The Regularization Illusion: Mass Legalization, Formalization, and the Null Labor Market

APEP Autonomous Research* @ai1scl

March 30, 2026

Abstract

Does mass regularization of immigrants affect host-country labor markets? I study Colombia's 2021 Estatuto Temporal de Protección para Migrantes Venezolanos (ETPV), which granted 10-year work permits to approximately 1.8 million Venezuelan migrants—the largest regularization program in Latin American history. Exploiting cross-department variation in Venezuelan concentration interacted with the ETPV's implementation, I estimate continuous-treatment difference-in-differences models with wild cluster bootstrap inference. The results are a precisely estimated null: regularization had no detectable effect on department-level employment rates ($\hat{\beta} = -0.310$, $p = 0.704$), unemployment rates ($\hat{\beta} = 0.026$, $p = 0.815$), or labor force participation ($\hat{\beta} = -0.293$, $p = 0.654$). Pre-trends are clean and all placebo tests pass. These findings suggest that converting irregular migrants to legal status—even at massive scale—need not disrupt aggregate labor markets, consistent with absorption through informality channels.

JEL Codes: J61, J46, O17, F22

Keywords: immigration regularization, informality, Venezuela, Colombia, ETPV, labor markets

*Autonomous Policy Evaluation Project. Correspondence: scl@econ.uzh.ch (cumulative: 35m).

1. Introduction

Governments worldwide face a persistent dilemma: millions of unauthorized immigrants live and work within their borders, and the decision to regularize them is among the most politically contentious in economic policy. The European Union extended temporary protection to over four million Ukrainians after 2022. Turkey hosts 3.6 million Syrians under subsidiary protection. The United States periodically debates pathways to legal status for its estimated eleven million undocumented residents. In every case, opponents warn that legalization will flood formal labor markets, depress wages, and displace native workers. Yet the empirical basis for these fears remains remarkably thin, because mass regularization events are rare, and those that do occur are difficult to study credibly.

The core problem is that existing immigration research overwhelmingly studies *arrival*—the labor market effects of new immigrant inflows—rather than *regularization*—the consequences of granting legal status to immigrants already present and working (Borjas, 2003; Card, 1990; Dustmann et al., 2017). These are fundamentally different policy margins. Arrival changes the quantity of labor supply; regularization changes the legal and institutional conditions under which an existing labor supply operates. The theoretical predictions diverge sharply. Models of task specialization (Peri and Sparber, 2009) and complementarity (Ottaviano and Peri, 2012) predict that arrival effects depend on skill composition, while regularization effects depend on labor market institutions—particularly the size and structure of the informal sector.

This paper studies the largest mass regularization in Latin American history: Colombia's 2021 ETPV (Estatuto Temporal de Protección para Migrantes Venezolanos), which granted 10-year temporary protection permits to approximately 1.8 million Venezuelan nationals. The policy was remarkable in its scale—covering roughly 3.6% of Colombia's total population—and in its institutional design. Prior to the ETPV, the vast majority of Venezuelan migrants in Colombia worked informally, outside the reach of labor regulations, social security, and minimum wage enforcement (Ibáñez et al., 2022). The ETPV did not create new immigration; the Venezuelans were already there. It created a legal pathway from informal to formal employment.

I exploit cross-department variation in Venezuelan migrant concentration, measured by ETPV pre-registrations per capita, interacted with the post-2021 implementation period. The treatment variable is continuous, reflecting the wide dispersion of Venezuelan settlement across Colombia's 23 departments with GEIH coverage—from 0.4% in Caquetá to 17.4% in Norte de Santander. I estimate two-way fixed effects models with department and year fixed effects, using wild cluster bootstrap inference (Cameron et al., 2008) to account for the small number of clusters.

The main finding is a well-identified null. The ETPV had no detectable effect on department-level employment rates ($\hat{\beta} = -0.310$, SE $= 0.400$, wild cluster bootstrap $p = 0.704$), unemployment rates ($\hat{\beta} = 0.026$, SE $= 0.106$, $p = 0.815$), or labor force participation rates ($\hat{\beta} = -0.293$, SE $= 0.364$, $p = 0.654$). The estimates are not merely imprecise zeros. Event study estimates show clean pre-trends (joint F-test $p = 0.291$ for employment, $p = 0.273$ for unemployment), and all placebo tests pass. The minimum detectable effect at 80% power is 1.12 percentage points per one-percentage-point increase in Venezuelan share—a substantively meaningful threshold given that the most exposed departments experienced Venezuelan shares exceeding 15%.

These results are robust across multiple specifications: including or excluding the COVID year (2020), adding department-specific linear trends, using binary rather than continuous treatment, leave-one-out jackknife excluding the most treated departments, and placebo treatment timing. The only heterogeneity that emerges is suggestive: in departments with above-median baseline employment rates, the interaction term is negative and statistically significant ($\hat{\beta} = -0.679$, $p < 0.05$), suggesting that regularization may generate modest competitive pressure specifically in already-tight labor markets. But even this effect is economically small relative to the scale of the regularization.

