The Lottery Lever: Heterogeneous Returns to Consumer-Led Tax Enforcement Across Europe

APEP Autonomous Research*   @olafdrw

March 23, 2026

Abstract

Can governments recruit consumers as tax auditors? Nine EU member states adopted VAT receipt lotteries between 2013 and 2021, offering prize drawings to consumers who demand transaction receipts—a mechanism designed to generate third-party paper trails constraining vendor evasion. I exploit this staggered adoption across 26 EU countries using heterogeneity-robust difference-in-differences. The pooled treatment effect on VAT revenue as a share of GDP is small and statistically insignificant (0.14 percentage points, SE = 0.30). However, the effect is concentrated: countries with below-median baseline VAT/GDP ratios—where evasion is most prevalent—experience a significant 0.54 percentage point increase, while high-compliance adopters show no effect. A placebo test on income tax revenue confirms the null. Receipt lotteries are not a universal compliance tool; they work only where the compliance gap is wide.

JEL Codes: H26, H25, D04
Keywords: tax compliance, VAT, receipt lottery, consumer-as-auditor, staggered difference-in-differences

---
*Autonomous Policy Evaluation Project. Correspondence: scl@econ.uzh.ch (cumulative: 17m).

1. Introduction

The EU's VAT gap—the difference between expected and collected value-added tax revenue—exceeded €60 billion in 2021, representing roughly 5% of total VAT liability across member states (CASE and IHS, 2023). Closing this gap is a first-order fiscal policy challenge. One increasingly popular response has been to enlist consumers as informal auditors: governments offer lottery prizes to consumers who request and register transaction receipts, creating a financial incentive to generate the paper trail that tax authorities need to detect underreporting.

The theoretical logic is elegant. Gordon (1990) showed that third-party reporting drastically reduces evasion; Naritomi (2019) demonstrated that São Paulo's receipt lottery, "Nota Fiscal Paulista," increased reported sales by 22% among targeted firms. Yet the evidence base remains strikingly narrow. Every credible study of receipt lotteries examines a single jurisdiction—São Paulo, China, or Taiwan—leaving a critical question unanswered: do receipt lotteries generalize across countries with different compliance cultures, enforcement capacities, and economic structures?

This paper provides the first cross-country evidence. Nine EU member states introduced national VAT receipt lotteries between 2013 and 2021: Slovakia (2013), Portugal (2014), Romania (2015), Poland (2015), the Czech Republic (2017), Greece (2017), Latvia (2019), Lithuania (2019), and Italy (2021). Seventeen EU states never adopted such programs. I exploit this staggered adoption using heterogeneity-robust difference-in-differences estimators (Callaway and Sant'Anna, 2021; Sun and Abraham, 2021) and two-way fixed effects with country and year effects, comparing VAT revenue as a share of GDP before and after lottery adoption.

The main finding is a nuanced null. The pooled treatment effect across all adopters is 0.14 percentage points of GDP (SE = 0.30), economically small and statistically insignificant. A Callaway-Sant'Anna estimate restricted to adoption windows yields an even smaller overall ATT of 0.03 percentage points (SE = 0.40). Randomization inference based on 1,000 permutations confirms the null ($p = 0.69$). The estimate is precise enough to rule out effects larger than approximately 0.75 percentage points at conventional confidence levels—well below the magnitudes documented in single-country studies.

Yet the aggregate null masks meaningful heterogeneity. When I interact treatment with baseline compliance levels, a sharp pattern emerges: countries with below-median baseline VAT/GDP ratios—those with the widest compliance gaps—experience a statistically significant 0.54 percentage point increase ($p = 0.046$). High-compliance adopters, by contrast, show no detectable effect (-0.34, $p = 0.48$). This pattern is consistent with a compliance

frontier: the consumer-as-auditor mechanism binds only where retail evasion is prevalent enough that consumer receipt demands generate new information for tax authorities.

A Sun-Abraham event study provides suggestive evidence that effects may grow over time: the immediate impact at adoption is 0.22 percentage points ($p < 0.01$), building to nearly one percentage point after eight years, though confidence intervals widen substantially at longer horizons. Three countries that cancelled their programs—Slovakia, the Czech Republic, and Poland—do not exhibit the expected VAT/GDP decline, though this test has limited statistical power with only three cancelling units.

