Shock and No Shift: Impact of the Expulsion Initiative on Foreign Crime Rates in Switzerland [1]

Souhil Khiat, Simon Schmidt-Ginzkey, Martin Sinishta, Tommaso Spreij, Timon van Engen[2]

HEC Lausanne
Faculty of Business & Economics
University of Lausanne

June 2025

Abstract

This paper examines whether Switzerland's 2016 expulsion law for foreign criminals affected violent crime rates among foreign residents. Using annual canton-level data (2010–2023) and a difference-in-differences design comparing foreigners to Swiss nationals, we find no statistically significant impact of the reform. This null result holds under various checks (pre-trend tests, placebo years, canton-specific effects, and alternative functional forms). Our findings imply that the enacted law, despite its stricter rhetoric, did not change criminal behavior among foreigners. Policies emphasizing integration and socioeconomic support may be more effective for reducing crime.

[1]The findings are data-driven and do not reflect any normative stance.

[2]souhil.khiat@unil.ch, simon.schmidt-ginzkey@unil.ch, martin.sinishta@unil.ch, tommaso.spreij@unil.ch, timon.vanengen@unil.ch

1 Introduction

On October 1st, 2016, Switzerland implemented a symbolic reform of its migration policy, marked by the implementation of the law resulting from the popular initiative of the *Union Démocratique du Centre* (UDC) party "Pour le renvoi des étrangers criminels" (For the expulsion of foreign criminals), which was approved by 52.9% of the electorate in 2010 [3]. The aim of this initiative was to enshrine in the Swiss Constitution the automatic expulsion of any foreign national convicted of certain serious offenses [11]. The intense public debate that emerged prior to the 2010 vote, as well as the parliamentary and legal debates that followed the vote, ultimately resulted in a final text that was more flexible than the one initially proposed: the law still provides for compulsory expulsion in many cases, but a rigour clause has been introduced giving judges the right to waive expulsion where it would have serious personal consequences [4].

The main objective of this reform was to reduce crime among foreigners by introducing a direct and systematic threat: expulsion from Swiss territory. These events provide an opportunity for econometric analysis, as the implementation of this law was sudden, nationwide, and applied only to one group of the population, non-nationals.

Therefore, our study raises the following research question:

Did the implementation of the expulsion law for violent crimes reduce criminality among foreigners in Switzerland?

Our project applies a Difference-in-Differences (DiD) method to the evolution of the crime rate of foreigners compared with that of Swiss nationals, before and after the implementation of the reform. To do this, we use a panel of cantonal crime data for the period 2010 to 2023, taken from the Police Crime Statistics (*Statistique policière de la criminalité* - SPC), available via the official and reliable STAT-TAB platform of the Federal Statistical Office[9](FSO). This rich database allows us to track crime objectively by canton, year, and population group. Our econometric model also includes fixed effects to control for unobserved factors that remain constant over time.

1.1 Previous literature

Despite the importance of the debates that have animated political life, no rigorous empirical study seems to have been carried out to estimate the concrete effect of this law on the criminality of foreigners. Our study aims to fill this gap. On the European side, Bell et al (2013)[2] compared two waves of migration to the UK, each of which had different access to the labour market. They concluded that access to the legal labour market plays a key role in reducing crime. Pinotti (2017)[10], on the other hand, studied the impact of regularising immigrants in Italy, based on a quota system. He compared migrants whose applications were accepted with those whose applications were narrowly rejected. He concluded that obtaining legal status, which gives access to the legal labour market, reduces crime, particularly among the most vulnerable. And recently in Switzerland, Auer et al (2024)[1] studied the impact of cantonal differences in the social assistance paid to refugees.

They concluded that more generous social assistance reduces criminality among those who receive it, even without professional integration. Financial support therefore plays an important role in preventing crime.

These various studies therefore show that the effects of migration policies on crime are highly dependent on the legal and economic background of migrants. Our project is part of this wider literature on the effects of migration policies on crime by foreigners and makes our contribution twofold: it is the first empirical evaluation of the implementation of this initiative, and our approach tests the deterrent effectiveness of this major reform.

The continuation of this work is structured as follows: we will present the institutional context and the data used and we will then detail our empirical strategy. We will continue by presenting and discussing our results, before concluding in a general way.

2 Context and Data

2.1 Context

2.1.1 Popular Initiative and Constitutional Amendment

The initiative's legal text approved by the swiss people inserted new paragraphs 3–6 into Article 121 of the Federal Constitution, and a transitional provision (Art. 197 § 8). The key additions are[11]:

- Art. 121 § 3: Foreign nationals, irrespective of their residence status, lose their permit and all rights to stay in Switzerland if they are convicted by a final judgment of murder, rape or any other serious sexual offence; other violent offences (robbery, trafficking in human beings or drugs, burglary); or the fraudulent receipt of social insurance or assistance benefits.

- Art. 121 § 4: Parliament must define the offences listed in § 3 in more detail and may extend the list.

- Art. 121 § 5: Those deprived of their residence rights under §§ 3–4 must be expelled and banned from re-entry for 5–15 years (20 years in case of repeat offence).

- Art. 121 § 6: Any person who violates the re-entry ban or enters Switzerland illegally commits a punishable offence; Parliament must enact the corresponding sanctions.

To give effect to the new constitutional rules, Parliament amended both the Foreign Nationals Act (LEtr, RS 142.20) and the Criminal Code on March 16th 2016 (effective October 2016). The Criminal Code now includes Art. 66a CP[4], which makes expulsion of foreign nationals convicted of specified serious offences mandatory for 5–15 years, and the Foreign Nationals Act tasks cantonal justice and police authorities, in coordination with the Federal Office for Migration, with executing expulsions and imposing re-entry bans in accordance with the constitutional mandates.

2.1.2 Geographical and Institutional Framework

Switzerland comprises 26 cantons, each with substantial autonomy over policing and sentence enforcement. While the Confederation sets the general legal framework for entry, residence, and expulsion, actual expulsions and entry-ban enforcement are carried out at the cantonal level. Although being part of the Schengen Area, Switzerland maintains spot checks at its borders and cooperates closely with neighboring EU/EEA states to identify and return individuals subject to entry bans.

2.1.3 Socio-Political Background

This initiative forms part of a broader series of Swiss popular votes on migration policy (e.g., the 9 February 2014 "Against Mass Immigration" initiative[13]). Advocates framed it as necessary for national security and sovereignty, while opponents warned of risks to social integration, rule of law, and Switzerland's international reputation.

2.2 Data - Key variables

Our dependent variable is the annual police-recorded rate of violent offences subject to automatic expulsion under Art. 66a (e.g., homicide, serious bodily harm, rape/sexual assault), per 100 000 residents in canton c, residency group g $\big(\text{Swiss} = 0,\ \text{foreign resident} = 1\big)$, and year t. Observations with a zero count are discarded to avoid undefined rates, so the series is strictly positive. Because the distribution is right-skewed, we analyse the variable in two forms: the original level $CrimeRate_{cgt}$ (policy-relevant units) and its natural logarithm $\log(CrimeRate_{cgt})$, which puts percentage changes on a linear scale and improves homoskedasticity. Our estimated econometric model is as follows:

$$\log\big(CrimeRate_{cgt}\big) = \beta\,\big(T_g \times P_t\big) + \alpha_c + \lambda_g + \gamma_t + \varepsilon_{cgt}\,.$$

Where:

- $\log(CrimeRate_{cgt})$ is the crime rate observed in canton c, for group g, in year t, expressed in natural logarithms.

- T_g is a binary variable which takes the value 1 for foreigners (the treated group) and 0 for Swiss nationals (the control group).

- P_t is a binary variable that takes the value 1 for post-reform years (from 2017 onwards) and 0 before that.

- $T_g \times P_t$ is the model's key interaction; its coefficient β captures the specific effect of the reform on the treated group (when both $T_g = 1$ and $P_t = 1$).

- α_c denotes canton fixed effects, controlling for all time-constant unobservable characteristics specific to each canton (e.g., demographic structure, average wealth, local culture, etc.).

- λ_g are group fixed effects capturing structural level differences between Swiss nationals and foreigners (average crime rates, underlying social or economic conditions, etc.).

- γ_t are time fixed effects absorbing all shocks common to all cantons and groups in a given year (e.g., national legislative changes, macroeconomic developments, COVID-19, etc.).