This paper contributes to three literatures. First, it advances the growing body of work on Venezuelan migration in Colombia and the region (Lebow, 2024; Peñaloza-Pacheco, 2023; Caruso et al., 2021; Rozo and Vargas, 2024). That literature has established that Venezuelan *arrival* increased informality among natives and reduced formal wages at the bottom of the distribution. My contribution is to show that the subsequent *regularization* of those same migrants—a conceptually distinct policy lever—produced no measurable aggregate labor market disruption. The arrival finding and the regularization finding are not contradictory; they illuminate different margins of adjustment.

Second, I contribute to the small but important literature on immigration regularization programs (Clemens et al., 2018; Kossoudji and Cobb-Clark, 2002; Cobb-Clark et al., 1995; Amuedo-Dorantes et al., 2020; Monras, 2020). Most existing evidence comes from the 1986 U.S. IRCA amnesty, which regularized roughly three million undocumented workers. Those studies find modest positive effects on regularized workers' wages and occupational upgrading, but limited evidence on aggregate market effects. The ETPV provides a test at far greater relative scale—1.8 million in a country of 50 million, compared to three million in a country of 240 million—and in an economy with a much larger informal sector, where absorption dynamics differ fundamentally.

Third, I speak to the broader literature on how labor markets absorb immigration shocks

(Borjas, 2017; Clemens and Hunt, 2019; Card, 2001). The canonical framework—from Borjas (2003) through Dustmann et al. (2017)—focuses on the substitutability between immigrant and native labor. But in economies with large informal sectors, the relevant margin of adjustment is not wage flexibility or displacement but rather the movement of workers between formal and informal employment (Ulyssea, 2020; Meghir et al., 2015). Colombia, where informality accounts for roughly half of total employment, is the natural laboratory for studying this channel.

I name the portable mechanism "the regularization illusion": the expectation that converting irregular workers to legal status will produce measurable disruption in aggregate labor markets, when in fact the informal sector has already absorbed the adjustment. The null result is informative precisely because the setting is maximally powered to detect an effect—a program covering 3.6% of the population, concentrated in identifiable geographic areas, with clean pre-trends and strong first-stage variation.

The remainder of the paper proceeds as follows. Section 2 describes the ETPV policy and institutional context. Section 3 presents the data. Section 4 details the empirical strategy. Section 5 reports results, including robustness and heterogeneity analyses. Section 6 discusses interpretations and limitations. Section 7 concludes.

2. Institutional Background: The ETPV

2.1 The Venezuelan Migration Crisis

Venezuela's economic and political crisis, which accelerated sharply after 2015, produced one of the largest displacement events in the Western Hemisphere. By 2021, over 5.9 million Venezuelans had left the country, with Colombia receiving the largest share—approximately 1.8 million by official estimates (R4V Inter-Agency Coordination Platform, 2021). The migration was overwhelmingly driven by economic collapse: hyperinflation, shortages of food and medicine, and the dismantling of formal labor market institutions under the Maduro government.

Colombian migration policy evolved through several phases. Initially, Venezuela-Colombia border crossing was relatively unregulated, reflecting decades of binational population flows. As the crisis deepened, Colombia introduced the Permiso Especial de Permanencia (PEP) in 2017, granting renewable two-year permits to Venezuelans who had entered legally. However, the PEP covered only a fraction of the migrant population, and by 2020, an estimated 56% of Venezuelan nationals in Colombia lacked any form of regular migration status (World Bank, 2021).

2.2 The Estatuto Temporal de Protección (ETPV)

On February 8, 2021, President Iván Duque announced the creation of the ETPV through Decreto 216, which took effect on March 1, 2021. The ETPV represented a dramatic policy shift, offering 10-year Temporary Protection Permits (Permiso por Protección Temporal, PPT) to all Venezuelan nationals who could demonstrate presence in Colombia before January 31, 2021, or who entered through regular channels within the following two years.

The PPT provides comprehensive labor market access: holders can work legally for any employer, enroll in the social security system (pension, health, and occupational risk insurance), open bank accounts, and access public services. Critically, the PPT converts irregular migrant workers—who previously could only work informally, without contracts or labor protections—into legally authorized participants in the formal labor market.

Implementation proceeded in two phases. A pre-registration phase (May–November 2021) collected biometric and demographic data from 2.46 million individuals, exceeding initial government estimates. The permit issuance phase began in October 2021, with 1.4 million PPTs delivered by September 2022 and continued rollout through 2023. The pre-registration data, published on Colombia's open data portal (datos.gov.co), provides granular geographic information on the distribution of ETPV beneficiaries by department.

2.3 Why the ETPV Is a Clean Natural Experiment

Several features make the ETPV well-suited for causal identification. First, the policy was announced and implemented rapidly—the decree was signed in February 2021 and pre-registration began in May—limiting the scope for anticipatory behavioral responses. Second, eligibility was universal for Venezuelan nationals present before the cutoff date, eliminating selection on individual characteristics within the migrant population. Third, the geographic variation in treatment intensity was determined by pre-existing settlement patterns that reflect historical migration networks and border proximity rather than contemporaneous labor market conditions. Fourth, the policy changed only the legal status of migrants already present and working in Colombia—it did not create new migration flows, alter deportation enforcement, or change the underlying skill composition of the migrant workforce.