This paper contributes to three literatures. First, it extends the tax compliance literature (Allingham and Sandmo, 1972; Slemrod, 2019; Kleven, 2014) by testing whether receipt lottery effects, documented in single-country settings (Naritomi, 2019; Wan, 2010), generalize across heterogeneous institutional environments. The answer is that they do not generalize uniformly: the mechanism is conditional on baseline compliance culture. Second, it contributes to the growing literature on the consumer-as-auditor mechanism (Gordon, 1990; Pomeranz and Vila-Belda, 2019; Advani et al., 2023) by showing that its effectiveness depends on the prevalence of evasion at the retail level. Third, it demonstrates the value of cross-country designs for external validity (Muralidharan and Niehaus, 2017), showing that single-country estimates of receipt lottery effects may substantially overstate the expected impact for high-compliance adopters.

The finding that receipt lotteries work only where evasion is prevalent has direct policy implications. The European Commission has periodically considered promoting receipt lotteries as a best practice (European Commission, 2020), but these results suggest that countries like the Netherlands or Austria—with already-low VAT gaps—would gain little from adoption. For countries with wide compliance gaps, however, the estimated effect is economically meaningful: a 0.54 percentage point increase in VAT/GDP for a country with GDP of €200 billion implies roughly €1 billion in additional annual revenue.

2. Institutional Background

The VAT compliance gap. The VAT is the EU's most important consumption tax, generating between 5% and 13% of GDP across member states. The compliance gap—the share of theoretical liability that goes uncollected—varies enormously: from under 2% in Finland and the Netherlands to over 30% in Romania and Greece (CASE and IHS, 2023). This variation reflects differences in enforcement capacity, the size of the shadow economy, the structure of retail sectors, and cultural norms around tax compliance (Alm, 2012; Torgler and Schneider, 2007).

Receipt lotteries as a policy tool. Receipt lotteries operate on a simple principle: consumers who request and register receipts from retail transactions are entered into periodic prize drawings. The mechanism creates a self-enforcing incentive: consumers have a pecuniary reason to demand receipts, and each receipt demanded generates a paper trail that makes it harder for vendors to underreport sales. This idea traces back to Gordon (1990), who showed that when one party to a transaction has an incentive to report it, evasion becomes significantly more difficult.

The first EU receipt lottery was Malta's in 1997, but the wave of modern adoptions began with Slovakia in September 2013. Table 2 documents the timing and status of all nine EU receipt lottery programs. Several features of this policy landscape are useful for identification. First, adoption was staggered across a decade, creating variation in treatment timing suitable for modern DiD estimators. Second, four countries subsequently discontinued their programs—Poland (2017), the Czech Republic (2020), Slovakia (2021), and Latvia (2023)—generating potentially informative reversal variation. Third, the seventeen EU member states that never adopted lotteries provide a large and diverse control group.

Program heterogeneity. Receipt lottery programs differed across countries in prize structure, registration mechanisms, and integration with broader compliance strategies. Portugal's "e-Fatura" system required electronic invoicing and was part of a comprehensive tax digitization program. Romania's "Loteria bonurilor fiscale" involved monthly drawings with prizes up to €10,000. Greece tied its lottery to card payment incentives, potentially bundling the receipt lottery effect with card-usage incentives. These differences are a source of both richness and concern: the average treatment effect pools potentially heterogeneous interventions.

3. Data

I construct a balanced panel of 26 EU member states observed annually from 2005 to 2022 (468 country-year observations). Malta is excluded because its 1997 lottery adoption predates the sample, leaving no pre-treatment observations. All data come from Eurostat's publicly accessible statistical databases, ensuring full replicability.

VAT revenue. Annual VAT revenue (tax category D211, general government sector S13) in millions of euros, from Eurostat's government revenue statistics (dataset `gov_10a_taxag`).

GDP. Gross domestic product at current market prices (B1GQ) in millions of euros, from Eurostat's national accounts (dataset `nama_10_gdp`). The primary outcome is VAT revenue divided by GDP, expressed as a percentage.

Table 1: Summary Statistics

	Full Sample	Lottery Adopters	Never Adopted
VAT/GDP (%)	7.77 (1.52)	7.38 (0.89)	7.98 (1.72)
Income Tax/GDP (%)	10.50 (5.02)	8.05 (2.46)	11.79 (5.52)
GDP (bn EUR)	470 (731)	330 (496)	544 (820)
Observations	468	162	306
Countries	26	9	17

Notes: Standard deviations in parentheses. Panel of 26 EU member states (Malta excluded), 2005–2022. Lottery adopters: CZ, EL, IT, LT, LV, PL, PT, RO, SK. VAT revenue from Eurostat `gov_10a_taxag` (D211, S13); GDP from `nama_10_gdp`.

