium status, which makes national identity more attractive to every type that remains eligible, so that total national identification can be non-monotone in the stringency of screening. Under a two-root condition, for which I provide primitive sufficient conditions, national identifiers form two intervals at the tails of the income distribution while a middle region retains class identity. A voting model delivers the redistributive implication: identifying with the nation lowers preferred taxation by a constant amount, the same for every type whose ideal points are not compressed by the bounds on the tax rate, which can generate a downward discontinuity at the lower identity cutoff.

*This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (2025S1A5A8010761). For proofreading, language editing, and assistance in preparing figures, I used ChatGPT (OpenAI) and Claude (Anthropic). I developed and verified all theoretical arguments, derivations, calculations, and substantive research decisions. I take full responsibility for the manuscript.

†School of Economics, Yonsei University, South Korea. Email: kim.dukgyoo@yonsei.ac.kr

Highlights

- National identity is income-non-excludable but composition-dependent in status.
- National identifiers can occupy both income tails, with a middle-income hold-out.
- Income-based screening raises equilibrium status and can nationalize the middle.
- Nationalizing middle-income hold-outs can reduce majority-supported redistribution.

1 Introduction

Why do the poor vote for the rich? The puzzling observation that some low-income citizens demand less redistribution than their material interests would dictate—often by voting for conservative parties—is a persistent question in political economy that has spawned numerous hypotheses. This paper investigates one such channel, motivated by the well-documented example of South Korean elderly voters. These so-called *Taegeukgi* (the South Korean flag) ralliers exhibit a conservatism tied to an encompassing national identity, characterized by a hardline stance toward North Korea and strong support for the Korea–US alliance. For these voters, economic class is not the primary determinant of political alignment; evidence shows that a substantial portion of them are low-income (Lee and You, 2019; Lee, 2021; Kim, Kang, and Ham, 2021). Why do they choose to identify themselves with the nation, and how is it sustainable? Specifically, this paper asks how an identity-based equilibrium is sustained when the identity value of the group depends on its composition, yet the group lacks the institutional ability to exclude low-income members.

In the standard Meltzer–Richard framework, preferred redistribution decreases monotonically with income, and majority voting therefore implements the decisive median voter's preferred tax rate: voters poorer than the median prefer at least as much redistribution as the equilibrium level, whereas richer voters prefer less (Meltzer and Richard, 1981). Departures from these benchmark predictions have been explained through multidimensional party competition (Roemer, 1998), the prospect of upward mobility (Benabou and Ok, 2001), voters' biased perceptions of their relative standing (Cruces et al., 2013), and heterogeneous beliefs about intergenerational mobility (Alesina et al., 2018), among others. However, the identity-utility tradition (Akerlof and Kranton, 2000) offers an internal mechanism. In the canonical model of Shayo (2009), agents trade off the *status* of a group against their perceived *distance* from it. Because the nation is assumed to be higher-status than the poor's class, low-income agents may identify nationally, which aligns them with conservative parties so that voting for them leads to weaker redistribution. The mechanism has received experimental support (Klor and Shayo, 2010; Dhami et al., 2021) as reviewed in Shayo (2020). However, a critical feature of this framework is that the status of the nation is treated as an exogenous parameter.

The novelty of this paper is to *endogenize* national status to the income composition of who visibly identifies with the nation. (From now on, I use the term "identifier" or "national identifier" as a shorthand for a citizen who chooses to make their national identity salient, in contrast to those who identify primarily with their economic class.) In this environment, an identifier whose income lies below the current composition status dilutes the very standing she seeks to borrow, imposing a *negative externality* on every other identifier. Typically, a group whose standing depends on who belongs to it maintains the status by controlling entry: For example, clubs charge income-scaled dues, professions impose credentials, and gated communities price out entrants (Buchanan, 1965). National identity, on the contrary, has no such instrument on the income dimension: conditional on formal citizenship, no gatekeeper can rule that a citizen is too poor to be a national identifier. This paper explores the theoretical consequences of this institutional feature, which I call *income-non-excludability*. Because national identity is non-excludable but its value is (de)graded by its composition, it operates analogously to a congested public good.

To capture the concept and the effect of the income-non-excludability, I introduce an income-based floor for having a national identity as a pure counterfactual benchmark. Note that this excludability floor is never meant to be proposed as a realistic policy prescription, but as a thought experiment designed to demonstrate the core reasoning that the absence of this screening mechanism allows the nation to serve as the "identity of last resort". While modern nations do not explicitly screen identity on income, the conceptual exercise is historically grounded. The analogy is institutional rather than literal. Historical franchise restrictions did not determine who could regard herself as a national, but they demonstrate that access to central institutions of national political membership was once explicitly conditioned on property or income (Acemoglu and Robinson, 2000). By treating the eligibility floor as a mathematical parameter, the model mechanically excludes the poorest identifiers, which raises the equilibrium composition status of the group of identifiers. This elevated status makes national identity more attractive to middle-income hold-outs. Consequently, total national identification can be non-monotone in the stringency of the screening.

The net utility gain from national identity varies systematically across types: the poor identify nationally to borrow standing derived from national eminence, while the rich do so because the opportunity cost of national salience declines past its middle-income peak. The middle class, by contrast, has the most to lose from submerging its local distinctiveness. Taken together, these forces motivate a net-gain function with two crossings. I formalize this through a two-root condition and provide primitive curvature conditions sufficient for it. Under this condition, national identifiers form two intervals at the *tails of the income distribution*, while the middle region holds out, retaining its class identity.

Furthermore, mapping this identity choice into a voting model delivers the redistributive implication. I assume that national identifiers perceive redistributive taxation as eroding part of the country's economic standing, so that a tax rate t leaves eminence $a(1-pt)$ in place of a, and that they weigh this perceived loss alongside their material payoff. Conditional on national salience, each identifier's preferred tax rate is therefore lower than under class salience by a constant amount, pa/d, whenever neither ideal point is compressed by the bounds on the tax rate; at the lower identity cutoff the schedule of ideal tax rates jumps. Because identity is fixed at the voting stage, preferences over the tax rate remain single-peaked, and the majority tax is the median of the ideal-tax distribution rather than the ideal point of the median-income voter. Because the hold-outs prefer higher taxes than they would under national salience, their class salience holds the majority tax above what it would otherwise be, and the low-income national identifiers depend on them for the redistribution they receive. National eminence thus plays two roles: a higher a makes national identity more attractive, and it raises the tax reduction that national identifiers demand once they have adopted it.

The model of this paper is technically related to models of group formation under status concern, where agents care about rank within a chosen group and providers sort them into rival horizontal groups (Staab, 2024; Moldovanu et al., 2007; Bernard et al., 2016). The model considered in Staab (2024) features mutually exclusive groups where a provider can optimally screen out low-status members to protect the group's standing. In contrast, the identity choice modeled in this paper operates within a nested structure: class membership is a proper subset of national membership. Because the nation lacks a centralized gatekeeper, the exclusionary tools that resolve status dilution in standard club models are institutionally unavailable. The paper also connects to dynamic models of identity (Yuki, 2023), dissonance-based identity trade-offs (Grossman and Helpman, 2021), international union fragility (Abramson and Shayo, 2022), belief-based cultural shifts (Bonomi et al., 2021), and identity-conflict feedback loops (Sambanis and Shayo, 2013).

This paper is organized as follows. Section 2 sets up the environment and establishes existence of the equilibrium. Section 3 provides the conditions for the two-tail characterization. Section 4 explores the counterfactual excludability, and Section 5 discusses the effect on redistribution. Section 6 provides a worked parametric example to illustrate the cutoffs of the two-tail characterization and comparative statics. Section 7 discusses testable implications, and Section 8 concludes.

2 A Model

2.1 Environment

Consider that a unit mass of economic agents (citizens) have income type $\theta \in \Theta = [\underline{\theta}, \overline{\theta}] \subset \mathbb{R}_{++}$, distributed by F with continuous and strictly positive density f on the interior of Θ. Every agent is both a member of an income class and a citizen of the nation. Since these two memberships are vertically nested, they cannot choose one membership exclusively. Instead, what the agent chooses is which of them is behaviorally salient. For example, if a middle-income person chooses to see herself as a citizen of the nation rather than a member of the middle-income class, she acts on, displays, and draws utility more from the national identity. Write $g \in \{c, n\}$ for class-salient and nation-salient identity respectively, and let $\mathsf{N} \subseteq \Theta$ denote the set of types for whom n is salient, which I call "national identifiers" or "identifiers" for short. The alternative of choosing class identity is modeled in reduced form as a type-dependent payoff $\delta(\theta)$. This payoff function captures the value of retaining class-based distinctiveness, without having to model a separate group with its own status dynamics. Where I need to distinguish those who choose c from others choosing n, I call them "hold-outs".

An agent's utility value of identifying nationally has two components:

$$\underbrace{a}_{\text{eminence}} + \underbrace{s(\mathsf{N})}_{\text{composition}}, \tag{1}$$

where $a \geq 0$ is an exogenous parameter capturing national prestige that does not depend on who identifies, such as international standing, military or economic weight, cultural achievement, and historical eminence, and $s(\mathsf{N})$ captures the perceived social standing of the *visible identifier group* itself. The distinction is important for interpretation. I do not claim that the average income of self-conscious national identifiers determines the international status of a country; that is what a is for. In contrast, $s(\mathsf{N})$ captures the endogenous status of the national identifiers: the social standing derived specifically from the income composition of those who actively make their national identity salient. Formally,

$$s(\mathsf{N}) = \frac{\int_{\mathsf{N}} \omega(\theta)\theta dF(\theta)}{\int_{\mathsf{N}} \omega(\theta) dF(\theta)}, \tag{2}$$

with weight $\omega(\theta) > 0$ continuous and non-increasing. Entry by a positive marginal mass of type θ lowers s when $\theta < s$ and raises it when $\theta > s$. Two points worth noting are that (1) whether a type counts as "low" is relative to current status, not to the population mean, and (2) because each individual has measure zero no single agent's choice moves s. (Later, these will be elaborated in Lemma 2.) The steepness of ω governs the *strength of dilution*: $\omega \equiv 1$ gives equal salience to every identifier, so $s(\mathsf{N})$ becomes the simple average of the identifier type; a strictly decreasing ω amplifies the influence of lower types, making standing especially sensitive to the least prestigious identifiers. Where a one-parameter family is needed, I use $\omega_\xi(\theta) = e^{-\xi\theta}$ with $\xi \geq 0$; $\xi = 0$ recovers the simple average as $\omega_0(\theta) = 1$, and larger ξ places more weight on the bottom. The general results only use that ω is non-increasing.

I interpret the decreasing weight as reduced-form bottom-weighting in observers' evaluations of the visible identifier community. For example, although the *Taegeukgi* ralliers include some wealthy voters, outsiders' perception disproportionately weights the low-income elderly voters. Weakest-link and downward-comparison (Wills, 1981) mechanisms provide related interpretations, but need not generate the exact same functional form.