3. Data

3.1 Labor Market Outcomes

Labor market outcomes come from the DANE Gran Encuesta Integrada de Hogares (GEIH) Department Annex, which reports annual labor market indicators for 23 Colombian departments from 2015 to 2024. The four primary outcomes are: the employment rate (*tasa de ocupación*), the unemployment rate (*tasa de desocupación*), the labor force participation rate (*tasa global de participación*), and the underemployment rate (*tasa de subempleo subjetivo*). These aggregated indicators capture the department-level equilibrium effects of regularization, which is the appropriate unit for studying labor market adjustment to a department-wide policy change.

I exclude 2020 from the baseline specification because COVID-19 lockdowns generated extreme labor market disruptions orthogonal to migration policy. The resulting panel comprises 207 department-year observations (23 departments $\times$ 9 years: 2015–2019, 2021–2024). Robustness checks including 2020 (230 observations) produce substantively identical results.

3.2 Treatment Variable

The treatment variable is constructed from Migración Colombia's ETPV pre-registration records, accessed via the Colombian open data portal (datos.gov.co). The dataset contains 2.46 million individual pre-registration records with department of residence. I compute treatment intensity as the number of ETPV pre-registrations in each department divided by the department's 2019 population from DANE projections.

This measure captures the department-level "dose" of regularization. It ranges from approximately 0.4% in Caquetá to 17.4% in Norte de Santander, with a mean of 4.4% and substantial cross-sectional variation (standard deviation of 4.0 percentage points). The distribution largely reflects geographic proximity to the Venezuelan border and the presence of established Venezuelan communities.

3.3 Summary Statistics

Table 1 presents pre-treatment (2019) characteristics by tercile of ETPV treatment intensity. High-treatment departments—those with Venezuelan shares above 5.5%—tend to be larger (average population 2.46 million versus 1.15 million in low-treatment departments) and have slightly higher employment and participation rates. These baseline differences motivate the inclusion of department fixed effects, which absorb all time-invariant departmental characteristics.

Table 1: Summary Statistics: Pre-Treatment Department Characteristics (2019)

	N	Ven. Share (%)	Empl. Rate	Unemp. Rate	Partic. Rate	Underempl. Rate	Pop. (000s)
Low	8	1.2	56.7	10.8	63.5	10.6	1154
Medium	7	3.7	56.2	10.7	62.9	8.5	1632
High	8	8.2	58.6	10.5	65.3	11.1	2456
Full Sample	23	4.4	57.2	10.6	63.9	10.1	1752

Notes: Departments split into terciles by ETPV pre-registration intensity (registrations per 2019 population). Venezuelan share ranges from 0.4% (Caquetá) to 17.4% (Norte de Santander). Source: DANE GEIH Department Annex (2024) and Migración Colombia ETPV Registry via datos.gov.co.

4. Empirical Strategy

4.1 Identification

The identifying assumption is that, absent the ETPV, labor market outcomes in high-Venezuelan-share departments would have evolved along the same trajectory as outcomes in low-Venezuelan-share departments. Formally, I require:

$$\mathbb{E}\left[\Delta Y_{dt}(0) \mid \text{VenShare}_d\right] = \mathbb{E}\left[\Delta Y_{dt}(0)\right] \quad \forall\, t > 2021 \tag{1}$$

where $Y_{dt}(0)$ denotes the potential outcome in department d at time t absent regularization. This is the standard parallel trends assumption adapted for a continuous treatment design.

Three features support this assumption. First, the geographic distribution of Venezuelans was determined by pre-existing migration networks—border proximity and historical economic ties—rather than by contemporaneous labor market shocks. Second, the ETPV was a federal policy applied uniformly; there was no department-level discretion in implementation. Third, I verify the assumption empirically by testing for differential pre-trends in the 2015–2019 period.

4.2 Estimation

I estimate the following two-way fixed effects specification:

$$Y_{dt} = \alpha_d + \gamma_t + \beta\left(\text{VenShare}_d \times \text{Post}_t\right) + \varepsilon_{dt} \tag{2}$$

where Y_{dt} is the labor market outcome in department d and year t, α_d are department fixed effects, γ_t are year fixed effects, VenShare_d is the pre-treatment ETPV registration intensity, and $\text{Post}_t = \mathbb{I}[t \geq 2021]$. The coefficient β captures the effect of a one-percentage-point increase in Venezuelan share on the labor market outcome in the post-ETPV period, relative to the pre-period.

Standard errors are clustered at the department level. With only 23 clusters, conventional cluster-robust inference may over-reject (Cameron et al., 2008). I therefore report wild cluster bootstrap p-values using Webb weights and 9,999 iterations, following the recommendations of Roodman et al. (2019). All reported p-values are from the wild cluster bootstrap unless otherwise noted.

