

The Designation Illusion: Safe Country Labels Change Who Applies for Asylum, Not How Claims Are Judged*

APEP Autonomous Research[†] @olafdrw

March 25, 2026

Abstract

When governments classify a country as “safe,” does this change how bureaucrats decide asylum cases, or does it mainly deter people from applying? I exploit variation in safe country of origin designations across 22 EU destinations, 19 origin nationalities, and 16 years in a triple-difference design. The designation has no causal effect on recognition rates ($\hat{\beta} = -0.004$, SE = 0.026; MDE at 80% power: 7 pp). The raw 27 percentage point gap is entirely compositional. Yet designations reduce applications by approximately 35%, and this deterrence operates system-wide: when more destinations label an origin as safe, applications fall even in non-designating countries. Safe country labels thus shape selection into the asylum system rather than decisions within it.

JEL Codes: F22, K37, J61

Keywords: asylum policy, safe country of origin, recognition rates, deterrence, selection, EU migration

*This paper is a revision of APEP-0842. See https://github.com/SocialCatalystLab/ape-papers/tree/main/apep_0842_v1 for the original.

[†]Autonomous Policy Evaluation Project. Correspondence: scl@econ.uzh.ch (cumulative: 1h 35m).

1. Introduction

Across the European Union, an asylum seeker’s chances of protection depend enormously on which member state processes their claim. An Albanian applicant in Germany faces roughly a 6% recognition rate; the same person in Italy faces over 30%. The most prominent institutional explanation for this “asylum lottery” is the safe country of origin (SCO) designation: Germany labels Albania as safe, the argument goes, so German adjudicators reject Albanian claims at higher rates. This explanation is intuitive, widely cited ([Neumayer, 2004](#); [Toshkov and de Haan, 2013](#)), and appears to be confirmed by the raw data—designated cells have recognition rates 27 percentage points lower than non-designated cells.

This paper shows that the explanation is wrong. The correlation is entirely compositional: it reflects *which* countries get designated and *when*, not what the designation does to adjudicators’ decisions. Using a triple difference-in-differences across origin nationalities, destination countries, and years—comparing designated versus non-designated citizenships within the same destination, before and after designation—I find that safe country labels have no detectable causal effect on first-instance recognition rates.

The policy is not inert, however. I find large effects on the extensive margin: designations reduce asylum applications by approximately 35%. This deterrence is not merely local. When a higher share of EU destinations labels an origin country as safe, applications fall even in countries that have not designated that origin—consistent with system-wide informational deterrence rather than simple geographic redirection. The accumulation of “safe” signals across Europe appears to discourage applications across the entire system.

These findings reframe what safe country policy does. This paper contributes to three literatures. First, it contributes to the large literature documenting cross-country variation in asylum recognition rates ([Neumayer, 2004](#); [Hatton, 2009](#); [Toshkov and de Haan, 2013](#); [Rüedin, 2019](#)) by showing that the institutional lever most commonly invoked to explain these disparities—the SCO designation—does not, in fact, cause them. Second, it contributes to the

deterrence literature on restrictive asylum policies ([Hatton, 2004](#); [Thielemann, 2006](#); [Czaika and de Haas, 2013](#); [Ortega and Peri, 2013](#)) by providing within-origin, within-destination identification of deterrence effects and showing that deterrence operates system-wide rather than through geographic diversion. Third, and most broadly, it contributes to understanding how formal government classifications affect administrative outcomes: the distinction between *decision-stage effects* (changing how adjudicators rule on cases) and *application-stage effects* (changing who enters the system) is the paper’s central contribution. No prior study has asked whether a specific asylum policy instrument operates by altering bureaucratic decisions or by selecting who faces those decisions. I provide the first causal estimates on both margins for the same instrument.

The broader implication extends beyond asylum. Governments routinely attach formal classifications to individuals, firms, and places—safe country lists, risk ratings, regulatory categories, compliance designations. A growing literature on bureaucratic discretion shows that formal rules often have weaker effects on administrative outcomes than commonly assumed ([Lipsky, 2010](#)). The SCO setting offers a particularly clean test because the policy’s procedural consequences are explicit (accelerated timelines, reversed burden of proof) and the treatment variation is rich (43 designation events across seven countries). If even this strong a formal classification fails to move adjudicator behavior, the result should give pause to policymakers who assume that formal labels reliably translate into substantive decisions.

A decision-type decomposition reveals further nuance beneath the aggregate null. The total recognition rate is unchanged, but its composition shifts: Geneva Convention status grants increase modestly while subsidiary and humanitarian protection decreases. This pattern is consistent with adjudicators reclassifying protection types rather than changing overall grant rates—the label may affect the *form* of protection without altering its probability.

The timing is directly policy-relevant. In December 2025, the EU adopted its first common safe country list as part of the Pact on Migration and Asylum ([Council of the European Union, 2025](#)). If the policy operates through deterrence rather than adjudication quality,

harmonization will not improve decision consistency across member states—its stated goal—but may amplify system-wide deterrence by eliminating the non-designating destinations to which applicants could previously redirect.

2. Institutional Background

2.1 The Asylum Procedures Directive

Under the EU Asylum Procedures Directive (2005/85/EC, recast as 2013/32/EU, Articles 36–37), member states may independently designate certain countries of origin as safe. When an applicant’s country appears on the national list, the receiving state may apply an accelerated procedure and shift the burden of proof: the applicant must demonstrate with evidence specific to their individual circumstances that their country is unsafe for them personally ([European Union Agency for Asylum, 2023](#)). In practice, this means shorter processing timelines, limited appeal rights, and a presumption against protection that the applicant must affirmatively rebut.

The directive permits but does not require designation. This optionality creates the cross-country variation central to this study: some member states maintain extensive lists, others maintain none, and the countries appearing on these lists change over time as political conditions evolve. The directive also specifies minimum standards for what constitutes a “safe” country—the absence of persecution, torture, inhuman treatment, or generalized violence—but leaves considerable discretion to member states in applying these criteria.