- ε_{cgt} is the cluster error term at the canton level, capturing specific unobserved shocks affecting the crime rate in canton c, group g, year t, after accounting for all fixed effects and the interaction. It includes random and unobserved influences that vary over time and across units.[3]

We define 2016 as pre-treatment because the expulsion law only took effect on October 1st 2016 and our data are annual, so including 2016 in the post-period would misattribute only a quarter of the year's exposure to the reform. By construction, $T_g \times P_t$ varies only across residency groups and over time, not within a canton–year cell. Throughout, β on $T_g \times P_t$ is interpreted as the average causal effect of the expulsion regime on foreign-resident crime relative to the Swiss counterfactual.

No external instruments are used. Identification relies entirely on the two-way-fixed-effects DiD framework outlined above and on the credibility of the parallel-trends tests reported elsewhere in the paper.

2.3 Data Sample

Our analysis begins with the complete set of offence counts and population totals released by the Swiss Federal Statistical Office's STAT-TAB portal [9]. Each raw record is uniquely defined by the canton in which the offence occurred (26 cantons), the residency status of the population (Swiss or foreign resident), and the calendar year ($t \in \{2010, 2011, \ldots, 2023\}$). Before any editing this structure yields 728 canton–group–year cells. All variables are harmonised prior to merging: French labels for residency status are translated into English, suppression symbols that STAT-TAB uses for confidentiality (X, x:, *) are recoded as missing, and numeric fields are coerced to numeric types. After the offence and population tables are joined, we construct the rate

$$CrimeRate_{cgt} \;=\; 100\,000 \times \frac{offence\_count_{cgt}}{population_{cgt}}$$

Rows are removed when a key variable is unobservable. Eight canton–group–year cells contain confidentiality codes in the offence-count column and are therefore dropped, leaving 720 observations prior to the low-rate screen. Because very small rates generate unstable logarithms, the analysis is limited to canton–group–year cells for which

$$\log(CrimeRate_{cgt}) \;\geq\; 0 \quad\Longleftrightarrow\quad CrimeRate_{cgt} \;\geq\; 1 \text{ per } 100\,000 \text{ inhabitants.}$$

Observations below this threshold, including all genuine zero–count cells, are omitted; the resulting loss of data is modest and has no bearing on the treatment definition. Rows are therefore excluded only for two reasons: confidentiality codes in the offence-count column and crime-rate values below the 1-per-100'000 cut-off.

[3]The statistical properties and implications of the error term are discussed in detail in section 4.2, which addresses potential threats to identification and validity.

A critical measurement convention of the Police Crime Statistics is that offences are allocated to the canton where the act was committed, irrespective of the offender's domicile. Thus a robbery perpetrated in Neuchâtel by a resident of Vaud is counted in Neuchâtel's tally. This *locus-delicti* rule is orthogonal to our year-based treatment indicator, so it does not contaminate the before-after contrast that identifies the policy effect, and we do not dwell on its implications here.

Descriptive statistics

Table 1: Descriptive statistics for the cleaned panel

Variable	Mean	Std. dev.	Min	Max
Crime rate per 100 000 – *Foreigners*	582.3	118.7	310.1	985.4
Crime rate per 100 000 – *Swiss*	429.8	92.5	210.7	772.2

The table reveals that, on average, foreign residents have higher violent crime rates than Swiss nationals, this underpins the credibility of the difference-in-differences strategy developed in the next section.

3 Descriptive Analysis

3.1 The Difference-in-Differences (DiD) framework

3.1.1 Justification and specification of the basic model

To estimate the causal effect of the reform on the crime rate of foreigners, we choose the empirical Difference-in-Differences (DiD) method. This method is well suited to a quasi-natural experiment framework. Indeed, the implementation of the initiative, i.e., an institutional reform, affects only one group of the population (foreigners) while the other group (Swiss) is not affected. This change occurs at a precise moment in time, and as previously mentioned, we take 2017 as the pivotal year in our model.

3.1.2 Interpretation of β

The basis of our identification is the β coefficient, estimated from a specification where the dependent variable is expressed as a logarithm. In this framework, β represents a semi-elasticity: it measures the percentage change in the crime rate of foreigners relative to that of the Swiss, following the introduction of the reform, having taken account of the general trends that the two groups have in common.

3.1.3 Justification for the control group

The choice of the control group (Swiss nationals) is strongly justified for several reasons. Although Swiss nationals are not subject to the same immigration rules as foreign residents, they reside in the same cantons and time periods, sharing many economic, social, and policing conditions. By including canton, year, and group fixed effects in our model, we account for these common factors. Consequently, the Swiss serve as a valid counterfactual, capturing the 'normal' crime trend absent the reform. This comparison bolsters the assumption that, in the absence of the law, both groups would have followed similar trajectories.[4]

3.1.4 Functional robustness

To ensure that our results do not depend on the specific form chosen for the dependent variable, we apply a Box–Cox transformation, a parametric method that empirically determines the best monotonic transformation of the data. This approach generalizes several usual forms: for example, a parameter $\lambda = 1$ corresponds to a specification in levels, while $\lambda = 0$ corresponds to a logarithmic transformation (because $\log(x)$ is the limiting case of Box–Cox when $\lambda \to 0$).

The Box–Cox transformation is defined as follows:

$$CrimeRate_{cgt}{}^{(\lambda)} = \begin{cases} \dfrac{CrimeRate_{cgt}{}^{\lambda} - 1}{\lambda}, & \lambda \neq 0, \\ \log\big(CrimeRate_{cgt}\big), & \lambda = 0. \end{cases}$$

This allows the data to determine the optimal functional form, maximising the log-likelihood of the model over the pre-processing period. In our case, the optimisation carried out over the years 2010 to 2016 indicates that the optimal value of the parameter is $\lambda \approx 0.25$, obtained using the maximum likelihood method, which consists of choosing the Box-Cox transformation that minimises the specification error, i.e. the one that maximises the statistical adequacy of the model (in particular the log-likelihood) over the estimation period, as shown by the likelihood profile illustrated in the figure below.[5]

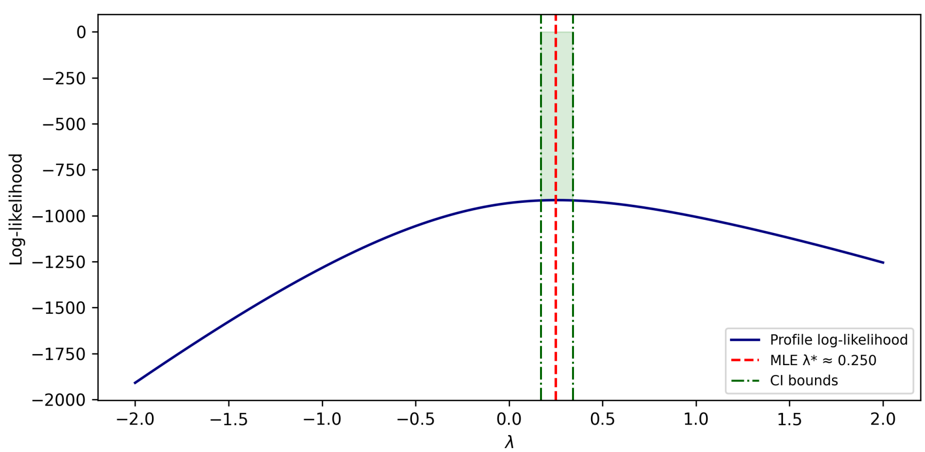

Figure 1: Box–Cox Profile with 95% CI (2010 - 2016)

[4]This identifying assumption is examined in greater detail in section 3.2
[5]The distribution of crime rates under alternative functional forms is shown in Appendix J

This value lies between the pure levels ($\lambda = 1$) and the logarithm ($\lambda = 0$), suggesting that neither of these two extreme specifications perfectly captures the relationship between treatment and crime rates. Nevertheless, as $\lambda = 0.25$ remains close to zero, the logarithmic form is an excellent approximation and has the advantage of being more directly interpretable (in percentage terms) and more commonly used in the empirical literature.

3.2 Identification hypothesis: parallel trends

3.2.1 Visualisation of pre-trends (logs)

The fundamental hypothesis of the DiD method is that of parallel trends. It assumes that, without the reform, Swiss and foreign nationals would have experienced similar trends in crime rates.