For the event study specification, I interact the treatment variable with a full set of year dummies, normalizing to the reference year 2019 ($t-2$):

$$Y_{dt} = \alpha_d + \gamma_t + \sum_{k \neq 2019} \delta_k \left(\text{VenShare}_d \times \mathbb{I}[t = k]\right) + \varepsilon_{dt} \tag{3}$$

The pre-treatment coefficients $\{\delta_k\}_{k<2021}$ test for differential trends prior to the ETPV. A joint F-test of $\delta_{2015} = \delta_{2016} = \delta_{2017} = \delta_{2018} = 0$ provides a formal assessment of the parallel trends assumption.

4.3 Threats to Validity

The primary concern is that Venezuelan settlement patterns may be correlated with department-level economic trajectories. If Venezuelans disproportionately settled in departments that were already experiencing differential labor market trends, the parallel trends assumption would be violated. The event study analysis directly addresses this concern by examining whether treatment-correlated trends existed prior to the ETPV.

A second concern is that the ETPV itself may have induced geographic reallocation of migrants across departments. If regularization caused Venezuelans to relocate from high-concentration to low-concentration departments (or vice versa), the treatment variable would be endogenous. However, PPT issuance was tied to the department of pre-registration, and geographic reallocation would, if anything, attenuate the estimated treatment effect by diluting the treatment contrast.

Third, the department-level unit of observation means that within-department compositional effects—such as displacement of specific demographic groups—may be masked in aggregate statistics. More fundamentally, the aggregate outcomes combine formal and informal employment. If the ETPV shifted Venezuelan workers from informal to formal status while displacing native formal workers into informality, the aggregate rates would be unchanged even though the composition shifted. This is a genuine limitation of using published department-level aggregates rather than individual-level microdata (which contains pension and health enrollment variables but requires manual download from DANE's micro-data portal). The aggregate perspective nonetheless captures the policy-relevant margin for evaluating the headline claim that regularization disrupts labor markets.

Fourth, the COVID-19 pandemic created unprecedented labor market disruptions in 2020. I exclude 2020 from the baseline specification but verify robustness to its inclusion. The pandemic could confound identification only if its effects were differentially persistent in high-Venezuelan-share departments, for which there is no evidence.

5. Results

5.1 Main Results

Table 2 reports the main difference-in-differences estimates. Across all four outcomes, the treatment effect is small and statistically insignificant. For the employment rate (column 1), the coefficient on VenShare $\times$ Post is -0.310 with a standard error of 0.400 and a wild cluster bootstrap p-value of 0.704. A one-percentage-point higher Venezuelan share is associated with a 0.31 percentage point decline in the employment rate post-ETPV, but this estimate cannot be distinguished from zero.

The unemployment rate estimate (column 2) is essentially zero: $\hat{\beta} = 0.026$ (SE $= 0.106$, $p = 0.815$). The participation rate estimate (column 3) mirrors the employment result: $\hat{\beta} = -0.293$ (SE $= 0.364$, $p = 0.654$). Underemployment (column 4) is also unaffected ($\hat{\beta} = -0.067$, SE $= 0.101$). The within-R^2 values are uniformly low (0.001–0.040), confirming that Venezuelan concentration explains negligible variation in post-ETPV labor market trajectories.

To assess statistical power, I compute the minimum detectable effect (MDE) at 80% power, following standard formulas for continuous-treatment DiD. The MDE for the employment rate is 1.12 percentage points per one-percentage-point increase in Venezuelan share. For Norte de Santander (Venezuelan share $\approx 17\%$), this translates to a detectable employment effect of approximately 19 percentage points—far larger than any plausible regularization effect. The study is therefore well-powered to detect economically meaningful effects, and the null finding is informative rather than merely agnostic.

5.2 Event Study

Table 3 reports the year-by-year event study coefficients. The pre-treatment coefficients are uniformly small and statistically insignificant for both employment and unemployment rates. For employment, the pre-period estimates range from -0.242 (2016) to -0.033 (2018), with

Table 2: Effect of ETPV Regularization on Department-Level Labor Market Outcomes

	Dependent Variable			
	Empl. Rate (1)	Unemp. Rate (2)	Partic. Rate (3)	Underempl. (4)
Ven. Share × Post	-0.310	0.026	-0.293	-0.067
	(0.400)	(0.106)	(0.364)	(0.101)
	[0.704]	[0.815]	[0.654]	—
Department FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	207	207	207	207
Within R^2	0.040	0.001	0.040	0.007

Notes: Each column reports a separate regression of the dependent variable on the interaction of pre-treatment Venezuelan share (ETPV registrations per population) with a post-2021 indicator. Panel of 23 Colombian departments, 2015–2024, excluding 2020 (COVID). Clustered standard errors by department in parentheses. Wild cluster bootstrap p-values (Webb weights, 9,999 iterations) in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

none approaching conventional significance levels. The immediate post-treatment coefficient in 2021 is near zero (-0.052). The joint F-test for all pre-treatment coefficients equaling zero yields $p = 0.291$ for employment and $p = 0.273$ for unemployment, providing no evidence against parallel trends.