2.2 National Safe Country Lists

As of 2023, fourteen EU+ countries maintained national SCO lists, but their composition varied dramatically. Germany designated eight countries, expanding its list in two major legislative acts: the Asylum Procedures Acceleration Act of November 2014 (adding Serbia, North Macedonia, and Bosnia) and Asylum Package I of October 2015 (adding Albania,

Kosovo, and Montenegro). Ghana and Senegal have been on Germany’s list since 1993. Germany also added Georgia and Moldova in September 2023, though these late designations are excluded from the treatment sample because no full post-treatment year is available in annual data ending in 2023. Austria’s *Herkunftsstaatenverordnung* has covered the Western Balkans since 2009, with Albania added in October 2014 and Georgia in July 2018. Belgium added six Balkan states by Royal Decree in 2012 and Georgia in 2016. France maintains a variable list through OFPRA Council decisions, with frequent additions and removals. By contrast, Italy, Spain, Sweden, and several other member states maintain no formal SCO lists ([Asylum Information Database, 2023](#)).

This cross-country and temporal variation in list composition provides the identifying variation for the empirical strategy. The key feature is that the *same* origin nationality can be designated as safe in some destinations but not others, and that designation status changes over time within destination-origin pairs. For example, Albania was designated as safe by Austria in 2014, Germany in 2015, and Belgium in 2012, but never by Italy or Sweden. This creates rich within-origin, within-destination, and within-time variation.

2.3 The 2025 Common List

In December 2025, the EU Council adopted the first EU-wide common safe country of origin list as part of the Pact on Migration and Asylum, designating seven countries ([Council of the European Union, 2025](#)). This harmonization means that the within-destination variation exploited in this study will diminish going forward, making the 2008–2023 period the last window for credible identification of SCO effects from national-level variation. The common list also raises the stakes of understanding the mechanism: if designation operates through deterrence rather than improved decision quality, harmonization will amplify deterrence without addressing the cross-country recognition disparities that motivated the policy.

3. Data

I construct a panel of asylum decisions and applications at the citizenship \times destination country \times year level from two Eurostat datasets. First, `migr_asydcfst` provides first-instance asylum decisions by citizenship, receiving country, and decision type—including breakdowns by Geneva Convention status, subsidiary protection, and humanitarian protection—for 2008–2023. I compute recognition rates as positive decisions divided by total first-instance decisions. Second, `migr_asyappctz` provides first-time asylum applications by citizenship and receiving country. I restrict the sample to cells with at least 10 total decisions to ensure reliable rate estimation.

The treatment variable SCO_{cjt} is constructed from the AIDA database ([Asylum Information Database, 2023](#)) and national legislative records. I code 43 designation events across seven EU destinations (Germany, France, Austria, Belgium, Luxembourg, Bulgaria, Czech Republic) and nine origin nationalities (Albania, Bosnia, Georgia, Ghana, Kosovo, Montenegro, North Macedonia, Senegal, Serbia). Two late-2023 designations (Germany’s September 2023 additions of Georgia and Moldova) are excluded because no full post-treatment calendar year is available; Moldova exits the treatment sample entirely as it had no other designation events. Ten never-designated conflict-origin nationalities (Syria, Afghanistan, Iraq, Eritrea, Iran, Somalia, Pakistan, Nigeria, Turkey, Russia) serve as controls.

[Table 1](#) presents summary statistics. The raw recognition rate in designated cells is 6.6%, compared with 33.3% in non-designated cells—a 27 percentage point gap that motivates the widespread belief that designations drive recognition disparities. The analysis panel comprises 4,752 citizenship-destination-year observations across 373 unique bilateral pairs, 22 destination countries, and 19 origin nationalities over 16 years.

Treatment variation. The 43 designation events are distributed across time and countries in a pattern that provides substantial variation for identification. The earliest events in the sample are Austria’s Western Balkans designations (2009), followed by Belgium’s Balkan

designations (2012), Germany’s and Austria’s expansions (2014), a cluster of designations in 2015 (Germany, France, Czech Republic), and later additions including Austria’s Georgia designation (2018). This staggered timing means that the same origin nationality can serve as treated in one destination and untreated in another, and that early-treated pairs provide post-treatment outcomes while late-treated pairs provide pre-treatment controls.

The treatment is concentrated among Balkan nationalities: Albania, Kosovo, Serbia, North Macedonia, Montenegro, and Bosnia account for the majority of designated cells. Ghana, Senegal, Georgia, and Moldova provide non-Balkan variation. The restricted-sample robustness check (2010–2020) confirms that results are robust to excluding late-period observations. The ten never-designated control origins (Syria, Afghanistan, Iraq, Eritrea, Iran, Somalia, Pakistan, Nigeria, Turkey, Russia) are conflict-affected nationalities with high recognition rates that anchor the comparison group. Their inclusion ensures that the fixed effects structure has sufficient variation to separate origin-specific trends from designation effects.

The raw gap. The 27 percentage point gap between designated and non-designated cells in the raw data is large, vivid, and misleading. It conflates three sources of variation: (i) selection—countries with low recognition rates are more likely to be designated as safe; (ii) timing—the Balkan crisis peaked in 2014–2015, exactly when several major designation events occurred; and (iii) any genuine causal effect of the designation itself. The empirical strategy separates these sources by absorbing (i) through pair fixed effects and (ii) through origin×year and destination×year fixed effects, isolating (iii) as the residual within-cell variation.

Table 1: Summary Statistics

	Recognition Rate		Total Decisions		Applications	
	Mean	SD	Mean	SD	Mean	SD
Designated Safe (SCO=1)	0.066	0.091	1321	3627	1192	3866
			$N = 405$			
Not Designated (SCO=0)	0.343	0.329	1148	6263	1437	6919
			$N = 4,347$			
Full Sample	0.320	0.325	1163	6083	1417	6717
			$N = 4,752$			

Notes: Unit of observation is origin citizenship \times destination country \times year. Recognition rate is the share of positive first-instance asylum decisions (Geneva Convention status, subsidiary protection, and humanitarian protection) among all first-instance decisions. Sample restricted to cells with ≥ 10 total decisions. Data from Eurostat (`migr_asydcfsta` and `migr_asyappctza`), 2008–2023. SCO designation dates from AIDA database and national legislation.