Individuals derive utility from their social identities by balancing two fundamental psychological needs, as formalized in the optimal distinctiveness theory of Brewer (1991): the need for *inclusion in a broader group* and the need for *differentiation to maintain uniqueness*. Identifying with the nation (n) satisfies the need for inclusion by allowing the individual to borrow the group's aggregate standing. However, submerging oneself in such a large group incurs a cost to individual distinctiveness. Conversely, identifying with one's economic class (c) provides no broad borrowed standing but allows the individual to retain their localized distinctiveness. Formally, identifying nationally yields utility from borrowed standing scaled by a type-dependent weight, net of the distinctiveness cost, while identifying with class yields the retained local distinctiveness:

$$U_n(\theta;\mathsf{N}) = \beta(\theta)[a + s(\mathsf{N})] - k(\theta), \qquad U_c(\theta) = \delta(\theta). \tag{3}$$

The primitives of these utility functions encode the inclusion–differentiation trade-off:

- $\beta(\theta) > 0$ decreases in θ, meaning that those with the least standing of their own place enjoy the highest marginal value on borrowed group standing.
- $k(\theta) > 0$ decreases in θ, meaning that agents whose individual standing is robust to group boundaries lose little by adopting an encompassing identity, whereas lower types face a higher cost of lost individuality.
- $\delta(\theta) > 0$: agents value retaining their class distinctiveness. I assume it is single-peaked with an interior peak at θ_m, capturing the intuition that while abandoning a

reference group is costly for all types, the middle class relies on it most. The poor have little class-based prestige to lose, and the rich possess individualized standing that transcends group labels, leaving middle-income types uniquely dependent on localized class boundaries for their social distinctiveness. This shape carries the economic narrative and generates the upper tail in the numerical illustration of Section 6, but this assumption is not what the main results rely on.[1]

Define the *net gain* from national identity at eminence a and composition status s,

$$\Delta(\theta; a, s) \equiv \beta(\theta)(a+s) - k(\theta) - \delta(\theta). \tag{4}$$

Throughout this paper, I treat the eligibility rule for identifying nationally as an exogenous feature of the environment rather than as the solution to a gatekeeper's optimization problem.[2] The *income-contingent floor* $\phi \in [\underline{\theta}, \overline{\theta}]$ admits only types $\theta \geq \phi$ to national identity, serving as a reduced form for income-scaled dues, credential requirements, or admission rules. Then the *income-non-excludable* institution corresponds to the limiting case where $\phi = \underline{\theta}$: conditional on formal membership, no rule bars a citizen from claiming national identity because she is poor. Section 4 treats movements in ϕ as a counterfactual institutional comparative static.

2.2 Equilibrium

Given the institutional floor ϕ, write $\mathsf{N}_\phi(s) = \{\theta \geq \phi : \Delta(\theta; a, s) \geq 0\}$ for the set of national identifiers at candidate status s, and $\Phi_\phi(s) = s(\mathsf{N}_\phi(s))$ for the induced composition status. For notational simplicity, subscript ϕ is omitted when it is set at $\phi = \underline{\theta}$. Define an *identity equilibrium* as follows:

Definition 1 (Identity equilibrium)**.** Given institution ϕ, an *identity equilibrium* is a set $\mathsf{N}^*_\phi \subseteq [\phi, \overline{\theta}]$ such that, with $s^* = s(\mathsf{N}^*_\phi)$, $\mathsf{N}^*_\phi = \{\theta \geq \phi : \ \Delta(\theta; a, s^*) \geq 0\}$.

The equilibrium mapping is well defined only for institutional floors that sustain a positive-mass identifier set at every candidate status. I refer to such floors as admissible.

[1] Proposition 2 uses only Assumptions 1 and 2; Lemma 1 offers one route to the latter, and needs only that δ be no more convex than $\beta(\theta)(a+s) - k(\theta)$, of which concavity of δ is a special case.

[2] Modeling a gatekeeper who maximizes the average standing of admitted members yields a trivial outcome: to elevate the group's status, the gatekeeper would simply raise the requirement to exclude everyone except the highest-income individuals. See more discussions in Section 4.

Definition 2 (Admissible floor)**.** An income-contingent floor ϕ is *admissible* if

$$\inf_{s\in[\underline{\theta},\overline{\theta}]}\int_{\mathsf{N}_{\phi}(s)} dF(\theta) > 0.$$

An inadmissible floor may leave no positive-mass set of eligible national identifiers at some candidate status, in which case the composition-status map Φ_{ϕ} is not defined on the entire candidate-status domain. It is worth noting that not every floor need be admissible. In particular, sufficiently stringent screening (for example, $\phi = \overline{\theta}$) may leave no positive-mass identifier community and therefore destroy an equilibrium of the form studied here.

Definition 1 can then be restated as a fixed-point condition: $\Phi_{\phi}(s^*) = s(\mathsf{N}_{\phi}(s^*)) = s(\mathsf{N}^*_{\phi}) = s^*$. The following conditions make Φ_{ϕ} well defined and continuous.

Assumption 1. (i) $\int_{\{\theta:\Delta(\theta;a,s)=0\}} dF(\theta) = 0$ for every $s \in [\underline{\theta},\overline{\theta}]$; and

(ii) $\Delta(\underline{\theta};a,\underline{\theta}) > 0$, which implies that $\phi = \underline{\theta}$ is admissible.

Condition (i) of Assumption 1 states that the set of indifferent types, $\{\theta : \Delta(\theta;a,s) = 0\}$, has measure zero, a property that holds whenever the net gain function is non-constant on any positive-measure set. Evaluating this condition across the entire domain of candidate statuses, rather than solely at equilibria, is necessary for ensuring the fixed-point mapping to be well-defined. Condition (ii) guarantees that the lowest type strictly prefers the national identity even under the most pessimistic composition status, $s = \underline{\theta}$. Economically, the requirement is satisfied when the lowest-income types place a high marginal value on borrowed standing (a large $\beta(\underline{\theta})$) while facing reasonably small distinctiveness costs ($k(\underline{\theta})$ and $\delta(\underline{\theta})$). Mathematically, because the net gain strictly increases in expected status ($\partial\Delta/\partial s = \beta(\theta) > 0$), condition (ii) of Assumption 1 ensures that $\Delta(\underline{\theta};a,s) > 0$ across all possible values of s. By continuity, condition (ii) extends to a neighborhood of types strictly above $\underline{\theta}$, guaranteeing a strictly positive mass of identifiers ($\inf_s \int_{\mathsf{N}(s)} dF(\theta) > 0$) for any candidate status. The strict inequality is essential, as a weak inequality would permit the identifying set to collapse to a zero-measure point.

Proposition 1 (Existence)**.** *Under Assumption 1(i) and for any admissible floor ϕ, an identity equilibrium exists. The set of equilibrium status levels is compact, so highest- and lowest-status equilibria are well defined.*

Proof. See Appendix A. □

Proposition 1 establishes the existence of an identity equilibrium. The intuition relies on a self-confirming expectations mechanism: individuals anticipate a hypothetical composition status and sort into the national identity accordingly, thereby generating a realized group status. Because the underlying type distribution is continuous and the mass of identifiers is bounded away from zero, the realized status responds continuously to changes in expectations. Such continuity guarantees at least one self-consistent outcome where the anticipated standing of the national identity and its actual composition align.

One might expect monotonicity could settle the fixed-point problem. Higher composition status raises the net gain from national identity for every type, since $\partial\Delta/\partial s = \beta(\theta) > 0$, and it does so most strongly for the poorest, whose β is largest. Individual incentives are therefore complementary, and a larger identifier set is a natural response to a higher s. However, such complementarity does not survive aggregation. The mapping Φ returns not the size of the identifier set but its ω-weighted mean income. A set that expands at both margins admits new low types alongside new high ones, so composition status may decrease as the pool grows. That failure of monotonicity is the composition externality restated at the aggregate level. Existence accordingly relies on continuity rather than a monotone-map argument. Because uniqueness is not guaranteed, subsequent results requiring equilibrium selection are evaluated at the maximal equilibrium status.

3 Two-Tail Pattern

Proposition 1 guarantees that an identity equilibrium exists for every admissible institutional floor, but it still leaves the question about how the national identifiers are characterized. I now characterize identity sorting under the institution of primary interest, the income-non-excludable baseline ($\phi = \underline{\theta}$). Under the baseline participation condition of Section 2.2, this floor is admissible. Given that low-income citizens derive the highest marginal utility from borrowed standing, a natural conjecture is a simple threshold equilibrium where national identifiers form a single lower tail. Such a conjecture entails an immediate question: if the poor heavily rely on national composition status, why do middle-income citizens not similarly adopt the encompassing identity? The divergence occurs because identity choice is driven by net rather than gross benefits. While the middle class places a moderate value on borrowed standing, they simultaneously face the highest opportunity cost of abandoning their distinctiveness within the class. High-income types, by contrast, face negligible distinctiveness costs, allowing them to opt back into the national group. To obtain a clean sorting characterization, I impose a two-root restriction on the net gain function. The restriction is stated directly in terms of Δ, after which Lemma 1 provides primitive curvature conditions sufficient for it.

Assumption 2 (Two roots)**.** For each relevant s, the function mapping θ to $\Delta(\theta;a,s)$ has at most two zeros on Θ.

Assumption 2 imposes a shape restriction directly on the net gain function $\Delta(\theta;a,s)$. The monotonicity assumptions on β and k give the underlying economic forces their intended directions, but do not by themselves rule out additional crossings. The two-root restriction prevents the identifier set from fragmenting into multiple disconnected regions. (For example, if $\Delta(\theta;a,s)$ "touches" zero once in the negative region, then any such tangent point would belong to the identifier set $\{\Delta(\theta;a,s)\geq 0\}$.) This restriction, however, need not be taken simply on faith; the following lemma establishes explicit primitive sufficient conditions that guarantee properties stated in Assumption 2.

Lemma 1 (Primitive sufficient condition for Assumption 2)**.** *Given a and s, if $\beta(\theta)(a+s)-k(\theta)$ is convex and $\delta(\theta)$ is concave on Θ, then $\Delta(\cdot;a,s)$ is convex. Consequently, the strict sublevel set $\{\theta\in\Theta:\Delta(\theta;a,s)<0\}$ is an interval and $\Delta(\cdot;a,s)$ changes sign at most twice. In addition, if $\int_{\{\theta:\Delta(\theta;a,s)=0\}} dF(\theta)=0$, as required by Assumption 1(i), then Δ has at most two zeros, and Assumption 2 holds.*

Proof. See Appendix A. □

These conditions are satisfied, for example, by the parametric family

$$\beta(\theta)=\frac{b}{\theta+\eta},\qquad k(\theta)=\frac{\kappa}{\theta+\eta},\qquad \delta(\theta)=\delta_0-\delta_1(\theta-\theta_m)^2, \tag{5}$$

with $\eta,b,\kappa,\delta_0,\delta_1>0$. Here, $\beta(\theta)(a+s)-k(\theta)=\frac{b(a+s)-\kappa}{\theta+\eta}$ is convex when $b(a+s)>\kappa$, and δ is concave. Numerical illustrations shown in Section 6 use this family of functional forms.

The following proposition characterizes the configurations of the national identifiers.

Proposition 2 (Configurations of the identifier set)**.** *Consider the income-non-excludable institution $\phi=\underline{\theta}$, and let s^* be an equilibrium status. Under Assumptions 1 and 2, one of the following holds, up to F-null sets:*

(i) (Pooling) $\min_{\theta\in\Theta}\Delta(\cdot;a,s^*)\geq 0$ *and* $\mathsf{N}^*=\Theta$.