To test this hypothesis visually, we plot the time trend in the logarithm of crime rates for each group between 2010 and 2016.

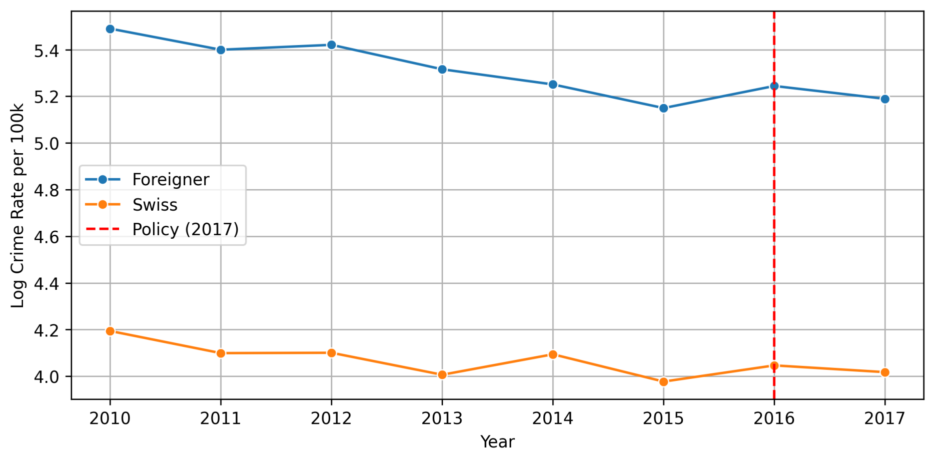

Figure 2: Pre-2017 Log Crime Rate by Group

The figure above shows that the two curves follow a similar logarithmic dynamic, except for 2014. One hypothesis to explain this anomaly is the reform of police crime statistics (SPC) that took place that year[5]. This restructuring of the databases may have led to a break in the series or a temporary change in the way offences were recorded, affecting the two groups differently.

3.2.2 Linear Interaction Test

In addition to the graphical inspection, we assess the parallel trends hypothesis more formally by estimating a model with a linear interaction between the group (foreign/Swiss) and time, restricted to the period 2010-2016. The aim is to examine whether the two groups had, on average, a significantly different slope over time before the introduction of the reform.

The estimated model is as follows:

$$\log(CrimeRate_{gt}) = \alpha + \gamma T_g + \lambda t + \beta(T_g \times t) + \varepsilon_{gt}$$

Where:

- $CrimeRate_{gt}$ is the crime rate for group g (foreigners or Swiss) in year t.
- T_g is a binary variable equal to 1 for foreigners and 0 for Swiss nationals.
- t is a time variable indicating the year in increasing order; it captures linear time trends.
- β measures whether the temporal slope differs for foreigners compared to Swiss nationals.

The estimated interaction coefficient is $\beta = -0.0255$, with a p-value of 0.101, which indicates that the slope of change for foreigners prior to 2017 is not significantly different from that for the Swiss. In simpler terms, on average, the two groups followed comparable linear dynamics before the reform[6].

This test supports the visual observations of the previous section and reinforces the idea that the hypothesis of parallel trends is reasonable in our empirical framework. It also provides a useful complement to the event-based test presented in the next section, which analyses the differences year by year.

3.2.3 Event-study test

To assess the robustness of the parallel trends hypothesis, we estimate an event-study model that measures, year by year, the difference in trajectory between foreigners and Swiss nationals before the introduction of the reform. This model refines our diagnosis by directly identifying any anticipations or divergences prior to 2017.

The estimated model is as follows:

$$\log(CrimeRate_{cgt}) = \gamma T_g + \sum_{y=2010}^{2016} \delta_y \cdot \mathbf{1}\{Year = y\} + \sum_{y=2010}^{2016} \theta_y \cdot (T_g \times \mathbf{1}\{Year = y\}) + \alpha_c + \varepsilon_{cgt}$$

Where:

- $CrimeRate_{cgt}$ is the crime rate in canton c, for group g (foreigners or Swiss), in year t.
- T_g is a binary variable equal to 1 if the group is foreign, and 0 if Swiss.
- δ_y are the year fixed effects, absorbing time-specific shocks common to all groups.
- θ_y are the interaction coefficients between the foreign group and year y; they measure the relative difference between foreigners and Swiss for each year, compared to the base year (here 2010, omitted in the estimation).

[6] Appendix A

The aim here is to check whether the coefficients θ_y for the years prior to 2017 are significantly different from zero. An absence of significant difference would indicate that the groups followed comparable trajectories before the reform, a key condition for the validity of the DiD design.

The figure below shows the estimates of the θ_y coefficients for the years 2011 to 2016 (2010 being the base year), with their 95% confidence intervals.

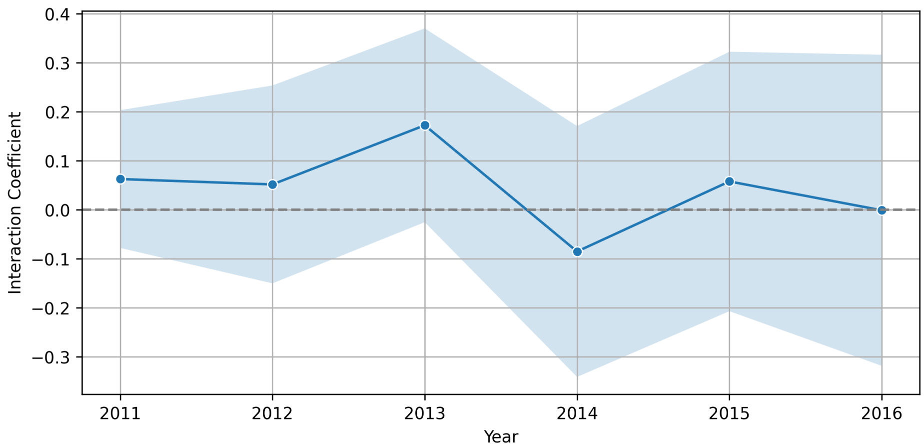

Figure 3: Event-Study Pre-Trends Dummies with 95% CI (log-lin specification)

We can see that all the coefficients are close to zero, and that the confidence intervals systematically include the zero value. For example, in 2013, although the coefficient is slightly positive, the uncertainty is such that it is not statistically different from zero. This lack of anticipated divergence supports the hypothesis of no pre-reform differential treatment.

We perform a joint nullity test of the coefficients θ_{2011} to θ_{2016}, formulated as follows:

$$H_0 : \theta_{2011} = \theta_{2012} = \theta_{2013} = \theta_{2014} = \theta_{2015} = \theta_{2016} = 0$$

This test aims to detect an anticipated dynamic or structural break before the reform. The Fisher test returns a p-value of 0.61, which prevents us from rejecting H_0. In other words, there is no statistical evidence of divergence between the groups before 2017.

These results, both visual and statistical, reinforce the credibility of our identification strategy: the hypothesis of parallel trends is here solidly validated[7].

3.2.4 Comparison with the level specification

To test the sensitivity of our results to the choice of functional form, we reproduced the previous analyses using the dependent variable $CrimeRate_{cgt}$ in absolute level rather than in logarithm.

[7] Appendix B

First, we can state that visually, the evolution of Swiss and foreign crime rates between 2010 and 2016 appears less parallel than in logarithm.

Moreover, the event-study of the coefficients θ_y for the years prior to 2017 in this level specification highlights a greater dispersion of effects. Some coefficients move away from zero, and their confidence intervals do not systematically contain the zero value, as we can see in the figure below.