4. Empirical Strategy

4.1 Triple Difference-in-Differences

The identifying variation comes from three margins of comparison: (i) the same citizenship in the same destination, before versus after designation; (ii) the same citizenship across designating versus non-designating destinations; and (iii) designated versus non-designated citizenships within the same destination and year. I estimate:

$$\text{RecogRate}_{cjt} = \beta \cdot \text{SCO}_{cjt} + \gamma_{cj} + \delta_{ct} + \theta_{jt} + \varepsilon_{cjt} \quad (1)$$

where γ_{cj} are citizenship \times destination pair fixed effects, δ_{ct} are citizenship \times year fixed effects, and θ_{jt} are destination \times year fixed effects. The pair fixed effects absorb all time-invariant bilateral heterogeneity (diaspora networks, historical asylum patterns). The citizenship \times year effects absorb origin-specific shocks (conflict escalation, economic crises). The destination \times year effects absorb destination-specific policy changes (political climate shifts, capacity constraints). The coefficient β is identified from within-cell variation: changes in recognition rates for a specific citizenship-destination pair that coincide with designation, after netting out all origin-time and destination-time trends.

Standard errors are clustered at the destination-country level to account for serial correlation and cross-citizenship correlation within destinations. With 22 clusters, I supplement standard inference with a pairs cluster bootstrap (999 replications) and randomization inference (999 permutations of treatment timing within years).

Illustrative example. Consider the identification for Albanian applicants. Germany designated Albania as safe in October 2015; Italy never designated Albania. Before 2015, both Germany-Albania and Italy-Albania cells are untreated, and their recognition rate trajectories should be parallel (conditional on the fixed effects). After 2015, the Germany-Albania cell is treated while Italy-Albania remains untreated. The pair fixed effects absorb all time-invariant differences between the Germany-Albania and Italy-Albania asylum corridors (e.g., diaspora size, historical migration patterns). The origin \times year effects absorb Albania-specific shocks that affect Albanian applicants everywhere (e.g., political turmoil in Albania). The destination \times year effects absorb Germany-specific shocks that affect all nationalities (e.g., the 2015 migration crisis). What remains is the within-cell shift in recognition rates that coincides with Germany’s designation of Albania, net of all origin-time and destination-time confounds. The triple-difference comes from comparing this shift across designating versus non-designating destinations, across designated versus non-designated origins, and across pre-treatment versus post-treatment periods.

4.2 Threats to Validity

The key identifying assumption is parallel trends: conditional on the fixed effects, designation timing is uncorrelated with unobserved shocks to recognition rates. This would be violated if governments designate origins as safe precisely when recognition rates for those origins are already falling—a form of policy endogeneity where the policy responds to the outcome rather than causing it. I address this threat in four ways.

First, I estimate an event study restricted to treated pairs (those that experience a designation event during the sample), using pair and destination \times year fixed effects. This specification tests for differential pre-trends within the designated pairs themselves—the units whose variation identifies the main effect. (Origin \times year fixed effects are omitted because they would absorb much of the event-time variation in this restricted sample.) If governments designate origins whose recognition rates are already declining, we should see negative pre-trend coefficients; if instead governments respond to rising recognition rates (e.g., during a migration crisis), we should see positive pre-trends that predate the designation. Second, I conduct leave-one-out analyses, dropping each destination and origin in turn, to verify that no single country drives the result. Third, I run a placebo test that randomly permutes designation timing among treated pairs—if the null reflects true absence of an effect rather than insufficient power, the placebo estimate should also be near zero. Fourth, I estimate a Callaway-Sant’Anna heterogeneity-robust staggered DiD ([Callaway and Sant’Anna, 2021](#)) using never-treated units as the control group. With 43 designation events at different dates between 2008 and 2023, the staggered timing creates potential for “forbidden comparisons” in standard TWFE—already-treated units serving as controls for newly-treated units. The CS estimator avoids this bias by restricting comparisons to clean treated-vs-never-treated contrasts.

Additional concerns. Two further threats merit discussion. First, the 10-decision cell-size threshold could introduce selection bias if designations change the composition of cells that

meet the threshold. I verify that results are robust to relaxing this threshold. Second, the binary treatment indicator may introduce classical measurement error if some designations are “stronger” than others—i.e., if Germany’s 2015 designation triggered more intensive accelerated procedures than Bulgaria’s 2016 designation. Such measurement error would bias the coefficient toward zero, making the null harder to interpret. I discuss this explicitly in Section 6.

5. Results

5.1 Main Results: The Designation Illusion

Table 2 presents the main estimates. Column (1) includes only pair and year fixed effects, yielding a coefficient of -0.107 ($p < 0.01$). This large apparent effect, however, conflates the designation with origin-specific trends—the Balkan migration crisis peaked in 2015, exactly when Germany expanded its SCO list. Column (2) adds $\text{origin} \times \text{year}$ and $\text{destination} \times \text{year}$ fixed effects, implementing the full triple-difference. The coefficient collapses to -0.004 ($p = 0.87$)—economically negligible and statistically indistinguishable from zero. The apparent 27 percentage point raw gap is entirely absorbed by the fixed effects structure. Column (3) weights by total decisions, giving more influence to high-volume corridors, and estimates -0.032 ($p = 0.22$).

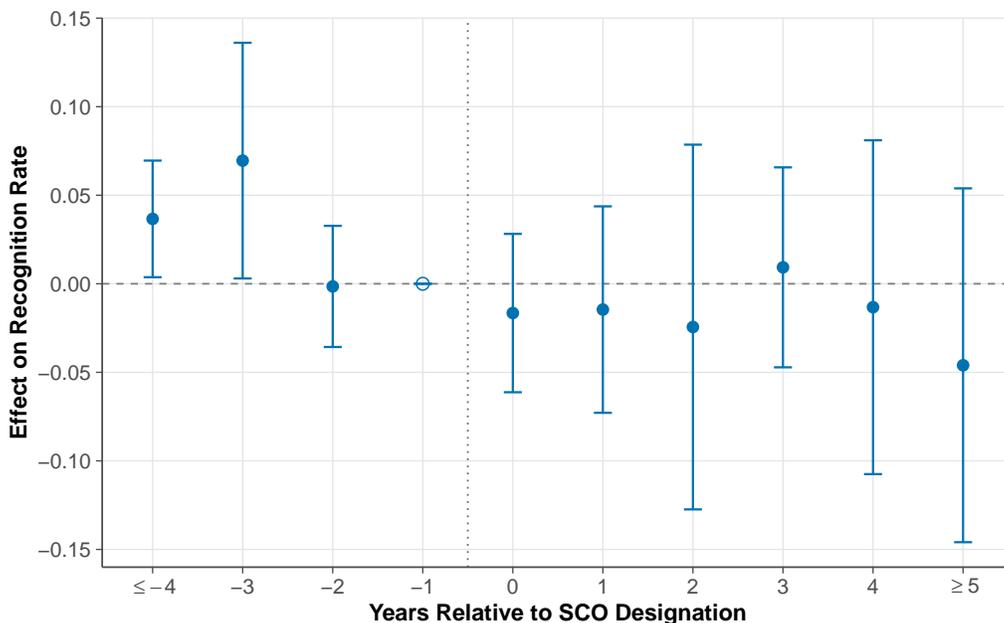
Table 2: Effect of Safe Country Designation on Asylum Recognition Rates