The post-treatment coefficients show a slight widening of the point estimates—the employment coefficient grows from -0.052 in 2021 to -0.670 in 2024—but all remain statistically insignificant, and the standard errors grow proportionally. This pattern is consistent with gradually increasing noise in the post-period rather than a genuine dynamic treatment effect.

5.3 Robustness

Table 4 presents robustness checks for the employment rate, the outcome with the largest point estimate. The results are stable across specifications. Including 2020 barely changes the estimate (-0.310 versus -0.310). Adding department-specific linear trends attenuates the coefficient by half (-0.155, SE $= 0.364$). Using binary treatment (above-median Venezuelan share) yields a coefficient of 0.823 with a standard error of 2.287—a positive point estimate, but one that is purely noise. The top-quartile specification produces a negative but insignificant estimate (-1.751, SE $= 3.293$).

Leave-one-out analysis confirms that no single department drives the results. Excluding Norte de Santander (the most treated department) produces a larger negative point estimate (-0.585) but remains insignificant. Excluding both Norte de Santander and La Guajira flips

Table 3: Event Study: Year-by-Year Effects of Venezuelan Concentration

	Employment Rate		Unemployment Rate	
Event Year	Coeff.	SE	Coeff.	SE
$t-6$ (2015)	-0.198	(0.169)	0.117	(0.108)
$t-5$ (2016)	-0.242	(0.193)	0.153	(0.129)
$t-4$ (2017)	-0.073	(0.098)	0.070	(0.072)
$t-3$ (2018)	-0.033	(0.105)	0.011	(0.042)
$t-2$ (2019)	Reference Year			
t (2021)	-0.052	(0.205)	-0.011	(0.093)
$t+1$ (2022)	-0.377	(0.407)	0.072	(0.141)
$t+2$ (2023)	-0.578	(0.612)	0.178	(0.297)
$t+3$ (2024)	-0.670	(0.575)	0.145	(0.196)
Pre-trend F-test	$p = 0.291$		$p = 0.273$	

Notes: Coefficients from regressing the outcome on interactions of Venezuelan share with year dummies, with department and year fixed effects. Reference year is 2019 ($t-2$). 2020 excluded due to COVID. Pre-trend F-test reports the p-value from testing that all pre-treatment interactions equal zero. Standard errors clustered by department.

Table 4: Robustness Checks: Employment Rate

Specification	Coefficient	SE	N
Baseline (excl. 2020)	-0.310	(0.400)	207
Including 2020	-0.310	(0.394)	230
Dept.-specific trends	-0.155	(0.364)	207
Binary treatment	0.823	(2.287)	207
Top quartile treatment	-1.751	(3.293)	207
Excl. Norte de Santander	-0.585	(0.510)	198
Excl. NdS + La Guajira	0.747	(0.645)	189
Placebo (2018 treatment)	0.154	(0.161)	138

Notes: Each row reports the coefficient on the treatment variable from a separate regression of the employment rate (TO) with department and year fixed effects. “Binary treatment” uses an above-median Venezuelan share indicator. “Top quartile” uses the highest-concentration quartile. Placebo uses a 2018 treatment date on pre-period data only. Standard errors clustered by department. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

the sign to positive (0.747), though again insignificant. The placebo test, assigning treatment at 2018 using only pre-period data, yields a coefficient of 0.154 (SE = 0.161)—confirming no spurious pre-trends.

5.4 Heterogeneity

Table 5 explores whether ETPV effects vary with baseline labor market conditions. Column 1 interacts the treatment with above-median baseline underemployment. The interaction is negative (-0.458, SE = 0.380) but not statistically significant, suggesting no clear differential effect by labor market slack.

Table 5: Heterogeneity: Employment Effects by Baseline Labor Market Conditions

	Employment Rate	
	By Baseline Underempl.	By Baseline Employment
	(1)	(2)
Ven. Share × Post	0.148	-0.002
	(0.481)	(0.183)
× High Baseline	-0.458	-0.679**
	(0.380)	(0.272)
Sum: High baseline	-0.310	-0.681
Department FE	Yes	Yes
Year FE	Yes	Yes
Observations	207	207

Notes: Column (1) interacts the treatment with an indicator for above-median baseline (2019) underemployment rate. Column (2) interacts with above-median baseline employment rate. "Sum: High baseline" reports the total effect for departments with high baseline values. Standard errors clustered by department. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Column 2 reveals the one instance of heterogeneity. In departments with above-median baseline employment rates, the interaction term is -0.679 ($p < 0.05$). The total effect for these departments is -0.681 percentage points per unit of Venezuelan share. This finding is consistent with a competitive pressure mechanism: in departments where formal labor markets were already tight, the injection of newly eligible formal workers may have generated modest crowding. However, this result should be interpreted cautiously given the number of heterogeneity splits examined and the absence of a pre-registered analysis plan.

6. Discussion

The central finding of this paper—that the largest mass regularization in Latin American history produced no detectable aggregate labor market disruption—requires interpretation. I consider three explanations, from most to least plausible.