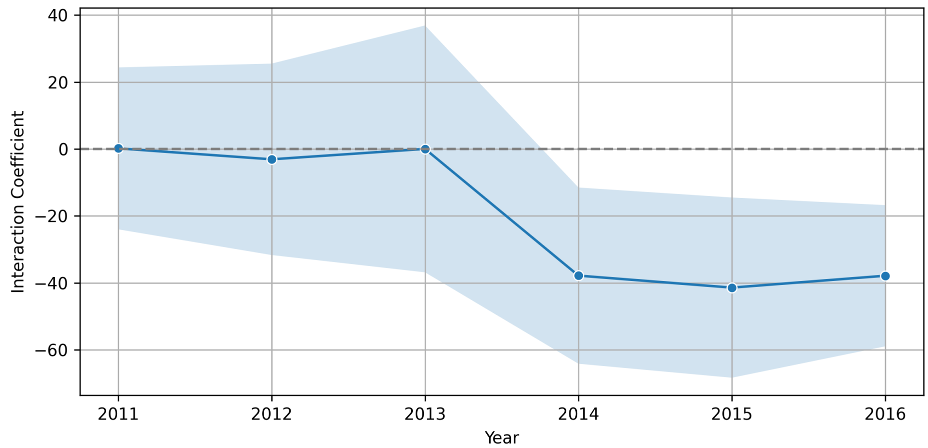

Figure 4: Event-Study Pre-Trends Dummies with 95% CI (lin-lin specification)

The joint nullity test of the coefficients θ_{2011} to θ_{2016} returns a p-value of less than 5%, suggesting a violation of the parallel trends hypothesis, which contrasts with the result obtained previously with the log model.

Both visually and statistically, the log specification therefore offers us better validity for our identification hypothesis, and so we decided to keep it for the remainder of the analysis[8].

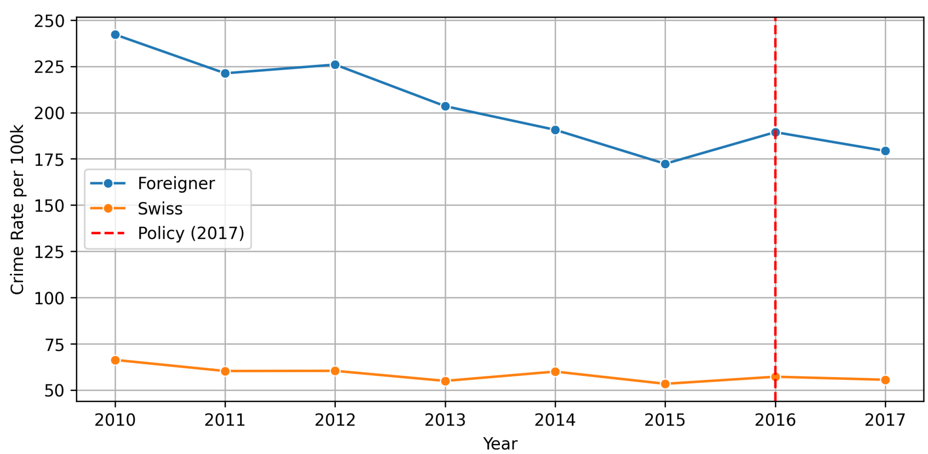

Figure 5: Pre-2017 Crime Rate by Group

[8] Appendix C

4 Results

4.1 Main estimates

4.1.1 Canton-Panel DiD

Using fixed effects by canton, year, group and clustering standard errors by canton, our baseline regression is:

$$\log\big(CrimeRate_{cgt}\big) = \beta\,(T_g \times P_t) + \alpha_c + \lambda_g + \gamma_t + \varepsilon_{cgt}$$

where T_g indicates foreign residents and P_t equals 1 for post-2016 years. Table 2 displays the results.

Table 2: OLS Regression Results: Canton-Panel DiD

Variable	coef	std err	z	P>z	[0.025	0.975]
Intercept	5.1324	0.094	54.402	0.000	4.947	5.317
$T_g \times P_t$	-0.0412	0.075	-0.553	0.581	-0.188	0.105

Statistic	**Value**	**Statistic**	**Value**
No. Observations	556	R-squared	0.782
Df Residuals	518	Adj. R-squared	0.766
Df Model	37	Covariance Type	cluster
Durbin–Watson	1.849	Jarque–Bera (JB)	1325.362
Skew	–1.296	Prob(JB)	1.59e–288
Kurtosis	10.105	Cond. No.	26.7

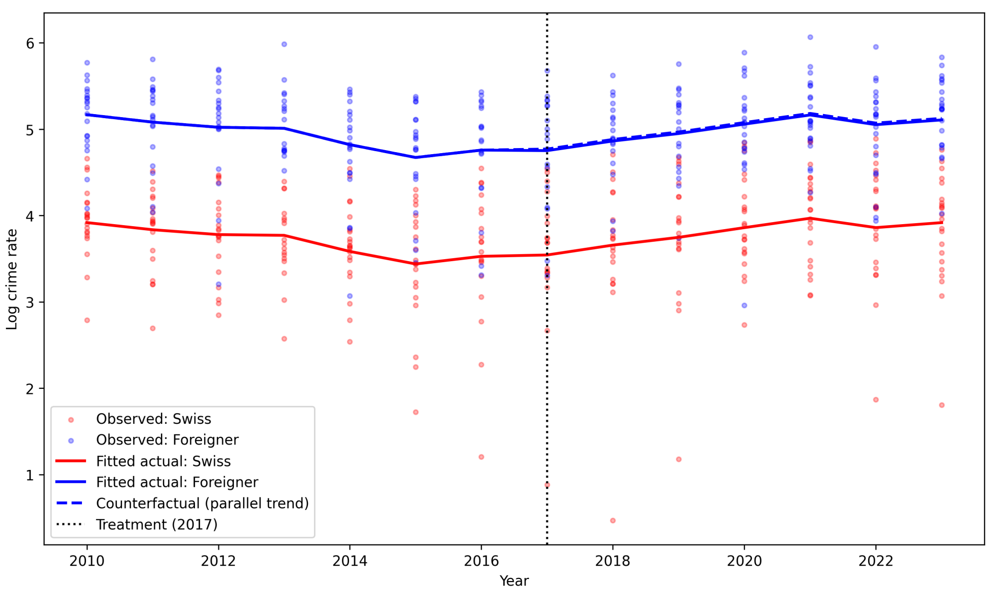

Figure 6: DiD with Group-Specific Trend: Actual vs. Parallel-Trend Counterfactual

The coefficient on $T_g \times P_t$ is -0.0412 (SE $= 0.075$, $p = 0.581$), corresponding to roughly a 4% decline in the foreign-resident crime rate after 2016, but this estimate is far from statistically significant.

In other words, once we control for canton, year, and group fixed effects (and cluster error terms within cantons), there is no detectable change in violent crime among foreign residents following the expulsion law. The model's R^2 of 0.782 indicates a strong overall fit.

Figure 6 above plots the observed and fitted log-crime rates for Swiss and foreign residents, along with the counterfactual series implied by a parallel trend; the vertical line marks the policy implementation in 2017. This visual helps illustrate the absence of any clear divergence between treated and control groups post-reform.

4.2 Threats to Validity and Mitigation

Below, we run a series of robustness checks, each targeting a different potential threat to ensure that our results hold under varying assumptions and specifications.

4.2.1 Parallelity hypothesis relaxation

A potential threat to validity is the possibility of differential secular trends between our groups. To address this, we extend our canton-panel DiD specification by allowing each group to follow its own linear time trend. Concretely, we estimate:

$$\log\bigl(CrimeRate_{cgt}\bigr) = \beta\left(T_g \times P_t\right) + \theta\left(T_g \times t\right) + \alpha_c + \lambda_g + \gamma_t + \varepsilon_{cgt}$$

where t is a linear time index (e.g., $t = 0, 1, \ldots$ for 2011–2023). The interaction $T_g \times t$ lets foreign and Swiss crime-rate trends differ linearly over time, thereby relaxing the strict parallel-trends assumption and capturing any secular drift.

Table 3: OLS Regression Results with relaxed parallelity hypothesis

Variable	coef	std err	z	P>|z|	[0.025	0.975]
Intercept	5.1372	0.091	56.301	0.000	4.958	5.316
$T_g \times P_t$	-0.0181	0.174	-0.104	0.917	-0.358	0.322
$T_g \times t$	-0.0032	0.018	-0.176	0.860	-0.039	0.033

Statistic	**Value**	**Statistic**	**Value**
No. Observations	556	R-squared	0.782
Df Residuals	517	Adj. R-squared	0.766
Df Model	38	Covariance Type	cluster
Durbin–Watson	1.849	Jarque–Bera (JB)	1320.292
Skew	-1.295	Prob(JB)	2.01e–287
Kurtosis	10.091	Cond. No.	123.