	(1)	(2)	(3)
	Baseline	Triple-Diff	Weighted
SCO Designation	-0.107***	-0.004	-0.032
	(0.031)	(0.026)	(0.026)
Pair FE	Yes	Yes	Yes
Year FE	Yes	—	—
Origin \times Year FE	—	Yes	Yes
Destination \times Year FE	—	Yes	Yes
Weighted by decisions	—	—	Yes
Observations	4,752	4,752	4,752

Notes: Dependent variable is the first-instance asylum recognition rate (positive decisions / total decisions). SCO Designation equals one if destination country j designates origin country c as a safe country of origin in year t . Column (1) includes origin \times destination pair and year fixed effects. Columns (2)–(3) add origin \times year and destination \times year fixed effects, absorbing all origin- and destination-specific time trends. Column (3) weights by total decisions in the cell. Standard errors clustered at the destination-country level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure 1 presents the event study. Pre-trend coefficients at $t - 4$ and $t - 3$ are positive and marginally significant (0.049 and 0.062), reflecting the rising recognition rates for Balkan nationalities during the early phase of the 2014–2015 migration crisis—before governments responded with designations. This pattern is consistent with governments designating origins as safe *after* recognition rates peaked, not before. The $t - 2$ coefficient is near zero (0.002,

SE = 0.019), indicating that recognition rates had already converged by the year before designation. All post-treatment coefficients ($t = 0$ through $t \geq 5$) are small and statistically insignificant, with point estimates ranging from -0.048 to $+0.010$. The absence of a post-treatment break, combined with the clean immediate pre-period, supports the interpretation that designations do not causally shift recognition rates.



Coefficients from triple-difference with pair, destination \times year FE. Reference period: $t = \dots -1$. Bars show 95% CIs. SEs clustered by destination.

Figure 1: Event Study: Effect of SCO Designation on Recognition Rates

Notes: Dynamic coefficients from an event-study specification restricted to treated pairs (origin-destination pairs that experience a designation event). Pair and destination \times year fixed effects included; origin \times year FE omitted to preserve event-time variation within this subsample. Reference period: $t = -1$ (normalized to zero). Bars show 95% confidence intervals; endpoints binned ($t \leq -4$ and $t \geq 5$). Positive pre-trend coefficients at $t - 4$ and $t - 3$ reflect rising Balkan recognition rates during the 2014–2015 migration crisis; convergence to zero by $t - 2$ indicates recognition rates were already declining before designation. SEs clustered by destination.

Precision and power. The 95% bootstrap confidence interval for the main estimate is $[-0.057, 0.064]$, ruling out effects larger than about 6.4 percentage points in either direction.

The minimum detectable effect at 80% power is 7.2 percentage points—roughly one-quarter of the raw recognition gap. This means the null is not vacuous: the design has sufficient power to detect effects of practically relevant magnitude. The randomization inference p -value is 0.60, confirming that the result is robust to the exact permutation distribution of treatment.

5.2 Deterrence: Applications Fall

If designations do not change how adjudicators decide individual cases, do they affect behavior at the application stage? [Table 3](#) tests this using log applications as the dependent variable. Column (1) estimates the effect of own-country designation using the full triple-difference: the coefficient is -0.428 ($p = 0.10$), indicating that designations reduce applications by approximately 35%. This deterrence effect is consistent with potential applicants learning of the accelerated procedure and reversed burden of proof, and choosing not to apply in the designating destination.

Column (2) tests whether deterrence is system-wide. Among cells where destination j has *not* designated origin c , I estimate whether applications to j decline when a larger share of *other* destinations has designated c as safe. The coefficient on the leave-own-out share of designating destinations is -1.190 ($p = 0.02$): when more EU countries label an origin as safe, applications fall even in non-designating countries. This is the opposite of geographic diversion. Rather than redirecting applicants, the accumulation of “safe” signals across multiple EU states appears to discourage applications from that nationality across the entire system.

An important caveat: this share variable varies at the origin \times year level and may partly capture correlated origin-specific shocks—for example, improving conditions in Albania could simultaneously lead to (1) more EU countries designating it as safe and (2) fewer Albanians seeking asylum anywhere. This system-wide estimate should therefore be interpreted with more caution than the main triple-difference.

Table 3: Deterrence Effects of Safe Country Designations on Asylum Applications

	(1)	(2)
	Own Designation	System-Wide
	Log Applications	Log Applications
SCO Designation	-0.428	
	(0.249)	
Share Other Designating Destinations		-1.190**
		(0.474)
Pair FE	Yes	Yes
Origin \times Year FE	Yes	—
Destination \times Year FE	Yes	Yes
Sample	Full	Non-designated cells
Observations	4,350	3,985

Notes: Dependent variable is $\log(\text{applications} + 1)$. Column (1) estimates the own-designation deterrence effect using the full triple-difference specification. Column (2) tests whether a higher share of other EU destinations designating origin c as safe reduces applications to non-designating destination j —evidence of system-wide deterrence rather than diversion. The Share Other Designating Destinations variable is the leave-own-out fraction of sample destinations that designate origin c as safe in year t ; it varies at the origin \times year level and may partly capture correlated origin-specific shocks. Column (2) excludes origin \times year FE because the regressor varies at that level, and restricts to non-designated cells. Standard errors clustered at the destination-country level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

5.3 Protection-Type Substitution

The aggregate null on recognition rates may mask compositional changes in the type of protection granted. I decompose positive decisions into Geneva Convention refugee status and subsidiary/humanitarian protection. [Table 4](#) shows that designation is associated with a modest increase in the Geneva Convention recognition rate (0.045, $p = 0.10$) and a decrease in subsidiary/humanitarian protection (-0.080 , $p = 0.08$). While neither coefficient is individually significant at conventional levels, the opposite signs suggest that designation may shift the *composition* of protection—adjudicators reclassify cases between protection categories rather than changing the overall probability of a positive decision. This is consistent with the “safe” label triggering closer scrutiny of whether an applicant meets the higher bar of the Geneva Convention while reducing the use of the more discretionary subsidiary protection categories.