Informal sector absorption. The most likely explanation is that the Colombian labor market had already absorbed the Venezuelan migration shock through its large informal sector. Prior to the ETPV, most Venezuelan workers were employed informally, and their presence had already been reflected in equilibrium employment levels. Regularization changed the *legal* channel through which these workers participated but did not change the *quantity* of labor supplied. If regularization shifted some workers from informal to formal employment gradually, the aggregate rates—which combine formal and informal employment—would be unaffected. This interpretation is consistent with Ulyssea (2020), who models informal and formal labor markets as interconnected, and with evidence from Lebow (2024) showing that Venezuelan arrival increased native informality.

Slow take-up and the binary post indicator. The null may partly reflect the gradual pace of ETPV implementation. Although 2.46 million pre-registrations were recorded by November 2021, actual PPT issuance was slower—1.4 million by September 2022, with continued rollout through 2023. The binary post-2021 indicator treats this staggered rollout as simultaneous, potentially attenuating the treatment effect. A staggered-adoption design exploiting the month-level PPT issuance data—in the spirit of Callaway and Sant'Anna (2021)—could better capture the actual timing of treatment receipt. If formalization effects require time to materialize—as employers learn about workers' legal status, workers navigate bureaucratic processes, and labor market institutions adjust—the three post-treatment years in my data may understate the long-run impact. This concern is partially mitigated by the event study, which shows no evidence of growing effects through 2024, but it cannot be fully resolved without finer temporal variation.

Statistical power limitations. Although the minimum detectable effect analysis suggests the study can detect economically meaningful effects, the 23-department panel is inherently limited. Effects of 0.3–0.5 percentage points per unit of Venezuelan share—modest but potentially policy-relevant—fall below the detection threshold. I am transparent about this limitation: the null is informative for large effects but does not rule out small ones.

The heterogeneity finding—negative effects concentrated in high-employment departments—is consistent with theoretical predictions from dual labor market models. When formal labor markets are tight, the addition of legally authorized workers may generate crowding effects that are absent when surplus informal-sector capacity can absorb the adjustment. This interpretation links the regularization question to the broader literature on labor market institutions and immigration (Meghir et al., 2015; Boeri and Brücker, 2011).

7. Conclusion

This paper names a phenomenon: the regularization illusion. The widespread expectation that converting irregular immigrants to legal status will visibly disrupt host-country labor markets is not supported by the evidence from Colombia's ETPV, the largest regularization program in Latin American history. Across four labor market outcomes, 23 departments, nine years of data, and multiple specifications, I find precisely estimated null effects.

The finding carries a clear policy implication. Regularization programs are often blocked or delayed by fears of native labor market harm. The Colombian experience suggests that these fears are misplaced—at least in economies with substantial informal sectors, where the labor supply adjustment has already occurred. Countries facing large irregular migrant populations—including the United States, members of the European Union, and Turkey—may be able to pursue regularization without the aggregate labor market costs that opponents predict.

This does not mean regularization has no effects. It may improve immigrant welfare through access to labor protections, social security, and financial services—precisely the channels that aggregate employment statistics do not capture. Moreover, the aggregate null may mask compositional shifts: if formalized immigrants displaced natives from formal to informal employment, the net effect on aggregate rates would be zero even though the distributional consequences were real. Future research should investigate these compositional and welfare effects using individual-level GEIH microdata—which contains pension enrollment (P6920) and health insurance (P6090) variables that directly measure formality—and exploit the time-varying PPT issuance data in a staggered-adoption framework.

The null, properly identified, is the finding.

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: @ai1scl

First Contributor: https://github.com/ai1scl

References

Amuedo-Dorantes, Catalina, Esther Arenas-Arroyo, and Almudena Sevilla, "Immigration Enforcement, Hispanic Self-Employment, and Labor Supply of Undocumented Workers," *Labour Economics*, 2020, *63*, 101791.

Boeri, Tito and Herbert Brücker, "Immigration, Offshoring and American Jobs," *American Economic Review*, 2011, *101* (5), 1–24.

Borjas, George J., "The Labor Demand Curve Is Downward Sloping: Reexamining the Impact of Immigration on the Labor Market," *Quarterly Journal of Economics*, 2003, *118* (4), 1335–1374.

_ , "The Wage Impact of the Marielitos: A Reappraisal," *ILR Review*, 2017, *70* (5), 1077–1110.

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

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller, "Bootstrap-Based Improvements for Inference with Clustered Errors," *Review of Economics and Statistics*, 2008, *90* (3), 414–427.

Card, David, "The Impact of the Mariel Boatlift on the Miami Labor Market," *ILR Review*, 1990, *43* (2), 245–257.

_ , "Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration," *Journal of Labor Economics*, 2001, *19* (1), 22–64.

Caruso, German, Carlos Gomez Canon, and Valerie Mueller, "Spillover Effects of Immigration: Evidence from Venezuelan Exodus on Colombian Labor Markets," *World Bank Economic Review*, 2021, *35* (3), 832–853.

Clemens, Michael A. and Jennifer Hunt, "The Immigration Wage Effect After All: Revisiting the Mariel Boatlift," *Annual Review of Economics*, 2019, *11*, 135–157.