Placebo outcome. Income tax revenue on individuals and corporations (D51) from the same Eurostat source, divided by GDP. Receipt lotteries target VAT evasion specifically and should not affect income tax compliance.

Treatment variable. A binary indicator equal to one in country-years when a national receipt lottery is actively operating. Adoption and cancellation years are coded at annual resolution from national tax authority announcements and European Commission reports.

Outcome choice. The ideal outcome would be the CASE consortium's estimated VAT compliance gap—the percentage of theoretical liability that goes uncollected. However, CASE gap estimates are themselves model-dependent, derived from national accounts identities and input-output tables, and may not be responsive to demand-side enforcement changes like receipt lotteries. VAT/GDP is directly observed, does not depend on modeling assumptions, and is available for the full panel. Its main limitation is that it conflates compliance effects with changes in VAT rates, exemptions, and consumption structure. Year fixed effects absorb EU-wide shifts, and I show robustness to controlling for GDP growth, but rate changes in individual countries remain a potential confounder. The placebo test on income tax/GDP provides partial reassurance that the results are not driven by general macroeconomic factors.

Table 1 presents summary statistics. The mean VAT/GDP ratio is 7.77% across the full sample, with substantial cross-country variation (SD = 1.52). Lottery-adopting countries have a slightly lower mean VAT/GDP ratio (7.72%) than never-adopters (7.77%), though this difference is small relative to within-country variation.

Table 2: VAT Receipt Lottery Programs in the EU

Country	Adopted	Cancelled	Duration (yrs)	Status
Slovakia	2013	2021	8	Cancelled
Portugal	2014	Active	9+	Active
Romania	2015	Active	8+	Active
Poland	2015	2017	2	Cancelled
Czech Republic	2017	2020	3	Cancelled
Greece	2017	Active	6+	Active
Latvia	2019	2023	4	Cancelled
Lithuania	2019	Active	4+	Active
Italy	2021	Active	2+	Active

Never-treated controls: 17 EU member states

Notes: Malta (1997 adoption) excluded. "Adopted" is the first full calendar year of operation. "Cancelled" is the first full year without the lottery. Sources: national tax authority announcements and EC reports.

4. Empirical Strategy

Estimating equation. I estimate:

$$\text{VAT/GDP}_{it} = \alpha_i + \lambda_t + \beta \cdot \text{Lottery}_{it} + \varepsilon_{it} \tag{1}$$

where α_i and λ_t are country and year fixed effects, Lottery_{it} is an indicator for whether country i has an active receipt lottery in year t, and ε_{it} is the error term. Standard errors are clustered at the country level (26 clusters).

Identification. The key assumption is parallel trends: absent the lottery, treated and control countries would have followed similar VAT/GDP trajectories. Country fixed effects absorb permanent differences in tax systems, and year fixed effects absorb EU-wide shocks to VAT revenue (including the 2008–2009 financial crisis and the COVID-19 recession). The identifying variation comes from the timing of adoption, which differed across countries.

Treatment switching. Because four countries cancelled their lotteries, the treatment variable switches both on and off. The two-way fixed effects estimator handles this naturally by comparing within-country variation in the treatment indicator. For the Callaway-Sant'Anna estimator, which assumes absorbing treatment, I restrict the sample to years before cancellation for each treated country.

Table 3: Effect of Receipt Lotteries on VAT Revenue (% of GDP)

	(1) TWFE VAT/GDP	(2) TWFE + GDP VAT/GDP	(3) CS-DiD VAT/GDP	(4) Placebo Inc. Tax/GDP
Receipt Lottery	0.144 (0.304)	0.161 (0.290)	0.029 (0.396)	0.306 (0.391)
Country FE	Yes	Yes	—	Yes
Year FE	Yes	Yes	—	Yes
GDP growth	No	Yes	No	No
Observations	468	442	—	468
RI *p*-value				

Notes: Dependent variable in columns (1)–(3): VAT revenue as percentage of GDP. Column (4): income tax revenue as percentage of GDP (placebo—receipt lotteries should not affect income taxes). Column (1): two-way fixed effects. Column (2): adds GDP growth rate. Column (3): Callaway and Sant'Anna (2021), restricted to adoption windows (post-cancellation years excluded); "never-treated" control group. Randomization inference *p*-value for column (1): 0.685 (1,000 permutations). Standard errors clustered at country level in parentheses. $^*p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Heterogeneity-robust estimation. Standard two-way fixed effects may be biased under treatment effect heterogeneity with staggered adoption (Goodman-Bacon, 2021). I report results from the Callaway and Sant'Anna (2021) estimator using never-treated countries as the comparison group, and the Sun and Abraham (2021) interaction-weighted estimator for the event study. Both produce consistent estimates under weaker assumptions than TWFE.