(ii) (Single tail) $\min_{\theta\in\Theta}\Delta(\cdot;a,s^*)<0$ *and* $\Delta(\overline{\theta};a,s^*)\leq 0$*; then there is a unique* $\tau\in(\underline{\theta},\overline{\theta})$ *such that* $\mathsf{N}^*=[\underline{\theta},\tau]$ *and* $\Delta(\tau;a,s^*)=0$.

(iii) (Two tails) $\min_{\theta\in\Theta}\Delta(\cdot;a,s^*)<0$ *and* $\Delta(\overline{\theta};a,s^*)>0$*; then there are* $\underline{\tau}$ *and* $\overline{\tau}$ *with* $\underline{\tau}<\overline{\tau}$ *in the interior of* Θ *such that* $\mathsf{N}^*=[\underline{\theta},\underline{\tau}]\cup[\overline{\tau},\overline{\theta}]$*, with both tails and the hold-out region* $(\underline{\tau},\overline{\tau})$ *having positive* F*-measure, and* $\Delta(\underline{\tau};a,s^*)=\Delta(\overline{\tau};a,s^*)=0$.

Proof. Appendix A. □

It is worth noting that while both the rich ($[\overline{\tau},\overline{\theta}]$) and the poor ($[\underline{\theta},\underline{\tau}]$) may adopt the national identity, as described in Proposition 2(iii), they do so for different reasons. The lower tail is driven by a high valuation of borrowed national standing, making $\beta(\theta)(a+s^*)$ the binding term for low-income types. Conversely, the upper tail is supported by the declining total opportunity cost of national salience, $k(\theta)+\delta(\theta)$, once the value of distinctiveness within a class has begun to fade at higher incomes. The model, however, does not isolate a single force driving this upper tail: national identity can be salient at the top because submersion becomes cheaper, because local class distinctiveness fades, or both. In the worked example of Section 6, the latter heavily dominates: over the upper half of the income range, δ drops by 1.00 while k falls by merely 0.04. However, the numerical illustration is strictly a parameterized exercise and should not be interpreted as empirical evidence. Fully disentangling these "opposite reasons" requires exogenous shifts that independently affect β and $k+\delta$, which directly motivates the experimental design suggested later in the paper.

Cases (i) and (ii) of Proposition 2 represent familiar benchmarks. Pooling arises when borrowed standing is attractive enough that no type declines it, and the single-tail configuration is the pattern shown in Shayo (2009), where national identification falls monotonically in income, which creates one crossing point. Case (iii) is the configuration of primary interest because it produces the coexistence of low- and high-income national identifiers with a middle hold-out region, which underlies the paper's identity-of-last-resort mechanism and subsequent comparative statics. As previously discussed, Lemma 1 provides primitive conditions ensuring that the required two-root restriction is satisfied. Whether the equilibrium is pooling, single-tail, or two-tail still depends on parameter values. In the worked family of Section 6, all three configurations arise as a varies, and the two-tail configuration obtains over a non-degenerate interval of national eminence rather than at a single parameter value. It is also worth noting that this two-tail identity pattern can be translated into a corresponding non-monotonicity in redistributive preferences, as will be shown in Section 5.

4 Income-Based (Non-)Excludability

This section serves as a counterfactual thought exercise to examine the role of income-non-excludability. National identifiers whose type lies below current composition status lower $s(\mathsf{N})$, and other identifiers would prefer to be without them. How, then, would the institutional floor affect the tension? Before comparing institutions, it is worth making that externality precise, since it is what the institutional comparison is all about. To represent the arrival of a small group of identifiers concentrated near a type $z \in \Theta$, let $\{I_\varepsilon\}_{\varepsilon>0}$ be a nested family of intervals containing z with $\int_{I_\varepsilon} dF(\theta) = \varepsilon$ and $I_\varepsilon \downarrow \{z\}$ as $\varepsilon \downarrow 0$; such a family exists because $f > 0$. The following lemma summarizes the effect of a newly added identifier group on $s(\mathsf{N})$. For notational convenience, write $\mathsf{N}_\varepsilon := \mathsf{N} \cup I_\varepsilon$.

Lemma 2 (Marginal composition externality)**.** *Let* N *have weighted mass* $W = \int_{\mathsf{N}} \omega(\theta) dF(\theta)$ *and status* $s = s(\mathsf{N})$*, and suppose* $I_\varepsilon \cap \mathsf{N} = \emptyset$*. Then*

$$s(\mathsf{N}_\varepsilon) = \frac{Ws + \int_{I_\varepsilon} \omega(\theta)\theta dF(\theta)}{W + \int_{I_\varepsilon} \omega(\theta) dF(\theta)}, \qquad \left.\frac{ds(\mathsf{N}_\varepsilon)}{d\varepsilon}\right|_{\varepsilon=0} = \frac{\omega(z)}{W}(z-s).$$

Hence entry by a marginal mass of types near $z < s$ *strictly lowers composition status, and the payoff effect on an incumbent national identifier of type* θ *is*

$$\beta(\theta)\frac{\omega(z)}{W}(z-s) < 0.$$

For a fixed entrant type $z < s$*, this effect is largest in absolute value for incumbents with the highest* β*, that is, for the lowest-income identifiers.*

Proof. See Appendix A. □

Lemma 2 describes the model's central tension, and is where the economic content of this section lies. Types below current status may privately prefer national identity; a positive mass of such entrants lowers the standing available to every incumbent identifier; and no agent internalizes this, because each individual is of measure zero and only a marginal *mass* moves s. How much a low-type entrant depresses status depends on how heavily the weight ω concentrates on the bottom of the identifier community. For the exponential family that I consider in Section 2, this dependence has a closed form.

Lemma 3 (Direct effect of dilution strength)**.** *Fix* N *with* $\int_{\mathsf{N}} dF(\theta) > 0$*, and let* $s_\xi(\mathsf{N})$ *be* (2) *with* $\omega_\xi(\theta) = e^{-\xi\theta}$. *Then*

$$\frac{\partial s_\xi(\mathsf{N})}{\partial \xi} = -\mathrm{Var}_{\mu_\xi}(\theta) \leq 0,$$

where μ_ξ *is the probability measure on* N *with density proportional to* $e^{-\xi\theta} f(\theta)$*, with strict inequality whenever* μ_ξ *is non-degenerate.*

Proof. See Appendix A. □

Greater weights on the bottom therefore lower the standing available to be borrowed from any fixed identifier set. The equilibrium response is also affected by the induced change in N, and I do not claim any general result about it; in the worked example of Section 6, however, the direction is the same in equilibrium. The question is then whether any institution can prevent the dilution that Lemma 2 describes. Stated at the level of the eligibility rule alone, the answer is straightforward, and I thus record it as a corollary.

Corollary 1 (Institutional exclusion of low-income identifiers)**.** *Consider an equilibrium under the income-non-excludable institution* $\phi = \underline{\theta}$*, and let* $[\underline{\theta}, \underline{\tau}]$ *denote the maximal lower interval contained in the equilibrium identifier set* N^**. Then:*

(i) under any institution with floor $\phi > \underline{\theta}$*, all baseline identifiers with* $\theta < \phi$ *are rendered ineligible for national identity; in particular, if* $\phi > \underline{\tau}$*, the entire baseline lower identifier interval* $[\underline{\theta}, \underline{\tau}]$ *is excluded;*

(ii) under $\phi = \underline{\theta}$*, every type in* $[\underline{\theta}, \underline{\tau}]$ *remains eligible and identifies whenever its private net gain is non-negative, irrespective of the composition externality it may impose on other identifiers.*

Thus, income-non-excludability prevents eligibility rules from selectively removing low-income identifiers, whereas an income-contingent floor can do so directly.

Proof. Both parts (i) and (ii) are immediate from the definition of $\mathsf{N}_\phi(s)$. □

Corollary 1 is substantive not because it compares two eligibility rules, which is straightforward given the characterization, but because it helps to articulate the role of the institutional floor: As ϕ is treated as a parameter, moving it generates an equilibrium effect that the definition alone does not deliver. Raising ϕ removes low-income identifiers, which raises composition status, which in turn makes national identity more attractive to every type that remains eligible. Income-non-excludability is the nested case where $\phi = \underline{\theta}$, not a separate object, and the effects of moving ϕ away from it are established below.

Again, it is critical to emphasize that the eligibility floor is a counterfactual institutional parameter rather than a proposed policy instrument. Modern nation-states do not ordinarily condition a citizen's ability to claim national identity on an income threshold. The purpose of varying ϕ is instead to isolate what income-non-excludability contributes to equilibrium: if low-income identifiers could be screened out, the composition of the national-identifier group, and accordingly the standing available to those who remain, would change endogenously. The following lemma makes this equilibrium effect precise.

Lemma 4 (Monotone institutional comparative static)**.** *Suppose Assumption 1(i) holds, and let $\phi' > \phi$ be two admissible floors. Then $\Phi_{\phi'}(s) \geq \Phi_{\phi}(s)$ for every $s \in \Theta$, and the largest equilibrium status $s^*_{\max}(\phi) = \max\{s \in \Theta : \Phi_{\phi}(s) = s\}$ is non-decreasing in ϕ.*

Proof. See Appendix A. □

Lemma 4 provides the equilibrium implication of the institutional comparison. A higher floor does not merely remove the excluded types mechanically. By raising equilibrium composition status, it raises the net gain from national identity for every type that remains eligible, since $\partial\Delta/\partial s = \beta(\theta) > 0$. Therefore, screening can bring former middle-income holdouts into the national identifiers while directly excluding low-income identifiers. The effect on the total mass of national identifiers is therefore ambiguous for moderate floors: the direct loss of excluded types and the induced expansion among remaining eligible types work in opposite directions. It is not ambiguous for stringent ones. Since $\int_{\mathsf{N}^*_\phi} dF(\theta) \leq \int_\phi^{\bar{\theta}} dF(\theta)$, the mass of identifiers vanishes as ϕ approaches $\bar{\theta}$ regardless of the induced entry, so sufficiently stringent screening reduces national identification below its baseline level. The worked example of Section 6 shows both: moderate screening moves the mass non-monotonically, and stringent screening drives it toward zero.

Three qualifications clarify the scope of this result. First, no optimizing gatekeeper is modeled. The floor is an exogenous institutional parameter used to compare environments, not the solution to an admission problem. Indeed, a gatekeeper maximizing only the average standing of admitted identifiers would have an incentive to push the floor toward the top of the income distribution, producing an uninformative extreme. Second, Lemma 4 is not a welfare result. A higher floor raises the standing available to eligible identifiers, but excluded agents lose access to national identification, so higher composition status does not imply a Pareto improvement. Third, income-non-excludability is explicitly dimension-specific. National communities may exclude through citizenship law, ethnicity, language, religion, and social recognition. The model takes formal membership as given and claims only that, conditional on such membership, a citizen cannot be made ineligible to regard national identity as salient merely because her income is too low. Thus the absence of an income-contingent gatekeeper is compatible with substantial exclusion on other dimensions.

5 Redistribution

The preceding analysis determines which citizens make national identity salient. Now I examine how the resulting identity configuration translates into redistribution policy. The central idea is that national salience changes the weight placed on the national consequences of taxation. This gives national eminence a a second role: It affects not only the attractiveness of national identity, as in the preceding sections, but also the policy preferences of those for whom that national identity becomes salient.