Estimating this specification on the full panel with canton-clustered standard errors yields $\hat{\theta} = -0.0032$ (SE = 0.018; $p = 0.860$) and $\hat{\beta} = -0.0181$ (SE = 0.174; $p = 0.917$). The coefficient θ is indistinguishable from zero, indicating no evidence of a differential linear trend between foreigners and Swiss before or after treatment. Meanwhile, β, which measures the DiD effect, remains essentially unchanged relative to the baseline model ($\hat{\beta} = -0.0412$, $p = 0.581$). Together, these results show that relaxing the parallel-trends assumption to allow for linear group-specific trends does not alter the conclusion: there is no significant policy effect, and the original DiD estimate is not biased by differential linear drifts.

4.2.2 Allowing for Time-Varying Treatment Effects (Event Study)

To relax the assumption of a single "jump" in treatment effect and allow for time-varying impacts, we estimate an event-study model. In this specification, we replace the fixed Post indicator with a series of leads and lags that capture the effect in each year relative to implementation. Concretely, we estimate:

$$\log(CrimeRate_{cgt}) = \sum_{\substack{k=-7 \\ k\neq -1}}^{6} \beta_k \cdot \mathbf{1}[t - 2017 = k] \times T_g + \alpha_c + \lambda_g + \gamma_t + \varepsilon_{cgt}$$

Where $\mathbf{1}\{\, t - 2017 = k \,\}$ is an indicator for being k years away from the policy year (with $k = -1$ omitted as the reference). Each coefficient β_k thus measures the differential change in log-crime rate for foreigners, k years before or after the reform, relative to Swiss nationals.

Table 4: Dynamic Treatment Effects (Event Study Specification)

Variable	**coef**	**std err**	**z**	**P>z**	**[0.025**	**0.975]**
Intercept	5.1147	0.206	24.818	0.000	4.711	5.519
lead_m7	0.0011	0.158	0.007	0.994	-0.309	0.311
lead_m6	0.0637	0.164	0.388	0.698	-0.258	0.386
lead_m5	0.0528	0.133	0.398	0.691	-0.207	0.313
lead_m4	0.1733	0.176	0.983	0.326	-0.172	0.519
lead_m3	-0.0841	0.179	-0.469	0.639	-0.435	0.267
lead_m2	0.0588	0.134	0.437	0.662	-0.205	0.322
lead_p0	-0.0094	0.171	-0.055	0.956	-0.344	0.325
lead_p1	0.0857	0.255	0.337	0.736	-0.413	0.585
lead_p2	0.0348	0.259	0.134	0.893	-0.472	0.542
lead_p3	-0.1162	0.142	-0.819	0.413	-0.394	0.162
lead_p4	-0.0470	0.199	-0.237	0.813	-0.436	0.342
lead_p5	-0.1257	0.237	-0.530	0.596	-0.590	0.339
lead_p6	0.1358	0.198	0.686	0.493	-0.252	0.524

Statistic	**Value**	**Statistic**	**Value**
No. Observations	556	R-squared	0.784
Df Residuals	506	Adj. R-squared	0.763
Df Model	49	Covariance Type	cluster
Durbin–Watson	1.825	Jarque–Bera (JB)	1246.023
Skew	–1.281	Prob(JB)	2.69e–271
Kurtosis	9.872	Cond. No.	36.9

In Table 4, the coefficients β_k for $k = -7, \ldots, -2$ (pre-treatment) are all small and statistically indistinguishable from zero, providing graphical evidence that no pre-existing divergence occurred between foreigners and Swiss before 2017. Similarly, the β_k for $k \geq 0$ (post-treatment) are also close to zero and not significant, indicating neither an immediate nor a delayed policy effect. Together, these results confirm that there was no anticipatory behavior leading up to the reform and no gradual impact after its implementation.

4.2.3 Canton heterogeneity

A key threat to our DiD inference is that cantons may follow different trends. This is due to local policing, demographic changes, enforcement priorities, which could bias the average treatment effect on our nation-wide DiD. However, the lack of data for certain cantons may cause an issue for our global inference.

A. Filtered-sample canton-panel DiD

In order to check the robustness of our nationwide regression, we limit attention to cantons with at least three pre-2017 and three post-2017 for a total of ≥ 6 observations per canton. We estimate:

$$\log\bigl(CrimeRate_{cgt}\bigr) = \beta\,\bigl(T_g \times P_t\bigr) + \alpha_c + \lambda_g + \gamma_t + \varepsilon_{cgt}\,.$$

Which is similar to our primary model with an extra filter.[9]

The estimated $\hat{\beta}$ is -0.0598 (SE $= 0.072$; $p = 0.408$)[10], indicating no significant average effect once we ensure that each canton contributes at least three years on each side of the cut. This supports the robustness of our baseline model[11], as both specifications lead to the conclusion that the effect is not statistically significant.

B. Canton-specific DiD

To examine heterogeneity in the treatment effect across cantons, potentially driven by institutional or demographic differences, we estimate a separate coefficient β_c for each canton. Formally, we estimate:

$$\log(CrimeRate_{cgt}) = \sum_c \beta_c\,(\mathbf{1}\{\text{canton} = c\} \times T_g \times P_t) + \alpha_c + \lambda_g + \gamma_t + \varepsilon_{cgt},$$

where $\mathbf{1}\{c\}$ is an indicator equal to 1 only when the observation belongs to canton c. Each β_c therefore captures the DiD estimate for "foreigners in canton c after 2017," allowing us to detect whether any single canton drives the overall null finding.

[9] section 4.1.1
[10] Appendix G
[11] section 4.1.1

Table 5: Estimated Treatment Effect by Canton

Variable	coef	std err	z	P>z	[0.025	0.975]
Intercept	5.1359	0.113	45.550	0.000	4.914	5.357
Aargau	-0.0872	0.190	-0.460	0.646	-0.460	0.280
Appenzell A. Rh.	0.2429	0.345	0.705	0.481	-0.434	0.920
Bern	0.2346	0.190	1.237	0.217	-0.138	0.600
Fribourg	-0.0792	0.190	-0.417	0.677	-0.452	0.290
Genève	-0.2644	0.190	-1.394	0.164	-0.637	0.100
Glarus	0.1266	0.222	0.571	0.569	-0.309	0.560
Grisons	0.1913	0.190	1.009	0.314	-0.181	0.560
Jura	-0.1469	0.192	-0.766	0.444	-0.524	0.230
Luzern	0.0012	0.190	0.006	0.995	-0.371	0.370
Neuchâtel	0.1228	0.190	0.647	0.518	-0.250	0.490
Nidwalden	-0.1636	0.292	-0.560	0.576	-0.738	0.410
Obwalden	0.5934	0.260	2.281	0.023	0.082	1.100
Schaffhausen	0.0088	0.192	0.046	0.964	-0.368	0.380
Schwyz	-0.1494	0.190	-0.788	0.431	-0.522	0.220
Solothurn	-0.1158	0.190	-0.610	0.542	-0.488	0.250
St. Gallen	-0.0553	0.190	-0.292	0.771	-0.428	0.310
Thurgau	-0.0237	0.190	-0.125	0.900	-0.396	0.340
Ticino	-0.1881	0.190	-0.992	0.322	-0.561	0.180
Uri	0.4278	0.364	1.175	0.241	-0.288	1.140
Valais	-0.0375	0.190	-0.198	0.843	-0.410	0.330
Vaud	-0.2291	0.190	-1.208	0.228	-0.602	0.140
Zug	-0.2760	0.190	-1.456	0.146	-0.649	0.090
Zürich	-0.2936	0.190	-1.548	0.122	-0.666	0.070

Statistic	**Value**	**Statistic**	**Value**
No. Observations	556	R-squared	0.791
Df Residuals	496	Adj. R-squared	0.766
Df Model	59	Covariance Type	nonrobust
Durbin–Watson	1.755	Jarque–Bera (JB)	1681.971
Skew	–1.397	Prob(JB)	0.00
Kurtosis	11.049	Cond. No.	31.2

In the full sample of 26 cantons, almost all $\hat{\beta}_c$ estimates are small and statistically insignificant, indicating that no single canton drives the overall result. Only Obwalden shows a significant deviation ($\hat{\beta}_{\text{Obwalden}} = 0.5934,\ p = 0.023$); since canton fixed effects only account for time-invariant differences, any local policy shift or demographic shock in Obwalden could be absorbed by $\hat{\beta}_c$. Moreover, Obwalden's small number of observations[12] makes this estimate sensitive to statistical noise, which further limits its causal interpretability. Several other cantons (e.g., Zug: $\hat{\beta} = -0.2760,\ p = 0.146$; Zürich: $\hat{\beta} = -0.2936,\ p = 0.122$) exhibit marginally large point estimates that nonetheless fail to reach conventional significance. Overall, the canton-specific estimates reinforce our finding that the pooled DiD estimate is not distorted by any individual region's behavior.