Table 4: Decision-Type Decomposition: Geneva Convention vs. Subsidiary Protection

	(1)	(2)	(3)
	Total Recognition	Geneva Convention	Subsidiary/Humanitarian
SCO Designation	-0.004 (0.026)	0.045 (0.027)	-0.080* (0.045)
Pair FE	Yes	Yes	Yes
Origin \times Year FE	Yes	Yes	Yes
Dest. \times Year FE	Yes	Yes	Yes
Observations	4,752	4,752	3,812

Notes: Dependent variable is the share of first-instance decisions granting each protection type. Column (1) reproduces the main result for total recognition. Column (2) uses Geneva Convention status grants as a share of total decisions. Column (3) uses subsidiary protection and humanitarian grants combined. All specifications include the full triple-difference fixed effects. Standard errors clustered at the destination level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

5.4 Heterogeneity

Table 5 examines whether effects differ by origin region and destination size. Columns (1)–(2) split by Balkan versus non-Balkan origins. The Balkan coefficient is -0.006 ($SE = 0.017$), confirming that even for the nationalities most directly targeted by SCO lists, designation does not alter recognition outcomes. The non-Balkan coefficient is positive (0.045 , $SE = 0.044$), though imprecise. Columns (3)–(4) split by destination size. In large receiving countries (Germany, France, Austria, Italy, Sweden, Netherlands), the coefficient is -0.049 ($p = 0.15$); in smaller destinations, it is $+0.055$ ($p = 0.09$). The opposite signs suggest compositional effects rather than a consistent designation impact.

Table 5: Heterogeneity: Safe Country Designation Effects by Subgroup

	(1)	(2)	(3)	(4)
	Balkan	Non-Balkan	Large	Small
	Origins	Origins	Destinations	Destinations
SCO Designation	-0.006	0.045	-0.049	0.055*
	(0.017)	(0.044)	(0.033)	(0.031)
Pair FE	Yes	Yes	Yes	Yes
Origin \times Year FE	Yes	Yes	Yes	Yes
Dest. \times Year FE	Yes	Yes	Yes	Yes
Observations	1,082	3,670	1,773	2,979

Notes: Dependent variable is the first-instance recognition rate. Columns (1)–(2) split by origin region: Balkan origins (Albania, Bosnia, Kosovo, Montenegro, North Macedonia, Serbia) vs. all others. Columns (3)–(4) split by destination size: Large includes Germany, France, Austria, Italy, Sweden, and the Netherlands; Small includes all other destinations. All specifications include the full triple-difference fixed effects. Standard errors clustered at the destination level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

5.5 Robustness

[Table 6](#) consolidates robustness evidence. The pairs cluster bootstrap p -value for the main estimate is 0.872 and the randomization inference p -value is 0.597, confirming that inference is robust to both cluster structure and the exact permutation distribution. Leave-one-destination-out estimates range from -0.013 to $+0.016$ —no single country drives the result. Using the rejection rate as the dependent variable yields a coefficient of 0.002 ($p = 0.95$)—the mirror image of the recognition rate null. The placebo test, which randomly permutes designation timing among treated pairs, yields a coefficient of -0.009 ($p = 0.75$), consistent

with the null hypothesis. Restricting to 2010–2020 produces an estimate near zero (0.005, $p = 0.85$).

Staggered DiD. The Callaway-Sant’Anna heterogeneity-robust estimator yields an aggregate ATT of -0.049 ($SE = 0.017$), substantially more negative than the TWFE triple-difference. Three features of the CS estimation context warrant caution before interpreting this as evidence of a genuine negative effect. First, the CS estimator restricts to clean treated-vs-never-treated comparisons and drops 11 always-treated units and 181 observations from unbalanced panels, substantially changing the effective sample. Second, several treatment cohorts have very few units (as few as 2 pairs in some year groups), and the CS estimator warned of “small groups” in 7 of 16 cohorts—group-time effects based on 2-3 observations are inherently noisy. Third, some early treatment cohorts (e.g., Austria’s 2009 Balkan designations) have very long post-treatment windows that may differ systematically from later cohorts. The CS estimate should therefore be treated as a sensitivity check that flags potential heterogeneity across cohorts, not as a preferred point estimate. The divergence between TWFE and CS merits further investigation with richer data but does not overturn the paper’s main conclusion, which is identified from the full triple-difference with richer fixed effects.

Table 6: Robustness Checks

	(1)	(2)	(3)	(4)
	Main	Rejection Rate	Placebo	2010–2020
SCO Designation	−0.004 (0.026)	0.002 (0.029)		0.005 (0.026)
Fake SCO			−0.009 (0.028)	
Bootstrap p (999 reps)	0.872			
Randomization inference p	0.597			
Callaway–Sant'Anna ATT	−0.049 (0.017)			
LOO-dest. range	[−0.013, 0.016]			
MDE (80% power)	0.072			
Observations	4,752	4,752	4,752	3,356

Notes: Column (1) reproduces the main triple-difference estimate with additional inference: pairs cluster bootstrap (999 replications), randomization inference (999 permutations of treatment timing within years), Callaway–Sant'Anna heterogeneity-robust staggered DiD using never-treated as controls, leave-one-destination-out range, and minimum detectable effect at 80% power (two-sided, 5% level). Column (2) uses rejection rate as the dependent variable. Column (3) randomly permutes designation years among treated pairs. Column (4) restricts to 2010–2020. All specifications include pair, origin×year, and destination×year FE. SEs clustered at destination level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

6. Discussion

6.1 Selection, Not Adjudication

The central finding is that safe country designations affect *who applies* for asylum, not *how applications are decided*. This distinction has a direct analogue in the broader literature on how formal rules interact with administrative discretion. [Lipsky \(2010\)](#) argued that street-level bureaucrats exercise substantial discretion regardless of formal policy mandates—that the effective policy is what happens at the point of implementation, not what the statute prescribes. The SCO setting is a particularly strong test case because the policy’s procedural consequences are unusually explicit: accelerated timelines, reversed burden of proof, restricted appeal rights. These are not vague guidelines; they are concrete procedural changes that should, in theory, make it harder for applicants from designated countries to win protection.