I treat identity formation and voting sequentially. Citizens first choose identity salience given the inherited level of national eminence a, generating an identity equilibrium. Conditional on that identity configuration, they then vote over redistribution. This timing separates the identity-composition mechanism from the subsequent political aggregation problem and keeps preferences over the policy instrument one-dimensional.

Let $t \in [0,1]$ denote a linear tax financing an equal lump-sum transfer, and let $\mu \equiv \mathbf{E}_F[\theta]$ denote mean income. The material payoff of an individual of type θ is

$$m(t;\theta) = (1-t)\theta + t\mu - \frac{d}{2}t^2, \qquad d > 0, \tag{6}$$

where the first two terms are the standard tax-and-transfer components and the quadratic term captures the efficiency cost of redistribution in reduced form. $d > 0$ captures the magnitude of the efficiency cost of redistribution: a larger d implies a more rapidly increasing deadweight loss as the tax rate rises.

Redistributive taxation has long been associated with a potential trade-off between distributional objectives and aggregate economic performance.[3] I therefore capture the growth-versus-distribution consideration as a *perceived* rather than mechanical relationship: national identifiers view higher redistributive taxation as potentially eroding some of the coun-

[3] Empirically, however, the evidence does not support a universal negative effect of redistribution on growth; the consequences appear to depend on the extent and design of redistribution (Ostry et al., 2014). At the same time, tax rates can generate economically meaningful behavioral and aggregate responses. For example, Mertens and Montiel Olea (2018) find that changes in marginal tax rates affect reported income and aggregate economic activity in the United States.

try's economic standing. Let the level of national eminence preserved under tax rate t be

$$a(t) = a(1 - pt), \qquad 0 < p \leq 1, \tag{7}$$

where a is the inherited level of national eminence introduced in Section 2, and p measures the perceived policy cost of redistribution in terms of national eminence, incorporating both the perceived sensitivity of national economic standing to taxation and the weight national identifiers place on that consequence. This formulation does not require taxation literally to reduce national standing one-for-one through current output. Rather, it summarizes the perception that a more redistributive tax system may come at the cost of weaker aggregate economic performance and, consequently, lower national economic eminence.

Relative to the inherited level a, the perceived loss associated with taxation is therefore $a - a(t) = pat$. The utility associated with redistribution policy t is then

$$V(t;\theta,g) = m(t;\theta) - pat\,\mathbf{1}\{g = n\}. \tag{8}$$

The roles of $\beta(\theta)$ and p are conceptually distinct. The former, introduced in $U_n(\theta;\mathsf{N})$, measures how much an individual values borrowing national standing when choosing which identity to make salient. The latter summarizes the perceived cost of redistribution policy in terms of national eminence: it captures, in reduced form, how strongly taxation is perceived to erode national economic standing once national identity is salient. In other words, p incorporates both the perceived sensitivity of national eminence to taxation and the importance attached to that consequence.[4]

Writing

$$x(\theta) \equiv \frac{\mu - \theta}{d}, \qquad h(a) \equiv \frac{pa}{d}, \tag{9}$$

the individual ideal tax rates are

$$t^*(\theta,c) = [x(\theta)]_0^1, \qquad t^*(\theta,n) = [x(\theta) - h(a)]_0^1, \tag{10}$$

where $[y]_0^1$ denotes y capped below at 0 and above at 1. Thus national salience weakly lowers

[4] The policy stage is treated as subsequent to the identity equilibrium. Thus a is inherited national eminence when identity is chosen, while $a - a(t)$ is the perceived loss associated with the subsequent policy choice. Allowing the policy outcome to feed back contemporaneously into identity formation would introduce an additional fixed-point problem that is not needed for the mechanism studied here.

each individual's preferred tax rate. Define

$$D_a(x) \equiv [x]_0^1 - [x - h(a)]_0^1. \tag{11}$$

For the parameter region $0 < h(a) \leq 1$, $0 \leq D_a(x) \leq h(a)$; the reduction equals $h(a)$ whenever $h(a) \leq x \leq 1$, and equals zero exactly when $x \leq 0$ or $x \geq 1 + h(a)$. The identity-induced reduction $D_a(x)$ can disappear at either boundary of the tax interval. Under either identity, citizens with sufficiently low (high) θ may prefer $t = 1$ ($t = 0$). National salience therefore need not alter tax demand at the extremes even though it lowers the preferred tax for intermediate types.

This observation is particularly useful for interpreting the upper tail of the two-tail equilibrium. If $\overline{\tau} \geq \mu$, then every upper-tail identifier satisfies $x(\theta) \leq 0$ and therefore prefers $t = 0$ regardless of identity. The identity channel then operates on those low- and middle-income citizens whose preferred tax rates are not already constrained at a boundary.

I now turn from individual policy preferences to political aggregation. Fix an identity equilibrium under institutional floor ϕ, and let $g_\phi^*(\theta) \in \{c, n\}$ denote the salient identity of type θ under ϕ. Define the associated ideal-tax schedule by

$$q_\phi(\theta) \equiv t^*(\theta, g_\phi^*(\theta)). \tag{12}$$

Although $q_\phi(\theta)$ is not necessarily monotone in income, each individual's preferences over t are strictly concave and hence single-peaked. Majority voting can therefore be characterized by the median of the distribution of ideal tax rates rather than by median income itself. The distinction is familiar from majority voting over a publicly provided good that households may supplement privately, where the pivotal voter can have income below the median (Epple and Romano, 1996a); here the separation between the pivotal voter and the median-income voter arises from identity salience rather than from private supplementation. Assuming universal turnout,[5] the standard median-ideal-point result therefore applies: the majority-rule equilibrium tax is a median of the distribution of $q_\phi(\theta)$. When this median is unique, I denote it by t_ϕ^{MV}.[6] Unless otherwise stated, I call the equilibrium tax rate determined by majority voting the majority tax.

[5] I assume that no voters abstain. Endogenous or strategic abstention could alter the identity of the pivotal voter and is left for future work.

[6] If the distribution of ideal tax rates has multiple medians, the majority-rule equilibrium is set-valued; equivalently, the set of any $t \in [0, 1]$ such that $\int_{\{\theta: q_\phi(\theta) \leq t\}} dF(\theta) \geq \frac{1}{2}$ and $\int_{\{\theta: q_\phi(\theta) \geq t\}} dF(\theta) \geq \frac{1}{2}$ is a majority-rule equilibrium tax set.

Under the income-non-excludable two-tail configuration of Proposition 2,

$$q_{\underline{\theta}}(\theta) = \begin{cases} [x(\theta) - h(a)]_0^1, & \theta \in [\underline{\theta}, \underline{\tau}], \\ [x(\theta)]_0^1, & \theta \in (\underline{\tau}, \overline{\tau}), \\ [x(\theta) - h(a)]_0^1, & \theta \in [\overline{\tau}, \overline{\theta}]. \end{cases} \tag{13}$$

At $\underline{\tau}$, the schedule jumps up by $D_a(x(\underline{\tau}))$ as identity switches from national to class; the jump is strict when $0 < x(\underline{\tau}) < 1 + h(a)$ and equals $h(a)$ when $h(a) \le x(\underline{\tau}) \le 1$.

The politically relevant object is therefore not necessarily the ideal tax of the median-income citizen, but the median of the entire ideal-tax distribution. In particular, redistributive policy is higher than it would be if the middle-income hold-outs identified nationally; in that sense, it is sustained by citizens who continue to make class identity salient.

This gives the middle hold-out region an additional economic role. It is not merely the set of citizens who decline to identify themselves nationally, but it can also provide the political support sustaining redistribution received by low-income national identifiers. Those low-income identifiers may themselves prefer less redistribution than they would under class salience, yet the majority outcome can remain relatively redistributive because the middle-income hold-outs demand higher taxes.

The institutional comparative static of Section 4 then has a political counterpart. Consider an increase in the floor from ϕ to $\phi' > \phi$, and select the largest identity equilibrium under each floor. By Lemma 4, $s^*_{\phi'} \ge s^*_{\phi}$. Since $\frac{\partial \Delta(\theta; a, s)}{\partial s} = \beta(\theta) > 0$, every type that identifies nationally under ϕ and remains eligible under ϕ' continues to identify nationally under the higher floor. Hence

$$\mathsf{N}^*_{\phi} \cap [\phi', \overline{\theta}] \subseteq \mathsf{N}^*_{\phi'}. \tag{14}$$

There is therefore no third group of still-eligible national identifiers who revert to class identity. The identity changes induced by the higher floor can be divided into two groups. Let

$$E_{\phi,\phi'} \equiv \mathsf{N}^*_{\phi} \cap [\phi, \phi') \tag{15}$$

denote previously national identifiers who become ineligible under the higher floor, and let

$$H_{\phi,\phi'} \equiv \mathsf{N}^*_{\phi'} \setminus \mathsf{N}^*_{\phi} \tag{16}$$

denote previously class-salient types who are induced to identify nationally by the resulting increase in composition status. For the first group in $E_{\phi,\phi'}$, preferred taxation changes from

$[x(\theta)-h(a)]_0^1$ to $[x(\theta)]_0^1$, leading to an upward shift of $D_a(x(\theta))$. For the second group in $H_{\phi,\phi'}$, the same quantity shifts in the opposite direction. Consequently,

$$q_{\phi'}(\theta)-q_{\phi}(\theta)=\begin{cases} D_a(x(\theta)), & \theta\in E_{\phi,\phi'}, \\ -D_a(x(\theta)), & \theta\in H_{\phi,\phi'}, \\ 0, & \text{otherwise.} \end{cases} \tag{17}$$

Equation (17) shows why the sign of the screening effect on the majority tax is not globally determined. Excluding previously national low-income citizens can raise their demand for redistribution, whereas drawing middle-income hold-outs into national identification lowers theirs. Which force matters politically depends on where these changes occur relative to the median of the ideal-tax distribution.

However, there is a notable case in which the first force disappears. The citizens with sufficiently low θ may already be at the upper bound $t=1$ under both identities. For such types, $D_a(x(\theta))=0$. Screening them out of national identity changes their identity but not their preferred tax. By contrast, if the higher composition status nationalizes middle-income hold-outs whose ideal points are away from the bounds, their preferred taxes fall strictly. This yields a particularly transparent channel through which income-based screening can reduce redistribution.

Corollary 2 (Screening with politically inactive exclusion)**.** *Consider two admissible floors $\phi<\phi'$, select the largest identity equilibrium under each, and suppose the majority-rule tax is unique under each floor. If $D_a(x(\theta))=0$ for every $\theta\in E_{\phi,\phi'}$, then $q_{\phi'}(\theta)\le q_{\phi}(\theta)$ for every θ, and consequently $t_{\phi'}^{MV}\le t_{\phi}^{MV}$.*

Proof. By (17), the condition $D_a(x(\theta))=0$ on $E_{\phi,\phi'}$ eliminates every upward change in the ideal-tax schedule. Types in $H_{\phi,\phi'}$ weakly lower their ideal taxes, while all other types have unchanged ideal taxes. Hence $q_{\phi'}(\theta)\le q_{\phi}(\theta)$ pointwise. The median of the ideal-tax distribution therefore cannot increase. □

The condition in Corollary 2 has a natural interpretation. If the citizens directly excluded by a higher floor are sufficiently poor that their preferred tax is already constrained at unity under either identity, their exclusion has no countervailing effect on tax preferences. The only politically relevant identity changes are then those of former middle hold-outs whose entry into national identification weakly lowers their demand for redistribution.