[12] Appendix H

C. Sample-composition (leave-one-out)

To assess the robustness of the estimated treatment effect to sample composition, we re-estimate the DiD model iteratively, each time excluding one canton from the sample. This allows us to verify whether the nationwide null result is driven by any single region. For each exclusion, we record the estimated coefficient on the treatment variable β, its standard error, and the associated p-value. The results appear in Table 6.

Table 6: Leave-One-Out DiD Estimates by Removed Canton

Removed Canton	**Coefficient** ($T_g \times P_t$)	**Std. Err.** ($T_g \times P_t$)	**p-value** ($T_g \times P_t$)
Aargau	-0.043799	0.078573	0.577234
Bern / Berne	-0.045348	0.078614	0.564041
Fribourg / Freiburg	-0.027739	0.077799	0.721435
Genève	-0.032194	0.078237	0.680708
Glarus	-0.075213	0.067208	0.263090
Graubünden / Grigioni / Grischun	-0.037274	0.078836	0.636348
Jura	-0.068641	0.071660	0.338130
Luzern	-0.031183	0.078467	0.691067
Neuchâtel	-0.048169	0.078058	0.537168
Nidwalden	-0.045498	0.075823	0.548472
Schaffhausen	-0.055669	0.076447	0.466487
Schwyz	-0.040294	0.078778	0.609010
Solothurn	-0.028233	0.077888	0.716990
St. Gallen	-0.036717	0.078703	0.640833
Thurgau	-0.037562	0.078636	0.632889
Ticino	-0.034846	0.078544	0.657292
Valais / Wallis	-0.036303	0.078596	0.644159
Vaud	-0.030708	0.078081	0.694111
Zug	0.002293	0.064629	0.971703
Zürich	-0.033141	0.078367	0.672366
Appenzell Ausserrhoden	-0.051052	0.074688	0.494269
Obwalden	-0.049475	0.075840	0.514165
Uri	-0.059817	0.072341	0.408304
None (Full Sample)	-0.041235	0.074627	0.580575

The leave-one-out estimates remain negative across all iterations (e.g., shifting from -0.0412 to -0.0752 when Glarus is excluded), but none reach statistical significance (all p-values > 0.26). This robustness check confirms that the null finding is not driven by any particular canton, reinforcing our conclusion that the expulsion initiative had no detectable effect on foreign-crime rates.

Overall, subsections (A), (B) and (C) indicate that by filtering cantons and estimating canton-specific effects, we both assess potential bias from differential time paths of cantons and verify that our nation-wide effect is not driven by outliers or lack of data.

4.2.4 Serial correlation

Year-to-year correlations in canton crime rates can induce serially correlated residuals. If we ignore this when estimating our DiD, OLS standard errors will be substantially underestimated, leading to misleading inference.

A. Non-clustered DiD (two-group model)

As a first pass, we estimated a "nationwide" DiD that included only group and year fixed effects (no canton clustering):

$$\log\left(CrimeRate_{gt}\right) = \beta\,(T_g \times P_t) + \lambda_g + \gamma_t + \varepsilon_{gt}.$$

Under this specification, the coefficient on $T \times P$ was -0.1461 (SE $= 0.0346$, $p = 0.001$), suggesting a significant negative effect.[13] However, two diagnostic statistics indicated that this model was misspecified:

- The Durbin–Watson statistic was 3.255, which flags negative autocorrelation in the residuals. Since residuals are ordered over time (one observation per group-year), the DW test remains interpretable because error is not clustered.
- The R^2 was extremely high (0.997), despite using only two groups and no canton fixed effects. This artificially tight fit often reflects omitted-variable bias when panel structure is ignored.

Because clustering standard errors was not feasible with only two clusters (Swiss vs. foreign), these diagnostics strongly suggested that serial correlation and omitted canton-level heterogeneity were biasing both the coefficient estimate and its standard error.

To address these concerns, we adopted our baseline specification [14]. The adjusted R^2 also drops (to 0.766), reflecting a more realistic fit once unobserved canton-level factors and serial correlation are properly accounted for. This change confirms that the earlier significance in the non-clustered model was driven by misspecified errors and unmodeled canton heterogeneity.

B. Wild cluster bootstrap

Although clustering at the canton level addresses serial correlation, 26 clusters still borders on the "few clusters" regime, where standard errors can remain unreliable. To verify our inference, we therefore implemented a wild cluster bootstrap with 1,000 replications, using independent Rademacher weights to reweight each canton's residuals, reconstruct the outcome, and re-estimate β in each draw. The wild-bootstrap p-value of 0.645 closely matches the original clustered-SE p-value, confirming that our DiD estimate remains indistinguishable from zero even under this more conservative inference procedure.[15]

[13] Appendix D
[14] section 4.1.1
[15] Appendix E

4.2.5 Functional form

To confirm that our DiD estimates aren't driven by how we transform the outcome[16], we re-ran the regression after applying the Box–Cox transform at its optimal power ($\lambda \simeq 0.25$)[17]. Namely, we replaced the raw crime rate by

$$CrimeRate_{cgt}^{(0.25)} = \frac{CrimeRate_{cgt}^{0.25} - 1}{0.25}.$$

The regression becomes the following:

$$CrimeRate_{cgt}^{(0.25)} = \beta \left(T_g \times P_t\right) + \alpha_c + \lambda_g + \gamma_t + \varepsilon_{cgt}\,.$$

Under the Box-Cox transformed *CrimeRate*, the treatment effect is estimated at $\hat{\beta}_{(\mathrm{BC})} = -0.104$ ($\mathrm{SE} = 0.186$, $p = 0.58$)[18]. By comparison, using a simple log-transformation yielded $\hat{\beta}_{(\log)} = -0.041$ ($\mathrm{SE} = 0.075$, $p = 0.58$).[19]

Both estimates are slightly negative, of similar magnitude in relative terms, and neither is statistically significant. Their confidence intervals overlap substantially, confirming that our DiD findings are not sensitive to the specific choice of transformation for the dependent variable.

4.2.6 Placebo Test

To further assess the validity of our identification strategy, we perform placebo tests by assigning the treatment to earlier years (2013, 2014, and 2015). For each placebo year Y, we redefine the treatment indicator P_t to activate from year Y onward and estimate the same difference-in-differences model used in the main analysis. The results are presented in Table 7.

Table 7: Placebo Tests: Estimated Coefficients for Pretreatment Years

Placebo Year	**Coefficient** ($T_g \times P_t$)	**Std. Err.** ($T_g \times P_t$)	**p-value** ($T_g \times P_t$)
2013	-0.032601	0.059227	0.582012
2014	-0.076770	0.052802	0.145969
2015	-0.034819	0.053329	0.513811

All estimated placebo coefficients are close to zero and statistically insignificant (all p-values > 0.14), suggesting no artificial break occurred before the actual reform.

[16] Appendix J
[17] section 3.1.4
[18] Appendix I
[19] section 4.1.1

5 Conclusion

The core empirical result of this study is straightforward: the Foreigner $\times$ Post coefficient is close to zero and far from statistical significance ($\hat{\beta} = -0.0412,\ p = 0.581$)[20].

Why did the promised deterrence not materialise? Combining institutional context with empirical results, we identify two likely explanations:

- Low perceived certainty of expulsion: the "rigour clause"[21] that allows judges to waive expulsion in certain cases, reduces the credibility of the threat.

- Marginal deterrence from harsher sanctions is limited when accounting for time and canton effects, consistent with evidence that factors such as legal work access (Bell et al., 2013 [2]) and welfare support (Auer et al., 2024 [1]) more effectively influence crime.

In short, the policy shifted legal discourse rather than offender behaviour. Alternative approaches, such as integration policies, early intervention programs, and targeted prevention efforts, may deliver better outcomes.

