Estimating the Impact of Norway's Public Smoking Ban on **Smoking Prevalence**: A Difference-in-Differences Approach with Spain as Control

Amadeus Linge, Ulrik Haugland
University of Bergen

October 30, 2025

1 Introduction

This project investigates whether the introduction of Norway's public smoking ban led to a reduction in **regular smoking prevalence** among the population. Using an econometric Difference-in-Differences (DiD) framework, we aim to identify the causal effect of the smoking ban (implemented in **June 2004**) on smoking behavior by comparing trends in Norway (Treatment Group) and Spain (Control Group) before and after the policy.

2 Research Question

Did the 2004 smoking place ban in Norway cause a reduction in the country's regular smokers' percentage relative to Spain?

This question shifts the focus from health outcomes to the **behavioral mechanism** by which such a policy operates.

3 Data Description

The analysis uses annual data from Norway (Treated) and Spain (Control) covering the period from **1978 to 2020**, with a focus on the pre- and post-ban periods around 2004. The data processing steps were as follows:

- **Data Cleaning and Merging:** Raw data on smoking prevalence for Norway and Spain were loaded and cleaned. For Norway, data was read from Excel and

columns were renamed and filtered. For Spain, data was provided as a dictionary and converted to a DataFrame.

- **Interpolation:** Both countries had missing years in their annual smoking prevalence series. We created a complete range of years and used linear interpolation to fill in missing values, ensuring a balanced panel for analysis. Tobacco price indices for both countries were also processed from monthly data, extracting January values for each year and interpolating as needed.

- **Panel Construction:** The data was reshaped to long format, resulting in one row per country-year with columns for year, country, smoking prevalence, and tobacco price. Additional variables were created: a Norway dummy, a post-2004 dummy, and a treatment interaction term for DiD analysis.

Example of processed/interpolated data:

```
   Year  Country  smokers_percentage  tobacco_price  Norway  Post2004  Treated
0  1996   Norway               34.50          30.90       1         0        0
1  1996    Spain               36.25          27.65       0         0        0
2  1997   Norway               34.00          34.80       1         0        0
3  1997    Spain               36.00          30.57       0         0        0
... (data continues for all years and both countries) ...
 extbf{With clustered standard errors:}
\begin{verbatim}
```

```
==============================================================================
                   coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept       47.2248      3.236     14.593      0.000      40.882      53.567
Norway          -1.9793      0.530     -3.734      0.000      -3.018      -0.940
Post2004        -0.3014      1.280     -0.235      0.814      -2.811       2.208
Treated          0.0450      0.655      0.069      0.945      -1.239       1.329
tobacco_price   -0.3398      0.089     -3.836      0.000      -0.513      -0.166
==============================================================================
Notes:
[1] Standard Errors are robust to cluster correlation (cluster)
```

extbfDiD Estimate (Treated): 0.0450, p-value: 0.9452
extbfWith robust (HC1) standard errors:

```
==============================================================================
                   coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept       47.2248      1.433     32.949      0.000      44.416      50.034
```

```
Norway             -1.9793       0.319      -6.208       0.000      -2.604      -1.354
Post2004           -0.3014       0.437      -0.690       0.490      -1.158       0.555
Treated             0.0450       0.705       0.064       0.949      -1.336       1.426
tobacco_price      -0.3398       0.034      -9.935       0.000      -0.407      -0.273
==============================================================================
Notes:
[1] Standard Errors are heteroscedasticity robust (HC1)
\item \textbf{Tobacco Price Index} (control variable, $\text{TobaccoPrice}_t$): The H
 extbf{DiD Estimate (Treated):} $0.0450$, p-value: $0.9490$
```

extbfInterpretation: In both specifications, the coefficient for `Treated` (the DiD estimator δ) is very close to zero and not statistically significant, indicating that the smoking ban in Norway was not associated with a statistically significant reduction in smoking prevalence relative to Spain. The results are robust to both clustered and robust (HC1) standard errors.

3.1   Conceptual Framework

The DiD strategy compares the change in smoking prevalence in Norway before and after 2004 with the corresponding change in Spain over the same period. This calculation removes the effects of general European trends and pre-existing country differences.

4   Econometric Specification

The main estimating equation is:

$$Y_{it} = \alpha + \beta_1 \text{Post}_t + \beta_2 \text{Norway}_i + \delta(\text{Post}_t \times \text{Norway}_i) + \gamma_1 \text{TobaccoPrice}_{it} + \varepsilon_{it} \qquad (1)$$

where:

- Y_{it} is the annual smokers' percentage in country i at time t.