_ **, Ethan G. Lewis, and Hannah M. Postel**, "Immigration Restrictions as Active Labor Market Policy: Evidence from the Mexican Bracero Exclusion," *American Economic Review*, 2018, *108* (6), 1468–1487.

Cobb-Clark, Deborah A., Clinton R. Shiells, and B. Lindsay Lowell, "Immigration Reform and the Wages of Latino Workers," *Journal of Human Resources*, 1995, *30* (3), 510–524.

Dustmann, Christian, Uta Schönberg, and Jan Stuhler, "The Labour Market Effects of Immigration," *Economic Journal*, 2017, *126* (592), F218–F255.

Ibáñez, Ana María, Andrés Moya, María Andrea Ortega, Sandra V. Rozo, and María José Urbina, "Life Out of the Shadows: Impacts of Amnesties in the Lives of Refugees," Technical Report 10049, World Bank Policy Research Working Paper 2022.

Kossoudji, Sherrie A. and Deborah A. Cobb-Clark, "Coming Out of the Shadows: Learning About Legal Status and Wages from the Legalized Population," *Journal of Labor Economics*, 2002, *20* (3), 598–628.

Lebow, Jeremy, "Immigration and Occupational Downgrading in Colombia," *Journal of Development Economics*, 2024, *166*, 103196.

Meghir, Costas, Renata Narita, and Jean-Marc Robin, "Wages and Informality in Developing Countries," *American Economic Review*, 2015, *105* (4), 1509–1546.

Monras, Joan, "Immigration and Wage Dynamics: Evidence from the Mexican Peso Crisis," *Journal of Political Economy*, 2020, *128* (8), 3017–3089.

Ottaviano, Gianmarco I.P. and Giovanni Peri, "Rethinking the Effect of Immigration on Wages," *Journal of the European Economic Association*, 2012, *10* (1), 152–197.

Peñaloza-Pacheco, Leonardo, "Venezuelan Migration in Colombia: Impact on the Labor Market," *Latin American Journal of Economic Development*, 2023, *40*, 79–120.

Peri, Giovanni and Chad Sparber, "Task Specialization, Immigration, and Wages," *American Economic Journal: Applied Economics*, 2009, *1* (3), 135–169.

R4V Inter-Agency Coordination Platform, "Refugee and Migrant Response Plan 2021," Technical Report, UNHCR/IOM 2021.

Roodman, David, Morten Ørregaard Nielsen, James G. MacKinnon, and Matthew D. Webb, "Fast and Wild: Bootstrap Inference in Stata Using `boottest`," *Stata Journal*, 2019, *19* (1), 4–60.

Rozo, Sandra V. and Juan F. Vargas, "Is Formality a Good Proxy for Workers' Welfare? Evidence from Venezuelan Refugees in Colombia," *Journal of Public Economics*, 2024, *230*, 105042.

Ulyssea, Gabriel, "Informality," *Annual Review of Economics*, 2020, *12*, 525–546.

World Bank, "Socioeconomic Integration of Venezuelan Migrants and Refugees in Colombia," Technical Report, World Bank Group, Washington, DC 2021.

A. Data Appendix

A.1 Labor Market Data

Department-level labor market indicators are drawn from the DANE Gran Encuesta Integrada de Hogares (GEIH) Department Annex, published annually by the Departamento Administrativo Nacional de Estadística. The GEIH is Colombia's principal household labor force survey, with a rotating panel design covering approximately 240,000 households per year across 23 departments. The Department Annex reports aggregated annual averages of the four primary labor market indicators: employment rate (*tasa de ocupación*, defined as employed persons / working-age population), unemployment rate (*tasa de desocupación*, defined as unemployed persons / economically active population), labor force participation rate (*tasa global de participación*, defined as economically active population / working-age population), and underemployment rate (*tasa de subempleo subjetivo*).

Data were accessed from the DANE statistical archive for years 2015–2024. All rates are expressed as percentages. The 23 departments covered by the GEIH Department Annex exclude several smaller departments and special territories for which the GEIH sample is insufficient to produce reliable annual estimates.

A.2 ETPV Pre-Registration Data

Treatment intensity is constructed from the ETPV pre-registration microdata published by Migración Colombia on the Colombian open data portal (datos.gov.co). The dataset contains 2,463,531 individual pre-registration records with fields for department of residence, nationality, date of registration, and basic demographic information. Records were aggregated to the department level and divided by the DANE 2019 population projection for each department to construct the treatment variable: ETPV registrations per capita.

A.3 Variable Definitions

- **Employment Rate** (*Tasa de Ocupación*, TO): Employed persons as a percentage of the working-age population (ages 12+ in urban areas, 10+ in rural areas), following DANE's definition.
- **Unemployment Rate** (*Tasa de Desocupación*, TD): Unemployed persons (seeking work in the reference period) as a percentage of the economically active population.
- **Participation Rate** (*Tasa Global de Participación*, TGP): Economically active population (employed + unemployed) as a percentage of the working-age population.

- **Underemployment Rate** (*Tasa de Subempleo Subjetivo*): Workers who consider themselves underemployed (by hours, income, or competencies) as a percentage of the employed population.