Our analysis relies on aggregated administrative data, which, while comprehensive at the cantonal level, presents several limitations. First, we lack individual-level information (e.g., age, gender, migratory background, criminal history), so we cannot explore heterogeneous effects within the foreign population. Second, small cantons like Obwalden have very low crime counts, making some estimates sensitive to statistical noise. Third, we assume uniform implementation of the reform from 2017 onward, even though enforcement may have varied across cantons. Without detailed data on legal proceedings or actual expulsion decisions, we cannot account for these differences. Finally, our crime data ends in 2023, so any effects emerging after that date remain unobserved.

Due to its decentralized system, central role of the popular vote, and margin of interpretation left to judges, Switzerland's institutional context is quite particular. The application of the reform may therefore have had a unique effect in Switzerland. These specific features limit the external validity of our results and make it difficult to compare our analytical framework with that of other countries. We must therefore remain cautious before drawing conclusions on the effect of similar measures in more centralized or uniform systems.

Future work could address additional robustness checks. For example, one could use crimes not subject to automatic expulsion (i.e., non-concerned crimes) as a secondary control group to verify that the reform had no spillover effects on unrelated offense categories. Monthly time series rather than annual aggregates would also allow for a more precise identification of timing and dynamic responses. Extending the analysis beyond 2023 would be essential for capturing delayed or longer-term effects that may emerge only after our current data window closes.

[20] section 4.1.1

[21] section 1

Acknowledgement

Our sincere thanks go to Rafael Lalive and Jeremy Zuchuat for their guidance and invaluable support throughout this project. We would also like to thank Hajer Badri, Tracy Frei, Louis Hubert and Rui Azevedo Leitao for their advice and availability throughout the project.

References

1. Auer, D., Slotwinski, M., Ahrens, A., Hangartner, D., Kurer, T., Kurt, S., & Stutzer, A. (2024). *Social assistance and refugee crime* (tech. rep. No. 11051). CESifo Working Paper. Retrieved May 27, 2025, from https://hdl.handle.net/10419/296140.
2. Bell, B., Fasani, F., & Machin, S. (2013). Crime and immigration: Evidence from large immigrant waves. *Review of Economics and Statistics*, *95*(4), 1278–1290. DOI: 10.1162/REST_a_00337.
3. Federal Chancellery. (2010). *Federal initiative 'for the expulsion of foreign criminals'* [Voted on 28 November 2010]. Retrieved May 27, 2025, from https://www.bk.admin.ch/ch/f/pore/vi/vis357t.html.
4. Federal Office of Justice (OFJ). (2016). *Implementing law introducing articles 66a and 66abis of the criminal code* [Entered into force on 1 October 2016]. Retrieved May 27, 2025, from https://www.bj.admin.ch/dam/bj/fr/data/sicherheit/gesetzgebung/archiv/ausschaffung/landesverweisung/vorentw-vo-f.pdf.
5. Federal Statistical Office (FSO). (2015). Police crime statistics (spc) – annual report 2014 [See Chapter 5.3.1 "Importation et sauvegarde"]. Retrieved May 27, 2025, from https://www.bfs.admin.ch/bfsstatic/dam/assets/2360912/master.
6. Federal Statistical Office (FSO). (2023a). *Cantonal unemployment rate (2023)* [Labour Market Statistics]. https://www.pxweb.bfs.admin.ch/pxweb/fr/px-x-0602010000_101/-/px-x-0602010000_101.px/.
7. Federal Statistical Office (FSO). (2023b). *Gdp per capita by canton* [Cantonal Economic Accounts]. https://www.bfs.admin.ch/bfs/fr/home/statistiques/economie-nationale/comptes-nationaux/produit-interieur-brut-canton.html.
8. Federal Statistical Office (FSO). (2023c). *Share of foreign residents by canton (2023)* [STAT-TAB platform]. https://www.pxweb.bfs.admin.ch/pxweb/fr/px-x-0102020000_202/-/px-x-0102020000_202.px/.
9. Federal Statistical Office (FSO). (2025). *Police crime statistics (spc)* [STAT-TAB platform]. Retrieved May 27, 2025, from https://www.pxweb.bfs.admin.ch/pxweb/en/px-x-1903020100_101/-/px-x-1903020100_101.px/.
10. Pinotti, P. (2017). Clicking on heaven's door: The effect of immigrant legalization on crime. *American Economic Review*, *107*(1), 138–168. DOI: 10.1257/aer.20150355.
11. Swiss Confederation. (2010). Proposed amendment to article 121 of the federal constitution [Constitutional proposal, no official document link provided].

12. Swiss Federal Chancellery. (2010). *2010 federal vote on the expulsion initiative* [Federal vote data].
https://www.bk.admin.ch/ch/f/pore/va/20101128/can552.html.
13. Swiss Federal Chancellery. (2014). *Popular initiative "against mass immigration"* [Voted on February 9, 2014].
https://fr.wikipedia.org/wiki/Initiative_populaire_%C2%AB_contre_l%27immigration_de_masse_%C2%BB.

Appendices

Appendix A - Parallel Trends, Group x Year Linear Interaction Test

Table 8: Slope Interaction Test for Parallel Trends

Variable	coef	std err	t	P>t	[0.025	0.975]
Intercept	53.9118	20.061	2.687	0.023	9.213	98.610
T	52.5384	28.370	1.853	0.094	-10.655	115.772
year	-0.0248	0.010	-2.484	0.032	-0.047	-0.003
T:year	-0.0255	0.014	-1.809	0.101	-0.057	0.006

Statistic	Value	Statistic	Value
No. Observations	14	R-squared	0.995
Df Residuals	10	Adj. R-squared	0.994
Df Model	3	Covariance Type	nonrobust
Durbin–Watson	1.889	Jarque–Bera (JB)	0.848
Skew	-0.103	Prob(JB)	0.654
Kurtosis	1.812	Cond. No.	5.30e+06

Appendix B - Parallel Trends, Event-Study Pre-Trends Test

Table 9: OLS Regression Results – Slope Interaction Test for Foreigners

Variable	coef	std err	z	P>z	[0.025	0.975]
Intercept	3.9209	0.074	53.283	0.000	3.777	4.065
FG	1.2064	0.058	20.694	0.000	1.092	1.321
2011	0.0626	0.072	0.874	0.382	-0.078	0.203
2012	0.0516	0.103	0.502	0.616	-0.150	0.253
2013	0.1722	0.101	1.708	0.088	-0.025	0.370
2914	-0.0852	0.131	-0.653	0.514	-0.341	0.171
2015	0.0577	0.135	0.427	0.669	-0.207	0.322
2016	-0.0011	0.162	-0.007	0.995	-0.318	0.316

Statistic	Value	Statistic	Value
No. Observations	272	R-squared	0.794
Df Residuals	236	Adj. R-squared	0.764
Df Model	35	Covariance Type	cluster
Durbin–Watson	1.607	Jarque–Bera (JB)	91.556
Skew	–0.586	Prob(JB)	1.32e–20
Kurtosis	5.590	Cond. No.	27.0

Joint Test for All Foreigners $\times$ Year Interactions = 0

F-statistic: 0.32

p-value: 0.5748

We have mixed evidence. To be fully credible, we must adjust our DiD for any residual trend differences.

Appendix C - Level Model

Table 10: Canton-panel DiD Regression Results

Variable	coef	std err	z	P>z	[0.025	0.975]
Intercept	148.7972	9.538	15.601	0.000	130.103	167.491
T_P	5.4134	6.879	0.787	0.431	-8.069	18.896

Statistic	Value	Statistic	Value
No. Observations	672	R-squared	0.720
Df Residuals	633	Adj. R-squared	0.703
Df Model	38	Covariance Type	cluster
Durbin–Watson	2.156	Jarque–Bera (JB)	588.782
Skew	0.928	Prob(JB)	1.40e–128
Kurtosis	7.193	Cond. No.	29.5

Table 11: Event-Study Pre-Trends Test

Variable	coef	std err	z	P>z	[0.025	0.975]
Intercept	44.6474	4.415	10.113	0.000	35.994	53.300
FG	129.0285	11.856	10.883	0.000	105.792	152.265
2011	0.1846	12.337	0.015	0.988	-23.995	24.364
2012	-3.0857	14.586	-0.212	0.832	-31.673	25.502
2013	0.0252	18.810	0.001	0.999	-36.841	36.891
2014	-37.8304	13.426	-2.818	0.005	-64.145	-11.516
2015	-41.4103	13.731	-3.016	0.003	-68.323	-14.498
2016	-37.8626	10.758	-3.519	0.000	-58.949	-16.776

Statistic	Value	Statistic	Value
No. Observations	272	R-squared	0.765
Df Residuals	236	Adj. R-squared	0.730
Df Model	35	Covariance Type	cluster
Durbin–Watson	1.902	Jarque–Bera (JB)	444.404
Skew	0.855	Prob(JB)	3.15e–97
Kurtosis	9.024	Cond. No.	27.0

Joint Test for All Foreigner$\times$Year Interactions = 0
F-statistic: 5.01
p-value: 0.0356
We have clear evidence of non parallel trends. A log-specification may be more appropriate.