The inequality can be strict when these downward shifts are sufficiently important around the median of the ideal-tax distribution.

Another worthwhile point is that a lower majority tax can in turn make low-income citizens materially worse off. From (6),

$$\frac{\partial m(t;\theta)}{\partial t} = \mu - \theta - dt = d\left[x(\theta) - t\right]. \tag{18}$$

Thus material utility is increasing in the tax rate up to the individual's material optimum $x(\theta)$. Since $x(\theta)$ is a decreasing function, the positive marginal utility of taxation ($x(\theta) > t$) implies that the income type should be lower given t. The following Corollary organizes this argument.

Corollary 3 (Material dependence on redistribution)**.** *Suppose an institutional change lowers the majority-rule tax from t_0 to $t_1 < t_0$. For any type θ satisfying $t_0 \leq x(\theta)$, we have $m(t_1;\theta) < m(t_0;\theta)$. Thus, a low-income national identifier can become materially worse off when political support for redistribution weakens, even if national identity remains salient for that individual.*

Proof. If $t_0 \leq x(\theta)$, then (18) is non-negative throughout $[t_1, t_0]$ and strictly positive except possibly at the endpoint $t_0 = x(\theta)$. Since $t_1 < t_0$, it follows that $m(t_1;\theta) < m(t_0;\theta)$. □

Corollary 3 reveals an interdependence that is not apparent from individual identity choice alone. A poor citizen may voluntarily make national identity salient because the national group offers valuable borrowed standing. Whenever $D_a(x(\theta)) > 0$, doing so lowers her redistributive demand relative to the class-salient benchmark. Yet the redistribution she actually receives can be sustained by middle-income citizens who remain class-salient and prefer higher taxation. Poor national identifiers can therefore depend materially on the continued existence of the middle hold-out region.

This interdependence becomes particularly stark when the identity shift of the poorest citizens is politically inactive because both of their ideal tax rates are clipped at unity. Such citizens can identify nationally without changing their own preferred tax, while the middle hold-outs remain decisive for the amount of redistribution actually implemented. If an institutional change subsequently draws those middle-income citizens into the national identifier group, the resulting decline in their tax demand can reduce redistribution and leave the poor materially worse off.

National eminence reinforces this tension along two margins. First,

$$\frac{\partial \Delta(\theta; a, s)}{\partial a} = \beta(\theta) > 0,$$

so greater eminence makes national identity more attractive. Second, $h(a) = \frac{pa}{d}$ is increasing in a, so conditional on national identification, the downward shift in preferred taxation becomes larger. Greater national eminence can therefore weaken redistributive support both by changing *who* identifies nationally and by changing *how strongly* national identifiers discount redistribution. Because the identity configuration and the politically decisive portion of the ideal-tax distribution can both change, the model does not imply a globally monotone relationship between national eminence and the majority tax.

The resulting configuration is related to "ends-against-the-middle" mechanisms (Epple and Romano, 1996b), but the source of non-monotonicity differs. In Epple and Romano (1996b), for example, access to a private alternative to the publicly provided service disrupts the standard ordering of policy preferences. Here, identity is fixed at the voting stage, so each voter's preferences over the tax rate remain single-peaked and the majority outcome can be characterized by the median ideal point. The non-monotonicity instead appears in the mapping from income to ideal tax rates because identity salience itself varies non-monotonically with income. Kim (2019) obtains a related alignment through positional concern and last-place aversion: low-income agents may oppose redistribution when its positional consequences outweigh its material benefits, while high-income agents oppose it for conventional material reasons. The mechanism here operates through neither private outside options nor rank comparisons, but through the composition-dependent standing of national identity. Its distinctive implication is therefore not simply that redistributive demand can be non-monotone in income, but that the political importance of the redistributive middle changes systematically with national eminence and with the institutional capacity to exclude low-income identifiers.

6 A Worked Example

This section fixes a parametric family satisfying every maintained assumption and computes the objects of Sections 2–5: the identity equilibrium and its configuration, the response of that configuration to national eminence, dilution strength and the eligibility floor, and the majority-rule tax. Its purpose is to show that the assumptions are jointly satisfiable and to make the mechanisms visible, not to calibrate to data.

6.1 Calibration

Income is distributed on $\Theta = [1,3]$ with density

$$f(\theta) = \tfrac{3}{2} - \tfrac{\theta}{2},$$

which integrates to one, is strictly positive on the interior of Θ, and is right-skewed: mean income is $\mu = 5/3 \approx 1.667$ and median income is $\theta_{\text{med}} = 3 - \sqrt{2} \approx 1.586$, so $\mu > \theta_{\text{med}}$ and the Meltzer–Richard benchmark tax is positive. The identity primitives are

$$\beta(\theta) = \frac{b}{\theta + \eta}, \qquad k(\theta) = \frac{\kappa}{\theta + \eta}, \qquad \delta(\theta) = \delta_0 - \delta_1(\theta - \theta_m)^2,$$

with $b = 2$, $\kappa = 0.5$, $\eta = 1$, $\delta_0 = 1.5$, $\delta_1 = 1$ and $\theta_m = 2$, so that β and k are positive and decreasing and δ is positive on Θ with $\delta(1) = \delta(3) = 0.5$. The status weight is $\omega_\xi(\theta) = e^{-\xi\theta}$ with $\xi = 2$ in the baseline. National eminence is $a = 0.7$ in the baseline; the reason for this value is given in Section 6.6. For the voting stage, $d = 0.5$ and $p = 0.05$, so that $h(a) = 0.1a$ and $h(0.7) = 0.07$, which lies below the benchmark tax $x(\theta_{\text{med}}) = 0.162$.

The net gain is

$$\Delta(\theta; a, s) = \frac{M}{\theta + 1} - \tfrac{3}{2} + (\theta - 2)^2, \qquad M \equiv b(a+s) - \kappa = 2(a+s) - \tfrac{1}{2},$$

and its zeros on Θ solve the cubic $\theta^3 - 3\theta^2 - \frac{3}{2}\theta + (\frac{5}{2} + M) = 0$. The maintained assumptions hold throughout. Assumption 1(ii) reads $\Delta(1; a, 1) = a + \frac{1}{4} > 0$, satisfied for every $a > 0$. Since $b(a+s) > \kappa$ for all $s \geq \underline{\theta}$, the term $\beta(\theta)(a+s) - k(\theta) = M/(\theta+1)$ is convex, and δ is concave, so Lemma 1 delivers Assumption 2. Composition status is computed exactly: for an interval $[u, v]$, $\int_u^v \omega_\xi f \, d\theta$ and $\int_u^v \omega_\xi \theta f \, d\theta$ are polynomial-times-exponential integrals with closed forms, and fixed points of Φ are located by a sign scan over 4001 candidate statuses followed by bracketing to a tolerance of 10^{-12}. Every figure below is drawn from these values.

6.2 The two-tail equilibrium

At $a = 0.7$ and $\xi = 2$ the status map has a single fixed point, $s^* = 1.2154$. The cubic has two roots in Θ, at $\underline{\tau} = 1.5562$ and $\overline{\tau} = 2.7878$, and $\Delta(1) = +1.165$, $\Delta(2) = -0.390$, $\Delta(3) = +0.333$, so the equilibrium is in case (iii) of Proposition 2:

$$\mathsf{N}^* = [1, 1.556] \cup [2.788, 3].$$

The low tail carries F-mass 0.479, the hold-out region 0.510 and the high tail 0.011, for a total identifying mass of 0.490; the high tail is thin because the skewed density places little mass at the top. Within the low tail, the types below equilibrium status—those whose presence dilutes the standing they borrow—make up 42.6 per cent of that interval. Both mean and median income lie in the hold-out region. Figure 1(a) plots $\Delta(\cdot; a, s^*)$ with the two identifying intervals shaded.

Over the upper half of the income range, from $\theta_m = 2$ to $\bar{\theta} = 3$, δ falls by 1.00 while k falls by 0.04, so the upper tail here is generated almost entirely by declining class distinctiveness. The slope of the status map at the fixed point is $\Phi'(s^*) = 0.158$: the feedback from status to identification to status is positive but damped.

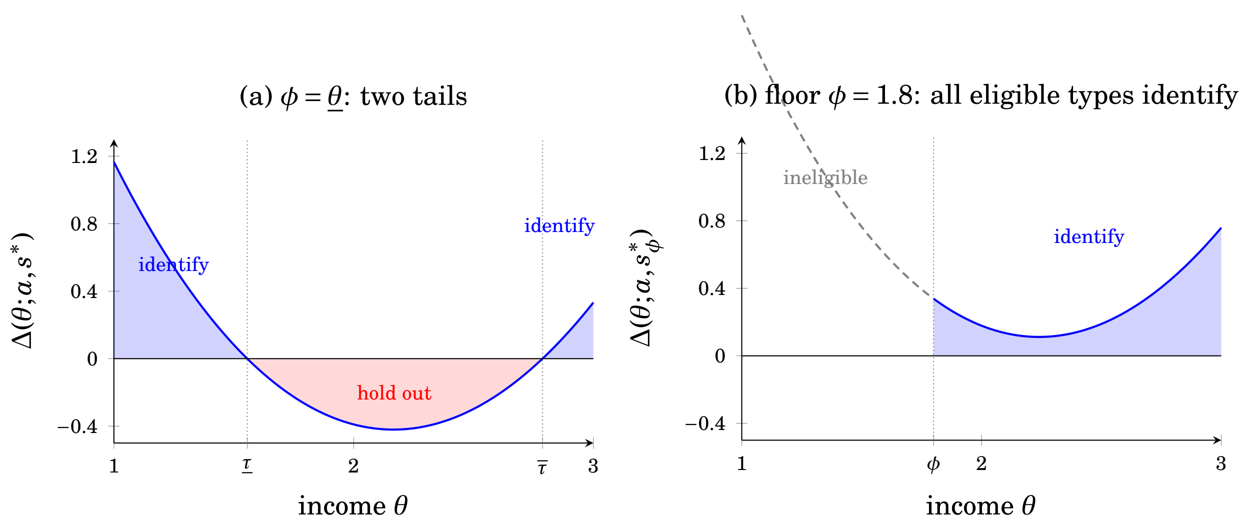

Figure 1: The net gain from national identity in the worked example ($a = 0.7$, $\xi = 2$). Panel (a): under income-non-excludability the equilibrium status is $s^* = 1.215$ and the identifier set is the union of the two shaded tails, with cutoffs $\underline{\tau} = 1.556$ and $\overline{\tau} = 2.788$; the middle region holds out. Panel (b): under a floor $\phi = 1.8 > \underline{\tau}$ the equilibrium is unique, its status is $s^*_\phi = 2.068$, the net gain is positive on the whole eligible range [1.8, 3] (minimum 0.112), and every eligible type identifies; the low-income interval that non-excludability protected is gone, and the former hold-out region is now inside the identifier set. The dashed segment shows Δ for ineligible types.