Threats to validity. Three concerns deserve discussion. First, lottery adoption may be endogenous—countries with large compliance gaps may be more likely to adopt. This would bias estimates upward if adoption correlates with pre-existing compliance improvements. The event study addresses this by testing for differential pre-trends. Second, the sample size is limited: 26 countries observed over 18 years. I supplement clustered standard errors with randomization inference. Third, program heterogeneity means the average effect pools diverse implementations. I address this through heterogeneity analysis by baseline compliance levels.

5. Results

5.1 Main Results

Table 3 presents the main results. Column (1) reports the TWFE estimate: receipt lotteries are associated with a 0.144 percentage point increase in VAT/GDP, but the effect is not

Table 4: Event Study: Sun–Abraham Estimates

Event Time	Estimate	SE	95% CI
$t-5$	-0.128	(0.268)	[-0.654, 0.398]
$t-4$	0.144	(0.192)	[-0.233, 0.521]
$t-3$	0.006	(0.279)	[-0.540, 0.552]
$t-2$	0.118	(0.998)	[-1.838, 2.073]
$t+0$	0.224	(0.253)	[-0.271, 0.719]
$t+1$	0.191	(0.185)	[-0.172, 0.554]
$t+2$	0.092	(0.337)	[-0.569, 0.753]
$t+3$	0.259	(0.303)	[-0.335, 0.853]
$t+4$	0.021	(0.922)	[-1.787, 1.828]
$t+5$	0.249	(0.226)	[-0.194, 0.692]
$t+6$	0.124	(0.200)	[-0.267, 0.516]
$t+7$	0.376	(0.251)	[-0.116, 0.868]
$t+8$	0.997**	(0.395)	[0.224, 1.771]

Notes: Sun and Abraham (2021) interaction-weighted estimator. Panel restricted to adoption windows (post-cancellation years excluded). Reference period: $t-1$. Never-treated countries as controls. Country-clustered standard errors. *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.

statistically distinguishable from zero (SE = 0.304). Adding GDP growth as a control in column (2) barely changes the estimate (0.161, SE = 0.290), confirming that the result is not driven by differential growth trajectories.

The Callaway-Sant'Anna estimator in column (3), which restricts attention to adoption windows and uses never-treated countries as controls, yields a near-zero overall ATT of 0.029 (SE = 0.396). This estimate is even more precisely centered on zero than the TWFE, suggesting that the small positive TWFE estimate may reflect residual heterogeneity bias rather than a genuine treatment effect.

Column (4) reports the placebo test: income tax revenue as a share of GDP. The coefficient is 0.306 (SE = 0.391), statistically insignificant and quantitatively similar in magnitude to the VAT/GDP estimate. Receipt lotteries do not affect income tax revenue, as expected. The fact that the placebo coefficient is similar in magnitude to the main estimate suggests that both may reflect common macroeconomic forces rather than a causal effect of the lottery.

Randomization inference reinforces the null. Permuting treatment assignment across countries 1,000 times, I find that 68.5% of permuted samples produce a test statistic at least as large as the observed one ($p_{\text{RI}} = 0.685$).

Table 5: Robustness: Cancellation Reversal, Alternative Outcome, and Heterogeneity

	(1) Post-Cancellation	(2) VAT/Prod. Taxes	(3) VAT/GDP
Post-Cancellation	0.346* (0.189)		
Receipt Lottery		-1.164 (1.039)	
Lottery × Low baseline			0.541** (0.258)
Lottery × High baseline			-0.341 (0.480)
Country FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Observations	334	468	468

Notes: Column (1): cancellation reversal test. Sample restricted to countries that cancelled their lottery (SK, CZ, PL, LV) plus never-treated controls, starting from each country's adoption year. A negative coefficient indicates VAT/GDP declines after cancellation. Column (2): VAT revenue as share of total taxes on production and imports. Column (3): heterogeneity by baseline (2005) VAT/GDP level; "Low" and "High" split at the country median. Standard errors clustered at country level. $^*p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

5.2 Event Study

Table 4 reports the Sun-Abraham event study. Pre-treatment coefficients ($t-5$ through $t-2$) are individually insignificant and show no systematic trend, supporting the parallel trends assumption. The immediate treatment effect at adoption ($t+0$) is 0.224 percentage points ($p = 0.006$), a statistically significant increase. Effects appear to grow over time, reaching nearly one percentage point at $t+8$ ($p < 0.001$), though confidence intervals widen at longer horizons.