Appendix D - Nationwide DiD with Group and Year Fixed Effects (Clustered)

Table 12: Nationwide Difference-in-Differences without Canton Fixed Effects

Variable	coef	std err	z	P>z	[0.025	0.975]
Intercept	5.0307	0.020	253.615	0.000	4.987	5.074
T:P	-0.1461	0.035	-4.208	0.001	-0.222	-0.070

Statistic	Value	Statistic	Value
No. Observations	28	R-squared	0.997
Df Residuals	12	Adj. R-squared	0.994
Df Model	15	Covariance Type	nonrobust
Durbin–Watson	3.255	Jarque–Bera (JB)	2.868
Skew	0.000	Prob(JB)	0.238
Kurtosis	1.432	Cond. No.	2.12e+16

Appendix E - Wild cluster bootstrap

Table 13: Wild Cluster Bootstrap Results

	Original Coef	Bootstrap p-value
β	-0.041235	0.645355

Appendix F - Descriptive Population Statistics by Canton

Table 14: Key demographic, political and economic indicators by canton

Canton	Abbr.	% Foreigners[8]	UDC Vote (%)[12]	Unemployment (%)[6]	GDP/capita (CHF)[7]
Aargau	AG	25.7	46.1	2.4	64,500
Appenzell Rh.-Ext.	AR	17.2	44.2	1.6	58,500
Appenzell Rh.-Int.	AI	10.9	59.2	1.3	48,000
Basel-Landschaft	BL	22.6	35.6	2.8	72,000
Basel-Stadt	BS	38.4	27.0	3.8	155,000
Bern	BE	17.2	42.9	2.7	67,500
Fribourg	FR	22.6	35.2	2.6	55,500
Geneva	GE	38.7	27.9	5.7	75,000
Glarus	GL	23.0	43.0	2.1	58,500
Graubünden	GR	19.4	43.1	1.6	63,500
Jura	JU	14.8	39.5	3.6	52,000
Lucerne	LU	20.2	41.7	2.0	63,500
Neuchâtel	NE	30.4	31.2	4.9	58,000
Nidwalden	NW	18.3	50.3	1.4	61,000
Obwalden	OW	17.4	50.5	1.4	60,000
Schaffhausen	SH	25.3	47.3	2.3	60,500
Schwyz	SZ	19.8	57.8	1.8	64,000
Solothurn	SO	23.3	40.6	2.8	57,000
St. Gallen	SG	24.5	44.3	2.3	61,000
Thurgau	TG	20.6	48.3	2.1	59,000
Ticino	TI	27.5	41.0	3.9	56,000
Uri	UR	12.5	53.1	1.5	50,500
Valais	VS	22.4	43.7	3.2	54,000
Vaud	VD	34.2	29.0	4.8	65,000
Zug	ZG	26.9	40.5	2.1	120,000
Zurich	ZH	27.1	36.5	3.4	79,000

Appendix G - Filtered-sample canton-panel

Table 15: Canton-panel DiD Regression Results

Variable	coef	std err	z	P>z	[0.025	0.975]
Intercept	5.1390	0.094	54.500	0.000	4.954	5.324
T_P	-0.0598	0.072	-0.827	0.408	-0.202	0.082

Statistic	Value	Statistic	Value
No. Observations	550	R-squared	0.784
Df Residuals	513	Adj. R-squared	0.769
Df Model	36	Covariance Type	cluster
Durbin–Watson	1.823	Jarque–Bera (JB)	1394.610
Skew	-1.306	Prob(JB)	1.46e–303
Kurtosis	10.351	Cond. No.	25.9

Appendix H - Obwalden data

Table 16: Obwalden crminilaty per group (2010–2023)

Year	Group	Crime Count	Population	Crime rate per 100k
2010	Foreigner	0	5162	0
2010	Swiss	0	30821	0
2011	Foreigner	0	5354	0
2011	Swiss	0	30969	0
2012	Foreigner	0	5407	0
2012	Swiss	0	31131	0
2013	Foreigner	0	5499	0
2013	Swiss	0	31414	0
2014	Foreigner	7	5599	125.0223
2014	Swiss	4	31609	12.6546
2015	Foreigner	8	5774	138.5521
2015	Swiss	3	31716	9.4589
2016	Foreigner	0	5772	0
2016	Swiss	0	31892	0
2017	Foreigner	0	5827	0
2017	Swiss	0	32056	0
2018	Foreigner	3	5870	51.1073
2018	Swiss	8	32296	24.7709
2019	Foreigner	0	5908	0
2019	Swiss	0	32361	0
2020	Foreigner	9	5872	153.2698
2020	Swiss	5	32475	15.3965
2021	Foreigner	8	6188	129.2825
2021	Swiss	9	32623	27.5879
2022	Foreigner	0	6868	0
2022	Swiss	0	32652	0
2023	Foreigner	22	7071	311.13
2023	Swiss	2	32835	6.0911

Appendix I - Canton Panel DiD - Box-Cox Optimal Transformation

Table 17: Canton-panel DiD Regression Results (BC)

Variable	coef	std err	z	P>z	[0.025	0.975]
Intercept	10.3053	0.257	40.085	0.000	9.801	10.809
T_P	-0.1036	0.186	-0.557	0.577	-0.468	0.261

Statistic	Value	Statistic	Value
No. Observations	556	R-squared	0.817
Df Residuals	518	Adj. R-squared	0.804
Df Model	37	Covariance Type	cluster
Durbin–Watson	1.804	Jarque–Bera (JB)	171.938
Skew	-0.177	Prob(JB)	4.62e–88
Kurtosis	5.701	Cond. No.	26.7

Appendix J - Specification plot

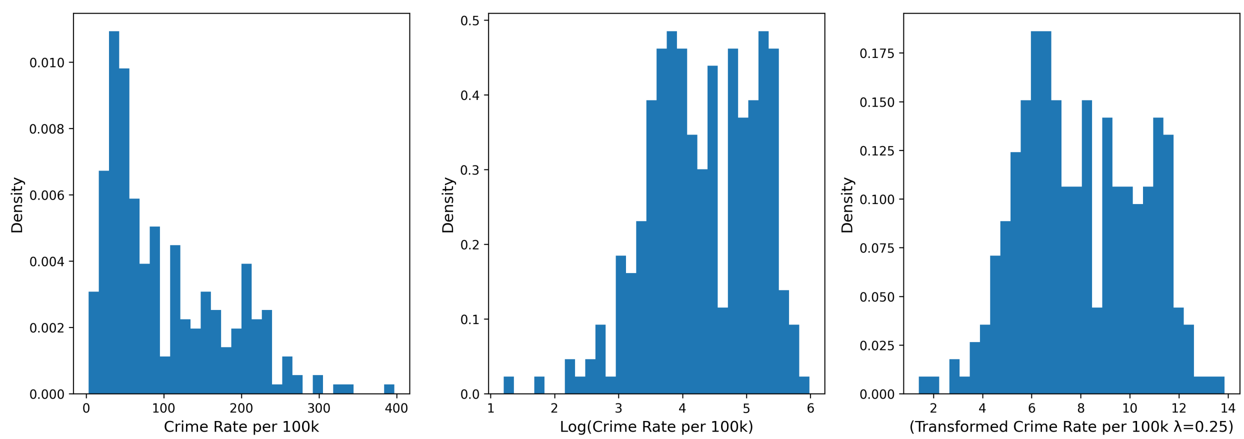

Figure 7: Distribution of crime rates under alternative functional forms