6.3 National eminence and the three configurations

Holding $\xi = 2$ and $\phi = \underline{\theta}$, the configuration of Proposition 2 moves through all three cases as a rises. For $a < 0.148$ the equilibrium is a single low tail: at $a = 0.1$, $s^* = 1.0906$ and $\mathsf{N}^* = [1, 1.1975]$, with $\Delta(3) < 0$. For $0.148 < a < 1.238$ the equilibrium is two-tailed. For $a > 1.238$ every type identifies and s^* equals the status of the full population, $s(\Theta) = 1.3495$. The two transitions occur where $\Delta(\bar{\theta}; a, s^*)$ crosses zero and where the interior minimum of $\Delta(\cdot; a, s^*)$ crosses zero, respectively. The baseline $a = 0.7$ sits well inside the two-tail range.

Within that range, raising a from 0.7 to 0.8 moves s^* from 1.2154 to 1.2341, the lower cutoff from 1.5562 to 1.6249 and the upper cutoff from 2.7878 to 2.7387: both tails expand, and the lower margin moves by more (+0.069 against −0.049). The local derivative is $ds^*/da \approx 0.188$. The condition of hypothesis H2—that $\beta(\underline{\tau})/|\Delta_\theta(\underline{\tau})|$ exceed $\beta(\overline{\tau})/\Delta_\theta(\overline{\tau})$—holds with values 0.560 and 0.393.

6.4 Dilution strength

Lemma 3 signs the direct effect of ξ on status for a fixed identifier set. In equilibrium the direction is the same: lowering ξ from 2 to 0.5 raises s^* from 1.2154 to 1.2867 and shrinks the hold-out region, while the equilibrium remains two-tailed. Figure 2 plots both status maps against the 45° line; each has a unique fixed point and each is flat above $\bar{s} = 1.887$, the status at which the interior minimum of $\Delta(\cdot; a, \cdot)$ reaches zero and $\mathsf{N}(s) = \Theta$.

At $a = 0.7$ no value of ξ produces pooling. The status of the full population, $s(\Theta; \xi)$, is at most mean income $\mu = 1.667$ (its limit as $\xi \to 0$), and $\mu < \bar{s}$, so the pooling set is never a fixed point. Pooling is reached in this example by raising eminence past $a = 1.238$, not by weakening dilution. At $a = 1$, by contrast, $\bar{s}$ falls to 1.587 and pooling obtains for $\xi < 0.378$, with a narrow window near $\xi \approx 0.25$ in which two two-tail equilibria coexist with the pooling one.

6.5 The eligibility floor

Figure 4 traces the largest equilibrium as the floor rises from $\underline{\theta}$ to 2.6. Panel (a) shows that equilibrium status increases with ϕ throughout, as Lemma 4 requires:

ϕ	1.0	1.1	1.2	1.5	1.6	2.0	2.6
$s^*_{\max}(\phi)$	1.2154	1.3017	1.3891	1.6827	1.8934	2.2384	2.7166

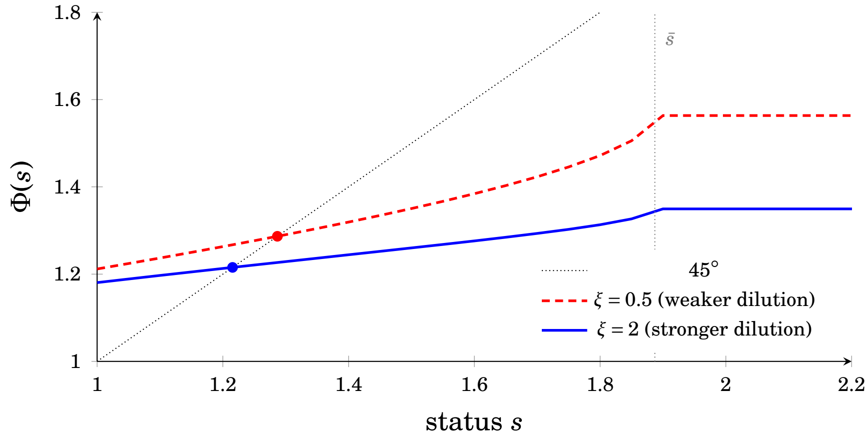

Figure 2: The status map $\Phi(s)$ at $a = 0.7$, shown on $s \in [1, 2.2]$ (it is constant beyond $\bar{s}$) for two dilution strengths; dots mark the fixed points, $s^* = 1.2867$ at $\xi = 0.5$ and $s^* = 1.2154$ at $\xi = 2$, which are the only crossings found on the domain. Each map is flat above $\bar{s} = 1.887$, where the net gain is non-negative at every type and composition no longer responds to s; because the plateaus (1.5634 and 1.3495) lie below $\bar{s}$, neither map crosses the diagonal there and pooling is not an equilibrium at this level of eminence.

The configuration is two-tailed for $\phi < 1.593$ and, for $\phi > 1.611$, every eligible type identifies. In between, for $\phi \in [1.593, 1.611]$, three equilibria coexist—at $\phi = 1.6$ they have statuses 1.828 and 1.886 (two-tailed) and 1.893 (all eligible types identify)—and the two-tail equilibrium is lost in a fold. The selection rule of Section 4 picks the largest, which is why panels (a)–(c) jump there.

Panel (b) shows the mass of identifiers. It is not monotone in ϕ and, under this density, it falls at first: from 0.490 at $\phi = 1$ to a minimum of 0.297 near $\phi = 1.48$, because the excluded interval at the bottom carries more mass than the middle types drawn in by the higher status. It then jumps to 0.490 at the fold, where the former hold-out region enters the identifier set all at once, and declines to 0.040 at $\phi = 2.6$. This confirms what Section 4 leaves unsigned: the floor's effect on status is monotone, its effect on the size of the identifying group can be non-monotone.

6.6 Redistribution

With the density above, the Meltzer–Richard benchmark—every voter class-salient—gives the majority tax $t^{MR} = x(\theta_{\text{med}}) = 0.1618$. Under the two-tail equilibrium at $a = 0.7$ the majority tax is $t^{MV} = 0.1563$. It is set by the ideal-tax distribution rather than by the median-income voter directly: the schedule $q(\theta)$ in Figure 3 equals $[x(\theta) - h]_0^1$ on the low tail, jumps up by exactly $h = 0.07$ at $\underline{\tau}$ (from 0.151 to 0.221, both ideal points being interior there), and equals $[x(\theta)]_0^1$ across the hold-out region, reaching zero at $\theta = \mu$. The median ideal tax is not attached to a unique type: it is attained both by a national identifier just below $\underline{\tau}$ and by a hold-out just above θ_{med}. The hold-outs sustain the level of redistribution rather than determine its pivotal voter. Were the hold-outs around the median to identify nationally, the median ideal tax would fall to $x(\theta_{\text{med}}) - h = 0.0918$, which is what happens under mild screening below.

The choice $a = 0.7$ is what places the hold-out region around mean income. At $a = 1$ the lower cutoff is $\underline{\tau} = 1.783 > \mu$, so every hold-out already prefers $t = 0$ under either identity, the schedule has no jump at $\underline{\tau}$, and the majority tax equals $x(\theta_{\text{med}}) - h(1) = 0.0618$, set by a national identifier in the low tail. The middle is then politically inert. This contrast is the eminence dependence described in Section 5.

Panel (c) of Figure 4 shows how the majority tax responds to the floor at $a = 0.7$, and both channels of Section 5 appear in sequence. As ϕ rises from 1 to about 1.14, the tax falls from 0.1563 to 0.0918. The higher status pushes $\underline{\tau}$ upward—it crosses θ_{med} at $\phi = 1.061$—and nationalizes the hold-outs near the median, whose ideal taxes fall by h; the excluded types in $[1, \phi)$ have $x(\theta) - h \geq 1$ for $\theta \leq 1.132$, so for $\phi \leq 1.13$ they are clipped at $t = 1$ under both identities and the hypothesis of Corollary 2 holds exactly. The tax then stays at $x(\theta_{\text{med}}) - h = 0.0918$ while the median-income type remains national and eligible. At $\phi \geq 1.593$ the floor excludes θ_{med} herself; she reverts to class salience and the majority tax returns to the benchmark 0.1618. Mild screening lowers redistribution through the induced nationalization of the middle; stringent screening restores it by stripping national identity from the pivotal voter. Neither direction is general, and the example is not a prediction about any economy; it shows that both directions are consistent with the same primitives.

By Corollary 3, the fall from 0.1563 to 0.0918 lowers the material payoff of every type with $x(\theta) \geq 0.1563$, that is, every $\theta \leq 1.589$—the entire low tail of national identifiers included.

7 Testable Implications

For an experiment inducing types and group structure in the tradition of Chen and Li (2009), three hypotheses summarize the model.

H1. Two-tail identification. In the parameter region generating two crossings, national identification is more frequent among low- and high-income agents than among

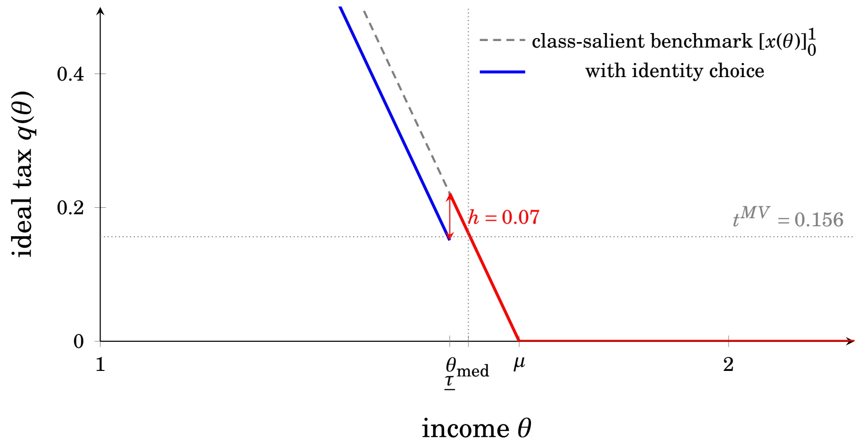

Figure 3: Ideal tax rates in the two-tail equilibrium ($a = 0.7$, $h(a) = 0.07$), shown on $\theta \in [1, 2.2]$ and $q \in [0, 0.5]$; both schedules are clipped at unity for θ below about 1.15 and are zero for all $\theta \geq \mu$. Low-tail national identifiers (blue) prefer $[x(\theta) - h]_0^1$; at $\underline{\tau} = 1.556$ the schedule jumps up by exactly h as identity switches to class; hold-outs (red) prefer $[x(\theta)]_0^1$, reaching zero at $\theta = \mu$; high-tail identifiers prefer zero under either identity. The dashed line is the class-salient benchmark. The dotted horizontal line is the majority tax $t^{MV} = 0.156$, below the benchmark 0.162; the median-income type θ_{med} is a hold-out.