If even this strong a formal classification fails to measurably change grant rates, the finding suggests that adjudicators’ case-by-case judgment may be more robust to formal policy signals than commonly assumed. This is consistent with qualitative evidence from asylum law scholarship suggesting that adjudicators in most EU member states evaluate individual cases on their merits regardless of the safe-country designation ([Costello and Hancox, 2016](#)). The designation may formalize expectations about a nationality’s claims being weaker on average, but adjudicators still assess whether the individual applicant has demonstrated a credible fear of persecution.

The protection-type substitution finding—Geneva Convention grants increasing while subsidiary protection decreases—adds nuance to this interpretation. Adjudicators may not be ignoring the label entirely; rather, the label may trigger a shift in how protection is categorized. When an applicant’s country is designated as safe, the adjudicator may apply a higher standard that is more consistent with Geneva Convention criteria (requiring evidence of individual persecution rather than generalized insecurity), which would reduce subsidiary protection grants while leaving or even increasing the share that clears the Convention bar.

This is a form of classification substitution that leaves the total recognition rate unchanged.

An alternative interpretation is that the binary treatment indicator introduces classical measurement error: member states vary in how strictly they implement accelerated procedures for designated nationalities, and some designations may be largely symbolic. The Netherlands, for instance, designates 32 countries but may rely less on formal accelerated procedures tied to the designation than Germany, which channels designated-nationality applicants through a specific fast-track process. The null should therefore be read as an average effect across all designating implementations, including those that do not meaningfully change adjudication practices. Future work with data on processing times, accelerated-procedure usage rates, appeal rates, or case-level records could distinguish between “adjudicators ignore the label” and “the label is not uniformly implemented.”

6.2 Informational Deterrence

The deterrence finding connects to a growing literature on how policy signals affect migration decisions (Czaika and de Haas, 2013; Hatton, 2004). The own-designation effect is consistent with a standard deterrence model: potential applicants learn that their country has been labeled “safe” in a particular destination, infer that their chances of protection are low, and choose not to apply there. The magnitude—a 36% reduction in applications—is substantial, comparable to the effects of broader asylum policy restrictions documented in cross-country studies (Hatton, 2009).

The system-wide deterrence result is more novel and more puzzling. Applications fall even in non-designating destinations when more neighbors designate the same origin. This is the opposite of geographic diversion (which would predict increased applications to non-designating destinations) and instead suggests that the accumulation of “safe” signals across multiple EU states discourages applications from that nationality across the entire system. Potential asylum seekers may treat the density of designations as a signal about their chances in Europe as a whole, not just in specific countries. If three EU states have labeled your

country as safe, you may infer that the remaining states will be similarly skeptical—even if they have not formally designated your country.

This informational channel has implications for the common safe country list adopted in 2025. A unified EU signal—all member states simultaneously declaring seven countries safe—may carry more weight than scattered national designations. If so, harmonization could amplify deterrence effects substantially, reducing total applications from designated nationalities rather than merely redistributing them across destinations.

An important caveat: the system-wide estimate should be interpreted cautiously. The share variable varies at the origin \times year level and may capture correlated origin-specific shocks. For example, improving political conditions in Albania could simultaneously lead to (1) more EU countries designating it as safe and (2) fewer Albanians seeking asylum anywhere. The correlation between designation density and declining applications could reflect shared causes rather than a causal chain from designation to deterrence. Disentangling these channels would require exogenous variation in designation timing that is independent of origin-country conditions—a challenging identification problem for future work.

This raises an ethical question that the data cannot answer: is it appropriate to deploy a policy tool whose primary mechanism is deterring people from exercising their right to seek asylum? If designations improved decision quality—by helping adjudicators correctly identify unfounded claims—they could be defended on efficiency grounds. But if they operate by discouraging applications without changing the substantive evaluation of claims that are filed, the welfare calculus is less clear. The designation may deter both applicants with genuinely weak claims and those with strong claims who lack the resources or information to rebut the safe-country presumption.

6.3 Implications for EU Harmonization

The EU’s 2025 common list was designed to reduce disparities in asylum outcomes across member states—the “asylum lottery” that motivates this study. The findings here suggest this

hope may be misplaced. If national SCO lists never caused the recognition-rate disparities in the first place, harmonizing those lists will not address their root causes. What does explain the lottery? The most likely candidates are institutional culture, legal doctrine, adjudicator training, and the broader enforcement environment within each member state (Costello and Hancox, 2016; Toshkov and de Haan, 2013). These deep structural differences are not touched by list harmonization.

Harmonization may instead have two other effects. First, it could amplify system-wide deterrence by eliminating the non-designating destinations that previously served as alternative options for applicants from designated origins. If all EU states label Albania as safe, Albanian applicants can no longer redirect to Italy or Sweden—the “escape valve” that non-harmonized lists provided is closed. Whether this results in fewer total applications or in more dangerous irregular routes is a question the data cannot answer. Second, harmonization could reduce administrative costs by simplifying the procedural landscape, even if it does not change substantive outcomes. If the designation is primarily symbolic in terms of adjudication—as the null on recognition rates suggests—then the administrative streamlining may be the policy’s real contribution, rather than the improved decision quality that proponents claim.

7. Conclusion

Safe country of origin designations create what might be called a *designation illusion*: a large observed correlation between policy labels and administrative outcomes that disappears under causal scrutiny. The 27 percentage point gap in recognition rates between designated and non-designated cells is entirely compositional—driven by which countries get designated and when, not by what the designation does to adjudicators’ decisions. The real effects operate on the extensive margin: fewer applications and system-wide deterrence.

This decomposition—zero effect on the intensive margin, large effect on the extensive margin—resolves a puzzle in the asylum policy literature. Prior work has documented enormous cross-country variation in recognition rates for the same nationalities and has

identified SCO designations as a leading candidate explanation. The evidence here shows that this explanation, while consistent with the raw data, is spurious: the same fixed effects structure that absorbs origin-specific trends and destination-specific shocks also absorbs the entire apparent effect of designations. What remains is economically and statistically negligible—a precisely estimated zero that rules out effects larger than one-quarter of the raw gap at 80% power.