- Post_t is the dummy variable equal to 1 for years ≥ 2004 (treatment period).

- Norway_i is the country dummy variable (1 for Norway, 0 for Spain).

- δ is the DiD estimator of interest, capturing the causal impact of the smoking ban.

- TobaccoPrice_{it} is the country-year specific control variable for tobacco price.

- ε_{it} is the error term.

The coefficient δ measures the average treatment effect of the smoking ban on the smokers' percentage in Norway relative to Spain, conditional on parallel trends and the control for price effects.

4.1 Difference-in-Differences Regression Results

We estimated the DiD model using OLS with robust (HC1) standard errors due to the small number of clusters (countries). The main regression included controls for tobacco price. The key results are summarized below:

```
================================================================================
                 coef     std err        t       P>|t|      [0.025     0.975]
--------------------------------------------------------------------------------
Intercept      44.0192     2.964     14.853      0.000      37.995     50.043
Norway         -6.0192     4.192     -1.436      0.163     -14.669      2.631
Post2004       -7.0192     4.192     -1.675      0.108     -15.669      1.631
Treated        -2.4808     5.927     -0.419      0.680     -14.667      9.705
tobacco_price   0.0192     0.192      0.100      0.922      -0.401      0.440
================================================================================
```

extbfInterpretation: The coefficient for `Treated` (the DiD estimator δ) is negative and not statistically significant at the 5% level, indicating that the smoking ban in Norway was associated with a reduction in smoking prevalence relative to Spain, but the effect is not statistically distinguishable from zero given the data and limitations.

extbfNote: Standard errors are robust (HC1). With only two clusters, clustered standard errors are unreliable and not reported.

5 Robustness and Limitations

5.1 Parallel Trends Assumption and Visual Check

The core assumption of the DiD design is that, in the absence of the ban, Norway and Spain would have followed similar trends in smoking prevalence. We conducted a thorough check using both visual inspection and a formal regression-based test:

- **Visual Inspection:** We plotted the pre-treatment trends in smoking prevalence for both countries. The lines appeared to move in parallel, suggesting the assumption may be reasonable.

- **Statistical Test:** We regressed smoking prevalence on year, country, and their interaction using only pre-treatment data. The interaction term tests for a difference in trends. The estimated coefficient for the interaction (Norway × Year) was small and statistically insignificant (p-value = 0.671), supporting the parallel trends assumption.

extbfRegression output (pre-trend test):

```
================================================================================
                coef      std err         t       P>|t|      [0.025      0.975]
--------------------------------------------------------------------------------
Intercept       37.6000    0.676      55.581     0.000      35.945      39.255
Year_ct         -0.2882    0.084      -3.433     0.014      -0.494      -0.083
Norway          -2.8000    0.957      -2.927     0.026      -5.141      -0.459
Year_ct:Norway  -0.0530    0.119      -0.446     0.671      -0.343       0.237
================================================================================
```

The interaction term's p-value (0.671) indicates no significant difference in pre-treatment trends.

5.2 Inference with Few Clusters

Since the treatment (the smoking ban) is applied at the country level, the standard errors should ideally be **clustered by country** to account for serial correlation within countries. However, the sample contains only two clusters (Norway and Spain), which renders the asymptotic properties of the clustered standard errors unreliable. The main regression results will be reported, but the p-values associated with the DiD coefficient (δ) must be interpreted with extreme caution due to the limited number of clusters.

6 Expected Results

We expect $\delta < 0$, indicating a **faster decline** in the regular smokers' percentage in Norway following the introduction of the public smoking ban compared to the change observed in Spain.

7 Potential Limitations

- **Low Cluster Count:** The primary econometric limitation is the small number of clusters (countries) preventing reliable inference with clustered standard errors.

- **Interpolated Data:** Reliance on interpolated data for smoking prevalence and tobacco prices may introduce measurement error and smooth out critical fluctuations.

- **Spain's Policy Context:** Spain introduced its first moderate smoke-free laws in 2006, which could contaminate the post-treatment period comparison.

8   Conclusion

This study uses a Difference-in-Differences design to estimate the effect of Norway's 2004 smoking ban on smoking prevalence, using Spain as a control and controlling for tobacco price changes. The analysis provides quantitative evidence on the policy's behavioral impact, though formal inference is limited by the sample size.