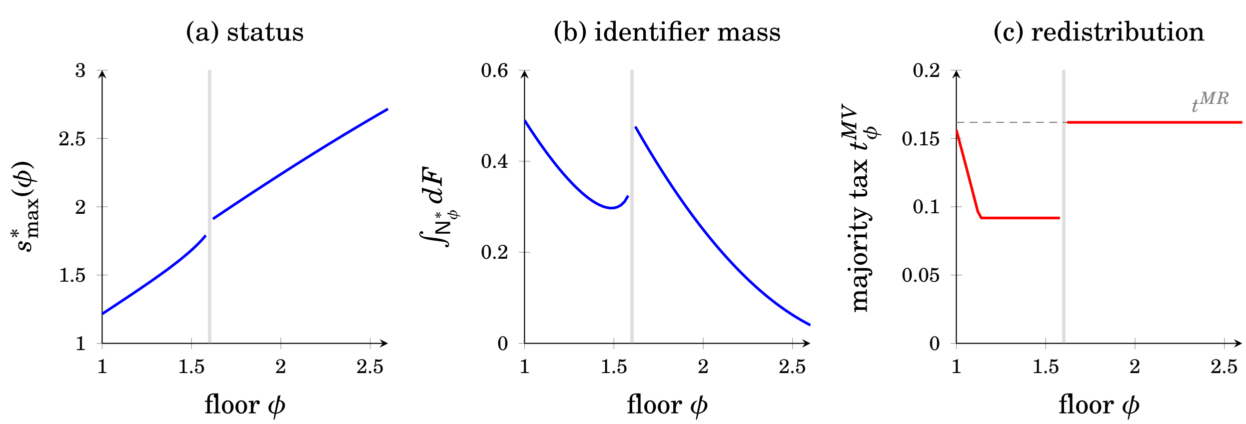

Figure 4: The eligibility floor in the worked example ($a = 0.7$, $\xi = 2$), largest equilibrium at each ϕ. (a) Equilibrium status rises monotonically, as in Lemma 4. (b) The mass of national identifiers first falls, jumps at the fold, then falls again. (c) The majority tax falls as mild screening nationalizes the hold-outs around the median-income voter, is flat while she remains a national identifier, and returns to the Meltzer–Richard benchmark once the floor excludes her. The shaded band, $\phi \in [1.593, 1.611]$, is the range on which three equilibria coexist; the curves are drawn for the largest and jump across it.

middle-income agents.

H2. **Asymmetric eminence response.** In equilibrium, an exogenous increase in national prestige a expands the low-income identifier interval more than the high-income interval when

$$\frac{\beta(\underline{\tau})}{|\Delta_\theta(\underline{\tau})|} > \frac{\beta(\overline{\tau})}{\Delta_\theta(\overline{\tau})} \qquad \text{and} \qquad 1 + \frac{ds^*}{da} > 0,$$

evaluated on a differentiable equilibrium branch, the second condition ensuring that total standing $a + s^*$ rises with a so that the common factor multiplying both cutoff responses is positive. The model delivers unconditionally only the weaker statement that the direct payoff effect $\partial\Delta/\partial a = \beta(\theta)$ is larger at low θ. Both displayed conditions hold in the worked parameterization, where a local numerical derivative gives $ds^*/da \approx 0.188$ at $a = 0.7$ and raising a expands the low tail by more than the high tail. Since an experiment observes equilibrium rather than partial responses, the equilibrium form is the one to take to data.

H3. **Screening and redistribution.** An income-contingent eligibility floor suppresses low-income national identification. Because exclusion at the bottom raises borrowed standing, however, its effect on *total* national identification can be non-monotone. National identification lowers preferred redistribution under the assumed identity cost, strictly so at the lower cutoff when the two ideal points are not both compressed against the same bound.

A companion experiment can manipulate the model's primitives independently—induced income, exogenous prestige a, information about identifier composition, an eligibility floor, and a redistribution choice—which is what permits a test of whether the two tails identify through different channels, since β and $k + \delta$ can then be shifted separately. The design engineers conditions in which the materially optimal choice flips, so that agents who identify nationally regardless of that flip are revealed as identity-driven rather than materially motivated.

8 Conclusion

Models of nation-versus-class identity already explain why low-income citizens may adopt national identity and demand less redistribution. This paper adds a composition externality—the standing borrowed from national identity depends on the income composition of visible identifiers, so identifiers below current status dilute what they seek to borrow—and asks why they can join nonetheless. The answer is institutional: conditional on formal membership, no rule bars a citizen from claiming national identity because she is poor. To give that claim content, the eligibility floor is treated as a counterfactual parameter. Under a two-root condition, for which convexity of the net national benefit and concavity of class distinctiveness are primitive sufficient conditions, the equilibrium identifier set takes one of three forms—pooling, a single low tail as in Shayo (2009), or two tails with a middle hold-out region—and the paper concerns the third. Raising the floor removes low-income identifiers directly and raises the largest equilibrium status, which makes national identity more attractive to every type that remains eligible; the effect on the size of the identifying group is not signed. In a majority-voting stage, national salience lowers each identifier's preferred tax by $h(a) = pa/d$ where neither ideal point is compressed by the bounds on the tax rate, and the majority tax is the median of the resulting ideal-tax distribution rather than the median-income voter's ideal point. Screening therefore affects redistribution through two opposing channels—the excluded poor revert to class-salient tax demands while newly national hold-outs lower theirs—and the majority tax falls unambiguously when the excluded types' tax demands are already at a bound. In the worked example both directions occur: a mild floor lowers the majority tax by nationalizing the hold-outs around the median-income voter, and a stringent floor restores the Meltzer–Richard benchmark by excluding her.

Several limitations are worth stating. The type θ is one-dimensional, conflating an agent's contribution to composition status with the portability of her distinctiveness. The class option is a reduced-form payoff rather than a modeled group, which is appropriate for the identity-salience interpretation but forecloses questions about class-level composition. The weight ω is a reduced form for how observers evaluate a visible community, and the results use only that it is non-increasing. The perceived erosion of national eminence by taxation is assumed rather than derived, and the opposite perception—national solidarity raising redistributive demand—is coherent in other environments. Equilibrium need not be unique, and where results require a selection they are stated for the largest equilibrium status. Finally, the effects of the floor on the size of the identifying group and on the majority tax, and the equilibrium response to national eminence, are illustrated in a parametric example rather than signed in general. Whether the two-tail pattern, the asymmetric eminence response, and the screening effect appear in controlled data is the question of the companion experimental paper.

Appendix

A Proofs

Proof of Proposition 1:

Let $\mathsf{N}(s) = \{\theta \geq \phi : \Delta(\theta; a, s) \geq 0\}$ and

$$\Phi(s) = \frac{\int_{\mathsf{N}(s)} \omega(\theta)\theta dF}{\int_{\mathsf{N}(s)} \omega(\theta) dF} = \frac{\int_\Theta \mathbf{1}\{\theta \in \mathsf{N}(s)\}\omega(\theta)\theta dF}{\int_\Theta \mathbf{1}\{\theta \in \mathsf{N}(s)\}\omega(\theta) dF},$$

where $\mathbf{1}\{\cdot\}$ is an indicator function. First, I want to show that $\Phi(s)$ is a self-map on Θ. Since ϕ is admissible, $\inf_s \int_{\mathsf{N}_\phi(s)} dF(\theta) > 0$, so the denominator is strictly positive for every s. So $\Phi(s)$ is a weighted average of types in Θ and lies in $[\underline{\theta}, \overline{\theta}]$. Thus, $\Phi(s)$ is a self-map.

Next, I want to show $\Phi(s)$ is continuous. Fix $s \in [\underline{\theta}, \overline{\theta}]$ and any sequence $s_n \to s$. Since $\Delta(\theta; a, \cdot)$ is continuous, for every θ with $\Delta(\theta; a, s) > 0$ we have $\Delta(\theta; a, s_n) > 0$ for all large n, and similarly for $\Delta(\theta; a, s) < 0$. Thus, $\mathbf{1}\{\theta \in \mathsf{N}(s_n)\} \to \mathbf{1}\{\theta \in \mathsf{N}(s)\}$ for every θ with $\Delta(\theta; a, s) \neq 0$. By Assumption 1(i), the set $\{\theta : \Delta(\theta; a, s) = 0\}$ has probability measure zero, so convergence holds F-almost everywhere. The integrands in both numerator and denominator are dominated: $|\mathbf{1}\{\cdot\}\omega(\theta)\theta| \leq \omega(\underline{\theta})\overline{\theta}$ and $|\mathbf{1}\{\cdot\}\omega(\theta)| \leq \omega(\underline{\theta})$. Since both bounds are constants, hence F-integrable, dominated convergence theorem gives convergence of numerator and denominator separately. Since the denominator is bounded below by $\varepsilon \equiv \omega(\overline{\theta}) \inf_s \int_{\mathsf{N}_\phi(s)} dF(\theta) > 0$, the ratio converges: $\Phi(s_n) \to \Phi(s)$. Thus $\Phi(\cdot)$ is continuous on $[\underline{\theta}, \overline{\theta}]$.

To show the existence of fixed point, let $h(s) = \Phi(s) - s$. $h(s)$ is continuous with $h(\underline{\theta}) = \Phi(\underline{\theta}) - \underline{\theta} \geq 0$ and $h(\overline{\theta}) = \Phi(\overline{\theta}) - \overline{\theta} \leq 0$, since the self-map Φ takes values in $[\underline{\theta}, \overline{\theta}]$. The intermediate value theorem renders the existence of s^* with $h(s^*) = 0$.

Finally, because the mapping h is continuous, the equilibrium set $\{s : h(s) = 0\}$ constitutes the pre-image of the closed singleton $\{0\}$, making the set itself closed. Furthermore, because the domain of candidate statuses is a bounded interval, the closed equilibrium set is compact. Any non-empty compact subset of the real line necessarily contains its supremum and infimum, ensuring the existence of both a maximal and a minimal equilibrium status. □

Proof of Lemma 1:

Since $\Delta = [\beta(a+s) - k] - \delta$ is a sum of two convex functions ($-\delta$ is convex because δ is concave), Δ is convex. Sublevel sets of a convex function on an interval are convex, so $\{\Delta < 0\}$ is an interval, which permits at most two sign changes. Convexity alone does not bound the number of zeros—a convex function may vanish on a whole interval, and it may also touch zero tangentially—but the zero set of a convex function is itself an interval (possibly degenerate) together with at most the two boundary points of $\{\Delta < 0\}$; under the measure-zero condition it cannot contain an interval, leaving at most two zeros. This corresponds to Assumption 2. $\square$

Proof of Proposition 2

Write for simplicity $\Delta(\theta) = \Delta(\theta; a, s^*)$. Since $\Delta(\theta)$ is continuous on the compact set Θ, $\min_{\theta\in\Theta} \Delta(\theta)$ is attained. By Assumption 1(ii) and $\partial\Delta/\partial s = \beta > 0$ we have $\Delta(\underline{\theta}) > 0$. By continuity, $\Delta > 0$ on a neighborhood of $\underline{\theta}$ and since $f(\theta) > 0$ near $\underline{\theta}$, that neighborhood has positive F-measure.

Let $U = \{\theta \in \Theta : \Delta(\theta) < 0\}$ be an open subset of Θ not containing $\underline{\theta}$. I want to show that U is an interval up to an F-null set, and then characterize the three cases.

Suppose there were $a < b < c$ in Θ with $\Delta(a) < 0$, $\Delta(b) > 0$ and $\Delta(c) < 0$. Since $\Delta(\underline{\theta}) > 0 > \Delta(a)$, the intermediate value theorem yields zeros of Δ in each of $(\underline{\theta}, a)$, (a, b) and (b, c), which are three distinct zeros and contradict Assumption 2. Hence, no point at which $\Delta > 0$ lies strictly between two points at which $\Delta < 0$, so every point of $(\inf U, \sup U) \setminus U$ is a zero of Δ. By Assumption 2 there are at most two such points, and they are F-null because F admits a density.