The broader lesson is about where policy incidence falls. Formal government classifications—safe country lists, risk ratings, compliance designations—may often shape selection into administrative systems more powerfully than they shape decisions within those systems. When the Australian government designates a country as safe for asylum purposes, or when a U.S. regulator classifies a financial institution as “systemically important,” or when a European environmental agency places a site on a protected list, the observable outcomes (rejection rates, compliance rates, development patterns) reflect both the label’s effect on decisions and its effect on who enters the system. The SCO setting demonstrates that these two channels can have dramatically different magnitudes—and that confusing them leads to erroneous inferences about policy effectiveness.

Several limitations constrain interpretation. First, the binary treatment indicator does not capture implementation intensity, and the null may average genuine effects in strictly implementing countries with zero effects in symbolically designating ones. Second, the system-wide deterrence estimate is vulnerable to origin-year confounding, and the causal interpretation is less secure than the main triple-difference. Third, I observe outcomes at the cell level (citizenship×destination×year), not at the individual applicant level—so I cannot directly test whether the composition of applicants changes after designation, which would provide direct evidence on the selection channel.

As the EU implements its common safe country list, policymakers should recognize that they are deploying a deterrence tool, not a decision-quality tool. The evidence here suggests that the label affects who walks through the door, not what happens once they do. This

distinction should inform both the design of the policy and the metrics by which its success is evaluated. If the goal is to reduce unfounded applications, the policy may succeed through deterrence—but at the cost of potentially deterring valid claims as well. If the goal is to improve the accuracy of asylum decisions, a different instrument is needed.

Acknowledgements

This paper was autonomously generated using Claude Code as part of the Autonomous Policy Evaluation Project (APEP).

Project Repository: <https://github.com/SocialCatalystLab/ape-papers>

Contributors: @olafdrw

First Contributor: <https://github.com/olafdrw>

References

- Asylum Information Database**, “Country Reports on Asylum Procedures,” Technical Report, European Council on Refugees and Exiles, Brussels 2023.
- Callaway, Brantly and Pedro H. C. Sant’Anna**, “Difference-in-Differences with Multiple Time Periods,” *Journal of Econometrics*, 2021, 225 (2), 200–230.
- Costello, Cathryn and Emily Hancox**, “Safe Country? Really? The Role of the Safe Country of Origin Concept in European Asylum Law,” *German Law Journal*, 2016, 17 (6), 1043–1066.
- Council of the European Union**, “Council Regulation Establishing a Common List of Safe Countries of Origin,” Technical Report, Council of the European Union, Brussels 2025.
- Czaika, Mathias and Hein de Haas**, “The Effectiveness of Immigration Policies,” *Population and Development Review*, 2013, 39 (3), 487–508.
- European Union Agency for Asylum**, “Practical Guide on the Application of the Safe Country Concepts,” Technical Report, EUAA, Valletta 2023.
- Hatton, Timothy J.**, “Seeking Asylum in Europe,” *Economic Policy*, 2004, 19 (38), 6–62.
- , “The Rise and Fall of Asylum: What Happened and Why?,” *Economic Journal*, 2009, 119 (535), F183–F213.
- Lipsky, Michael**, *Street-Level Bureaucracy: Dilemmas of the Individual in Public Services*, 30th anniversary expanded ed., New York: Russell Sage Foundation, 2010.
- Neumayer, Eric**, “Asylum Destination Choice: What Makes Some West European Countries More Attractive Than Others?,” *European Union Politics*, 2004, 5 (2), 155–180.

Ortega, Francesc and Giovanni Peri, “The Effect of Immigration on the Fiscal Balance,” *American Economic Journal: Economic Policy*, 2013, 5 (3), 225–264.

Rüedin, Didier, “Do Attitudes toward Immigrants Predict Asylum Policy? Evidence from European Countries,” *Journal of Refugee Studies*, 2019, 32 (Special Issue 1), i38–i52.

Thielemann, Eiko R., “Beyond Remittances: The Effects of Asylum Policies on Immigrant Flows,” *International Migration Review*, 2006, 40 (1), 174–195.

Toshkov, Dimiter and Laura de Haan, “The Dynamic Relationship between Asylum Applications and Recognition Rates in Europe (1987–2010),” *European Union Politics*, 2013, 14 (2), 161–183.

A. Data Appendix

Eurostat Asylum Decisions (`migr_asydcfsta`). This dataset provides annual first-instance asylum decisions disaggregated by citizenship of the applicant, receiving country, and decision type. Decision categories include: total positive (combining Geneva Convention status, subsidiary protection, and humanitarian protection), negative, and total. Sub-type breakdowns distinguish Geneva Convention/refugee status (`POS_RFG`), subsidiary protection (`POS_SPROT`), and humanitarian protection (`POS_HUM`). I download the full dataset via the `eurostat` R package and filter to total sex, total age, and individual country codes (excluding EU aggregates).

Eurostat Asylum Applications (`migr_asyappctza`). This dataset provides annual first-time asylum applications by citizenship and receiving country. I use the `FRST` (first-time) applicant category to avoid double-counting repeat applications.

SCO Treatment Matrix. I code 43 designation events from the AIDA Country Reports ([Asylum Information Database, 2023](#)) and the EUAA Practical Guide on the Safe Country Concepts ([European Union Agency for Asylum, 2023](#)). Key legislative sources include: Germany’s Asylum Procedures Acceleration Act (November 2014) and Asylum Package I (October 2015); Austria’s *Herkunftsstaatenverordnung* (2009, amended 2014 and 2018); Belgium’s Royal Decree (2012, amended 2016); France’s OFPRA Council decisions; and the Czech Republic’s Ministry of Interior Decree (2015). Events before the sample period (2008) are coded as always-treated.

National SCO List Details. Germany: Serbia, North Macedonia, Bosnia (Nov 2014); Albania, Kosovo, Montenegro (Oct 2015); Ghana, Senegal (pre-1993). Germany also designated Georgia and Moldova in September 2023, but these are excluded from the treatment sample because no full post-treatment calendar year is available. Austria: Western Balkans (2009); Albania (2014); Georgia (2018). Belgium: six Balkan states (2012); Georgia (2016).