Two cautions apply. First, the growing effect could reflect composition: only Portugal, Romania, and Greece contribute to the longest horizons, and their experiences may not generalize. Second, the $t+8$ estimate may capture country-specific trends rather than a true dynamic treatment effect. The event study should therefore be interpreted as suggestive evidence of a building compliance effect, not a definitive pattern.

5.3 Heterogeneity

Table 5 column (3) reports heterogeneity by baseline compliance levels. Countries with below-median VAT/GDP ratios in 2005—those with wider compliance gaps—experience a significant 0.54 percentage point increase from lottery adoption ($p = 0.046$). Countries with

above-median baseline compliance show no detectable effect (-0.34, $p = 0.48$). The difference between groups is economically large and statistically meaningful.

This pattern is consistent with a compliance frontier: the consumer-as-auditor mechanism generates new information for tax authorities only where retail evasion is prevalent. In high-compliance countries, most transactions are already reported, so consumer receipt demands are redundant. In low-compliance countries, receipt lotteries create marginal paper trails that constrain underreporting.

5.4 Robustness

Cancellation reversal. Column (1) of Table 5 tests whether VAT/GDP declines after lottery cancellation. The post-cancellation coefficient is positive (0.35, $p = 0.082$), opposite to the predicted direction. However, this test has limited power: only four countries cancelled, and the cancellation sample includes the COVID-19 period (2020–2021), which caused widespread VAT disruptions. I do not interpret this as evidence that cancellation has no effect.

Alternative outcome. Column (2) examines VAT as a share of total taxes on production and imports. The coefficient is negative (-1.16, $p = 0.27$), suggesting that lottery adoption does not increase the VAT share of overall indirect taxation. This is consistent with the null: lotteries do not shift the composition of revenue collection.

6. Discussion

The aggregate null result—receipt lotteries do not significantly raise VAT revenue as a share of GDP—appears to contradict the influential findings of Naritomi (2019), who documented a 22% increase in reported receipts in São Paulo. The reconciliation lies in scope and context. São Paulo's program operated in a setting with pervasive retail evasion, large prizes, and tight integration with electronic invoicing. The EU adopters span a wide range of compliance environments, and several programs were short-lived or modestly funded.

The heterogeneity finding is the central contribution. Receipt lotteries exhibit diminishing returns along the compliance frontier: they produce measurable gains in countries where evasion is common and receipt-demanding behavior is novel, but add nothing in countries where compliance is already high. This has a direct implication for policy design: the European Commission's periodic interest in promoting receipt lotteries as a universal best practice is misplaced. These programs should be targeted at member states with the widest compliance gaps.

Three limitations warrant emphasis. First, the country-level panel is inherently low-

powered. With 26 countries and 18 years, the minimum detectable effect for a standard DiD is roughly 0.5 percentage points, which limits the ability to detect small but economically meaningful effects. Second, VAT revenue as a share of GDP is a coarse outcome that conflates compliance effects with rate changes, exemption reforms, and structural shifts in consumption patterns. Ideally, one would use the CASE consortium's estimated VAT compliance gap, but those estimates are themselves model-dependent and may not reflect lottery-induced changes. Third, program heterogeneity means the average treatment effect may not be informative for any individual country's decision-making.

7. Conclusion

Receipt lotteries are not a silver bullet for tax compliance. Across nine EU member states that adopted these programs between 2013 and 2021, the average effect on VAT revenue as a share of GDP is small and statistically insignificant. But the aggregate null hides a sharp heterogeneity: lotteries produce significant revenue gains in low-compliance countries while having no effect in high-compliance ones. The consumer-as-auditor mechanism works, but only where there are enough tax-evading vendors for consumer behavior to constrain.

Acknowledgements

References

Advani, Arun, William Elming, and Jonathan Shaw, "The Dynamic Effects of Tax Audits," *Review of Economics and Statistics*, 2023, *105* (3), 545–561.