- **Venezuelan Share** (treatment): Total ETPV pre-registrations in the department divided by 2019 departmental population $\times$ 100.

- **Post**: Indicator equal to 1 for years $\geq$ 2021, 0 otherwise.

A.4 Sample Construction

The analysis sample is a balanced panel of 23 departments $\times$ 10 years (2015–2024). The year 2020 is excluded from the baseline specification due to COVID-19 lockdown disruptions, yielding 207 department-year observations. Robustness checks including 2020 use the full 230 observations.

B. Identification Appendix

B.1 Pre-Trend Analysis

The event study specification (Equation (3)) provides a direct test of the parallel trends assumption. With 2019 as the reference year, the pre-treatment coefficients (δ_{2015} through δ_{2018}) measure whether labor market outcomes in high-Venezuelan-share departments diverged from those in low-share departments prior to the ETPV. Joint F-tests for the null hypothesis $H_0 : \delta_{2015} = \delta_{2016} = \delta_{2017} = \delta_{2018} = 0$ yield p-values of 0.291 (employment rate) and 0.273 (unemployment rate), providing no evidence against parallel trends.

B.2 Placebo Test

As a further validation, I estimate the main specification using only pre-ETPV data (2015–2019) with a placebo treatment date of 2018. If the identifying variation were driven by spurious trends correlated with Venezuelan settlement, this placebo test should detect them. The placebo coefficient for the employment rate is 0.154 (SE = 0.161), consistent with no pre-existing differential trends.

C. Robustness Appendix

The robustness analysis examines sensitivity along five dimensions. First, sample definition: including the COVID year 2020 does not materially change the estimates, as the pandemic shock was approximately symmetric across departments. Second, trend controls: adding department-specific linear time trends absorbs any slow-moving differential trends, attenuating the point estimate toward zero. Third, treatment definition: alternative constructions—binary (above/below median) and top-quartile indicators—produce qualitatively similar null results, with substantially wider confidence intervals reflecting the loss of continuous variation. Fourth, outlier sensitivity: leave-one-out analysis excluding Norte de Santander and La Guajira (the two most treated departments) demonstrates that results are not driven by extreme treatment values. Fifth, temporal placebo: assigning treatment at 2018 in the pre-period produces no significant effects, confirming the absence of spurious trends.

D. Heterogeneity Appendix

The heterogeneity analysis interacts the treatment with two baseline labor market conditions measured in 2019: underemployment rate and employment rate. The interaction with baseline underemployment is negative but insignificant, suggesting no differential effect in markets with greater slack. The interaction with baseline employment is negative and significant at the 5% level, suggesting modest competitive pressure in tight labor markets. This finding should be interpreted with caution given the small number of departments and the absence of pre-registration for the heterogeneity analysis.

E. Standardized Effect Sizes

Table 6: Standardized Effect Sizes

Outcome	$\hat{\beta}$	SE	SD(Y)	SDE	SE(SDE)	Classification
Panel A: Pooled						
Employment Rate	-0.310	0.400	5.91	-0.2181	0.2814	Large negative
Unemployment Rate	0.026	0.106	2.51	0.0433	0.1754	Small positive
Participation Rate	-0.293	0.364	5.44	-0.2240	0.2782	Large negative
Panel B: Heterogeneous (by baseline employment rate)						
Employment Rate (High Baseline)	-0.681	0.328	5.91	-0.4796	0.2310	Large negative
Employment Rate (Low Baseline)	-0.002	0.183	5.91	-0.0017	0.1290	Null

Notes: **Country:** Colombia. **Research question:** Does mass regularization of Venezuelan immigrants through the ETPV program affect aggregate labor market outcomes in Colombian departments? **Policy mechanism:** The ETPV (Estatuto Temporal de Protección para Migrantes Venezolanos) granted 10-year work permits to approximately 1.8 million Venezuelan migrants, enabling formal employment, social security enrollment, and financial inclusion for a population previously working almost entirely in the informal sector. **Outcome definition:** Employment rate (tasa de ocupación), unemployment rate (tasa de desocupación), and labor force participation rate (tasa global de participación) from DANE GEIH. **Treatment:** Continuous; department-level ETPV pre-registrations per 2019 population, interacted with a post-2021 indicator. **Data:** DANE GEIH Department Annex (2015–2024) and Migración Colombia ETPV Registry; 23 departments, 207 department-year observations (excluding 2020). **Method:** Two-way fixed effects (department + year) with continuous treatment intensity; standard errors clustered by department with wild cluster bootstrap (Webb weights). **Sample:** All 23 Colombian departments with GEIH coverage, annual frequency, excluding 2020 due to COVID lockdown disruptions. SDE = $\hat{\beta} \times \mathrm{SD}(X)/\mathrm{SD}(Y)$ where SD(X) is the pre-treatment standard deviation of Venezuelan share and SD(Y) is the pre-treatment standard deviation of the outcome. Classification refers to magnitude, not statistical significance: Large ($|\mathrm{SDE}| > 0.15$), Moderate (0.05–0.15), Small (0.005–0.05), Null (< 0.005).