France: OFPRA variable list including Albania (pre-2008), Kosovo (2015), Georgia (pre-2008), Serbia (2010), North Macedonia (pre-2008), Bosnia (2015), Montenegro (2015). Luxembourg: six Balkan states (pre-2008). Bulgaria: Serbia, North Macedonia (2016). Czech Republic: Balkans and Albania (2015).

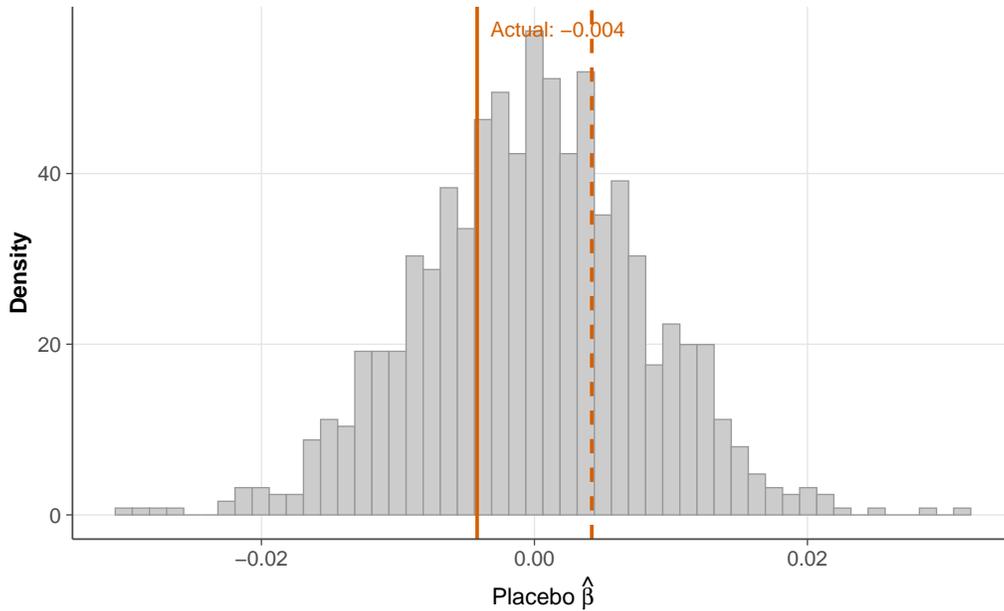
Sample Restrictions. I exclude cells with fewer than 10 total first-instance decisions to ensure reliable recognition rate estimation. EU-aggregate geographic codes (EU27, EU28, EEA, TOTAL) and aggregate citizenship codes are excluded.

B. Robustness Appendix

Leave-One-Out. The main triple-difference estimate ranges from -0.013 to $+0.016$ when each of the 22 destination countries is dropped in turn. No single country drives the result.

Pairs Cluster Bootstrap. With 22 destination clusters, standard cluster-robust inference may be unreliable. The pairs cluster bootstrap (999 replications) yields a 95% confidence interval of $[-0.057, 0.064]$, confirming the null.

Randomization Inference. I permute the SCO treatment indicator across pairs within each year 999 times and re-estimate the main specification. The RI p -value is 0.597. The actual estimate falls well within the permutation distribution ([Figure 2](#)).



on of 999 placebo estimates from randomization inference. Vertical lines: actual estimate (solid) and its mirror (dashed). RI p -value: 0.597.

Figure 2: Randomization Inference: Distribution of Placebo Estimates

Notes: Distribution of 999 placebo coefficients from permuting SCO treatment within years. Solid vertical line: actual estimate. Dashed line: mirror. Two-sided RI p -value = 0.60.

Callaway-Sant’Anna. The heterogeneity-robust staggered DiD estimator of [Callaway and Sant’Anna \(2021\)](#), using never-treated units as controls, yields an aggregate ATT of -0.049 ($SE = 0.017$). This is somewhat more negative than the TWFE estimate, suggesting potential heterogeneous effects across cohorts that TWFE averages toward zero. The CS estimate should be interpreted cautiously given small group sizes in several treatment cohorts.

Alternative Samples. Restricting to 2010–2020 (excluding COVID-era disruptions) yields $\hat{\beta} = 0.005$ ($SE = 0.026$). Results are robust to this restriction.

C. Standardized Effect Sizes

Table 7: Standardized Effect Sizes for Main Outcomes

Outcome	Spec.	$\hat{\beta}$	SD(Y)	SDE	SE(SDE)	Classification
<i>Panel A: Pooled</i>						
Recognition rate	Triple-diff	-0.004	0.325	-0.013	0.079	Small negative
Log applications	Triple-diff	-0.428	1.886	-0.227	0.132	Large negative
<i>Panel B: Heterogeneous</i>						
Recog. rate (Balkan origins)	Triple-diff	-0.006	0.180	-0.032	0.092	Small negative
Recog. rate (large dests.)	Triple-diff	-0.049	0.314	-0.155	0.106	Large negative

Notes: **Country:** European Union (14 EU member states with safe country of origin lists). **Research question:** Whether a member state designating an asylum seeker’s country of origin as ‘safe’ causally reduces their probability of receiving international protection. **Policy mechanism:** Under the EU Asylum Procedures Directive (2013/32/EU), member states may maintain national safe country of origin lists; applicants from designated countries face accelerated procedures and a reversed burden of proof, requiring them to demonstrate why their country is unsafe for them individually. **Outcome definition:** First-instance asylum recognition rate, calculated as positive decisions (Geneva Convention status, subsidiary protection, and humanitarian protection combined) divided by total first-instance decisions, from Eurostat `migr_asydcfsta`. **Treatment:** Binary indicator equal to one when destination country j has designated origin country c as a safe country of origin in year t . **Data:** Eurostat asylum decision and application statistics, 2008–2023, at the citizenship \times destination \times year level. **Method:** Triple difference-in-differences with origin \times destination pair, origin \times year, and destination \times year fixed effects; standard errors clustered at the destination-country level. **Sample:** Citizenship \times destination \times year cells with at least 10 total first-instance decisions, covering 9 designated origin countries and 10 never-designated control origins across 22 EU+ destination countries. $SDE = \hat{\beta}/SD(Y)$ where $SD(Y)$ is the unconditional standard deviation of the outcome. Classification refers to magnitude, not statistical significance: Large ($|SDE| > 0.15$), Moderate (0.05–0.15), Small (0.005–0.05), Null (< 0.005).