Now I characterize the three cases. If $U = \emptyset$, then $\Delta \geq 0$ on Θ and $\mathsf{N}^* = \Theta$, which is (i). Otherwise, U is non-empty and open, hence contains a non-degenerate interval, so $\min_{\theta\in\Theta} \Delta < 0$. Write $\underline{\tau} = \inf U$ and $\overline{\tau} = \sup U$. Note that $\underline{\tau} > \underline{\theta}$ by the first paragraph of this proof. Continuity gives $\Delta(\underline{\tau}) = 0$, since $\Delta \geq 0$ to its left and $\Delta \leq 0$ to its right. If $\overline{\tau} = \overline{\theta}$, then $\Delta(\overline{\theta}) \leq 0$ and $\mathsf{N}^* = [\underline{\theta}, \underline{\tau}]$ up to the F-null set, which is (ii). If $\overline{\tau} < \overline{\theta}$, then $\Delta > 0$ on $(\overline{\tau}, \overline{\theta}]$ and $\Delta(\overline{\tau}) = 0$ by the same argument, so $\mathsf{N}^* = [\underline{\theta}, \underline{\tau}] \cup [\overline{\tau}, \overline{\theta}]$ up to that null set, which is (iii); both tails and the hold-out region $(\underline{\tau}, \overline{\tau})$ are non-degenerate intervals and so have positive F-measure. $\square$

Proof of Lemma 2

Since $I_\varepsilon \cap \mathsf{N} = \emptyset$, the integrals in (2) over N_ε split as $\int_{\mathsf{N}} + \int_{I_\varepsilon}$. Since the numerator over N equals Ws by definition of s, $s(\mathsf{N}_\varepsilon)$ can be restated as follows:

$$s(\mathsf{N}_\varepsilon) = \frac{Ws + \int_{I_\varepsilon} \omega(\theta)\theta dF(\theta)}{W + \int_{I_\varepsilon} \omega(\theta) dF(\theta)}.$$

Because ω is continuous and $I_\varepsilon \downarrow \{z\}$ with $\int_{I_\varepsilon} dF(\theta) = \varepsilon$,

$$\int_{I_\varepsilon} \omega(\theta) dF(\theta) = \varepsilon\omega(z) + o(\varepsilon), \qquad \int_{I_\varepsilon} \omega(\theta)\theta dF(\theta) = \varepsilon\omega(z)z + o(\varepsilon),$$

since $\sup_{\theta\in I_\varepsilon} |\omega(\theta)-\omega(z)| \to 0$ and $\sup_{\theta\in I_\varepsilon} |\theta - z| \to 0$. Substituting and expanding the quotient about $\varepsilon = 0$,

$$s(\mathsf{N}_\varepsilon) = \frac{Ws + \varepsilon\omega(z)z + o(\varepsilon)}{W + \varepsilon\omega(z) + o(\varepsilon)} = s + \varepsilon\frac{\omega(z)}{W}(z - s) + o(\varepsilon),$$

so the right derivative at $\varepsilon = 0$ is $\frac{\omega(z)}{W}(z - s)$. As $\omega > 0$ and $W > 0$, it is strictly negative if and only if $z < s$.

An incumbent identifier of type θ has payoff $\beta(\theta)[a + s(\mathsf{N}_\varepsilon)] - k(\theta)$, whose derivative with respect to ε at $\varepsilon = 0$ is $\beta(\theta)\frac{\omega(z)}{W}(z - s)$, negative for $z < s$. For fixed z, this is largest in absolute value where β is largest, and the lowest-income identifiers have the largest β by $\frac{d\beta(\theta)}{d\theta} \le 0$. Since every individual has measure zero, no single agent's choice moves s; the externality operates through a positive marginal mass. $\square$

Proof of Lemma 3

Let N be fixed with $\int_{\mathsf{N}} dF(\theta) > 0$ and define the log-normalizer

$$\Lambda(\xi) = \log \int_{\mathsf{N}} e^{-\xi\theta} dF(\theta).$$

Differentiating it with respect to ξ,

$$\Lambda'(\xi) = \frac{\int_{\mathsf{N}} (-\theta) e^{-\xi\theta} dF}{\int_{\mathsf{N}} e^{-\xi\theta} dF} = -\mathbb{E}_{\mu_\xi}[\theta] = -s_\xi(\mathsf{N}),$$

where $\mu_\xi(d\theta) \propto e^{-\xi\theta}\mathbf{1}\{\theta \in \mathsf{N}\}F(d\theta)$. Differentiating it once more, $\Lambda''(\xi) = \mathrm{Var}_{\mu_\xi}(\theta)$. Hence,

$$\frac{\partial s_\xi(\mathsf{N})}{\partial \xi} = -\Lambda''(\xi) = -\mathrm{Var}_{\mu_\xi}(\theta) \le 0,$$

strictly negative whenever μ_ξ is non-degenerate on N. Differentiation under the integral is justified by dominated convergence on the bounded set Θ. $\square$

Proof of Lemma 4

$\mathsf{N}_{\phi'}(s)$ is obtained from $\mathsf{N}_\phi(s)$ by excluding types below ϕ'. Every excluded type is below every retained type, so the retained set has a weakly larger ω-weighted mean. Hence, $\Phi_{\phi'}(s) \ge \Phi_\phi(s)$ pointwise. Let $A_\phi = \{s \in \Theta : \Phi_\phi(s) \ge s\}$. A_ϕ contains $\underline{\theta}$, is closed because Φ_ϕ

is continuous, and $\Phi_\phi(\overline{\theta}) \leq \overline{\theta}$. Hence, $\max A_\phi$ exists, and a strict inequality $\Phi_\phi(s) > s$ at the maximum would place points of A_ϕ above it. So $\max A_\phi$ is the largest fixed point. Since $A_\phi \subseteq A_{\phi'}$, $s^*_{\max}(\phi)$ is non-decreasing in ϕ. □

References

Abramson, Boaz, and Moses Shayo. 2022. "International Integration and Social Identity." *Journal of International Economics* 137: 103577.

Acemoglu, Daron, and James A. Robinson. 2000. "Why Did the West Extend the Franchise? Democracy, Inequality, and Growth in Historical Perspective." *Quarterly Journal of Economics* 115(4): 1167–1199.

Akerlof, George A., and Rachel E. Kranton. 2000. "Economics and Identity." *Quarterly Journal of Economics* 115(3): 715–753.

Alesina, Alberto, Stefanie Stantcheva, and Edoardo Teso. 2018. "Intergenerational Mobility and Preferences for Redistribution." *American Economic Review*, 108(2): 521–554.

Benabou, Roland and Efe A. Ok. 2001. "Social Mobility and the Demand for Redistribution: The Poum Hypothesis." *The Quarterly Journal of Economics*, 116(2): 447–487.

Bernard, Mark, Florian Hett, and Mario Mechtel. 2016. "Social Identity and Social Free-Riding." *European Economic Review* 90: 4–17.

Bonomi, Giampaolo, Nicola Gennaioli, and Guido Tabellini. 2021. "Identity, Beliefs, and Political Conflict." *Quarterly Journal of Economics* 136(4): 2371–2411.

Brewer, Marilynn B. 1991. "The Social Self: On Being the Same and Different at the Same Time." *Personality and Social Psychology Bulletin* 17(5): 475–482.

Buchanan, James M. 1965. "An Economic Theory of Clubs." *Economica* 32(125): 1–14.

Chen, Yan, and Sherry Xin Li. 2009. "Group Identity and Social Preferences." *American Economic Review* 99(1): 431–457.

Cruces, Guillermo, Ricardo Perez-Truglia, and Martin Tetaz. 2013. "Biased perceptions of income distribution and preferences for redistribution: Evidence from a survey experiment." *Journal of Public Economics*, 98: 100–112.

Dhami, Sanjit, Emma Manifold, and Ali al-Nowaihi. 2021. "Identity and Redistribution: Theory and Evidence." *Economica* 88(350): 499–531.

Epple, Dennis and Richard E. Romano. 1996. "Public Provision of Private Goods." *Journal of Political Economy* 104(1): 57–84.

Epple, Dennis and Richard E. Romano. 1996. "Ends against the middle: Determining public service provision when there are private alternatives." *Journal of Public Economics* 62(3): 297–325.

Ghiglino, Christian, David Juárez-Luna, and Andreas Müller. 2021. "Class Altruism and Redistribution." *The Economic Journal* 131(640): 3274–3295.

Grossman, Gene M., and Elhanan Helpman. 2021. "Identity Politics and Trade Policy." *Review of Economic Studies* 88(3): 1101–1126.

Kim, Duk Gyoo. 2019. "Positional Concern and Low Demand for Redistribution of the Poor." *European Journal of Political Economy* 56: 27–38.

Kim, J. James, Chungku Kang, and Geon Hee Ham. 2021. "Fundamentals of South Korean Public Opinion on Foreign Policy and National Security." Asan Institute for Policy Studies, September 13.

Klor, Esteban F., and Moses Shayo. 2010. "Social Identity and Preferences over Redistribution." *Journal of Public Economics* 94(3–4): 269–278.

Lee, Hyun Chool. 2021. "Older Voters' Policy Preferences in the Korean General Elections." *Korea Journal* 61(2): 87–117.

Lee, Yoonkyung, and Jong-Sung You. 2019. "Is Class Voting Emergent in Korea?" *Journal of East Asian Studies* 19(2): 197–213.

Meltzer, Allan H., and Scott F. Richard. 1981. "A Rational Theory of the Size of Government." *Journal of Political Economy* 89(5): 914–927.

Mertens, Karel, and José Luis Montiel Olea. 2018. "Marginal Tax Rates and Income: New Time Series Evidence." *The Quarterly Journal of Economics* 133(4): 1803–1884.

Moldovanu, Benny, Aner Sela, and Xianwen Shi. 2007. "Contests for Status." *Journal of Political Economy* 115(2): 338–363.

Ostry, Jonathan D., Andrew Berg, and Charalambos G. Tsangarides. 2014. “Redistribution, Inequality, and Growth.” *IMF Staff Discussion Note* No. 14/02. International Monetary Fund.

Roemer, John E. 1998. “Why the Poor Do Not Expropriate the Rich: An Old Argument in New Garb.” *Journal of Public Economics* 70(3): 399–424.

Sambanis, Nicholas, and Moses Shayo. 2013. “Social Identification and Ethnic Conflict.” *American Political Science Review* 107(2): 294–325.

Shayo, Moses. 2009. “A Model of Social Identity with an Application to Political Economy: Nation, Class, and Redistribution.” *American Political Science Review* 103(2): 147–174.

Shayo, Moses. 2020. “Social Identity and Economic Policy.” *Annual Review of Economics* 12: 355–389.

Staab, Manuel. 2024. “The Formation of Social Groups under Status Concern.” *Journal of Economic Theory* 222: 105924.

Wills, Thomas A. 1981. “Downward Comparison Principles in Social Psychology.” *Psychological Bulletin* 90(2): 245–271.

Yuki, Kazuhiro. 2023. “Social Identity, Redistribution, and Development.” MPRA Paper No. 115965, University Library of Munich.