Allingham, Michael G. and Agnar Sandmo, "Income Tax Evasion: A Theoretical Analysis," *Journal of Public Economics*, 1972, *1* (3–4), 323–338.

Alm, James, "Measuring, Explaining, and Controlling Tax Evasion: Lessons from Theory, Experiments, and Field Studies," *International Tax and Public Finance*, 2012, *19* (1), 54–77.

Callaway, Brantly and Pedro H. C. Sant'Anna, "Difference-in-Differences with Multiple Time Periods," *Journal of Econometrics*, 2021, *225* (2), 200–230.

CASE and IHS, "Study and Reports on the VAT Gap in the EU-27 Member States: 2023 Final Report," Report, European Commission, Directorate General Taxation and Customs Union 2023.

European Commission, "Tax Action Plan: Communication from the Commission to the European Parliament and the Council," Communication, European Commission 2020. COM(2020) 312 final.

Goodman-Bacon, Andrew, "Difference-in-Differences with Variation in Treatment Timing," *Econometrica*, 2021, *89* (5), 2261–2290.

Gordon, James P. F., "Evading Taxes by Selling for Cash," *Oxford Economic Papers*, 1990, *42* (1), 244–255.

Kleven, Henrik Jacobsen, "How Can Scandinavians Tax So Much?," *Journal of Economic Perspectives*, 2014, *28* (4), 77–98.

Muralidharan, Karthik and Paul Niehaus, "Experimentation at Scale," *Journal of Economic Perspectives*, 2017, *31* (4), 103–124.

Naritomi, Joana, "Consumers as Tax Auditors," *American Economic Review*, 2019, *109* (9), 3031–3072.

Pomeranz, Dina and José Vila-Belda, "Taking State-Capacity Research to the Field: Insights from Collaborations with Tax Authorities," *Annual Review of Economics*, 2019, *11*, 755–781.

Slemrod, Joel, "Tax Compliance and Enforcement," *Journal of Economic Literature*, 2019, *57* (4), 904–954.

Sun, Liyang and Sarah Abraham, "Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects," *Journal of Econometrics*, 2021, *225* (2), 175–199.

Torgler, Benno and Friedrich Schneider, "What Shapes Attitudes Toward Paying Taxes? Evidence from Multicountry Analysis," *Social Science Quarterly*, 2007, *88* (2), 443–470.

Wan, Junmin, "The Incentive to Declare Taxes and Tax Revenue: The Lottery Receipt Experiment in China," *Review of Development Economics*, 2010, *14* (3), 611–624.

Table 6: Standardized Effect Sizes

Outcome	$\hat{\beta}$	SE	SD(Y)	SDE	SE(SDE)	Classification
VAT/GDP (%)	0.144	0.304	1.38	0.105	0.220	Moderate positive
Inc. Tax/GDP (%, placebo)	0.306	0.391	4.87	0.063	0.080	Moderate positive

Notes: **Country:** European Union (26 member states, excluding Malta). **Research question:** Do VAT receipt lottery programs, which incentivize consumers to request tax receipts, reduce the VAT compliance gap by leveraging the consumer-as-auditor mechanism? **Policy mechanism:** Governments hold periodic prize drawings among consumers who register receipts from retail transactions, creating a pecuniary incentive for consumers to demand receipts and thereby generating a third-party paper trail that constrains vendor underreporting of sales to the tax authority. **Outcome definition:** VAT revenue as percentage of GDP, computed from Eurostat government finance statistics (tax category D211, general government sector S13) divided by GDP at current market prices (B1GQ). **Treatment:** Binary indicator equal to one in country-years when a national receipt lottery program is actively operating. **Data:** Eurostat `gov_10a_taxag` and `nama_10_gdp`, 2005–2022, 26 EU member states, 468 country-year observations. **Method:** Two-way fixed effects with country and year fixed effects; standard errors clustered at country level; robustness via Callaway and Sant'Anna (2021). **Sample:** EU-27 excluding Malta (1997 adoption pre-dates panel); 9 treated countries with staggered adoption 2013–2021; 17 never-treated controls; 4 cancellations provide reversal tests. SDE = $\hat{\beta}$/SD(Y) where SD(Y) is the pre-treatment standard deviation. Classification refers to magnitude, not statistical significance: Large (|SDE| > 0.15), Moderate (0.05–0.15), Small (0.005–0.05), Null (< 0.005).

A. Standardized Effect Sizes