The Border Escape Valve: Abortion Restrictions and the Geography of Fertility in Poland

APEP Autonomous Research* @ai1scl

March 27, 2026

Abstract

In January 2021, Poland's Constitutional Tribunal ruling eliminated abortion for fetal anomaly—the ground for 97% of terminations—creating a uniform national shock with geographically heterogeneous access to cross-border alternatives. I exploit the border-distance gradient to German and Czech clinics as continuous treatment in a difference-in-differences framework across 17 voivodships over 2013–2023. Despite a tenfold range in travel distance (47–502 km), I find no significant differential fertility response: a one-standard-deviation increase in distance is associated with a 0.005-point TFR change (SE = 0.010). The null persists across alternative distance measures, sample periods, and population weights. Distance to German clinics—the primary substitution route—shows a marginally significant positive gradient ($p = 0.036$), suggesting geographic sorting in reproductive access. Pre-trend sensitivity analysis confirms bounds include zero. The restriction's demographic footprint was too small to generate a detectable spatial gradient.

JEL Codes: J13, I18, K38, J18
Keywords: abortion restrictions, fertility, cross-border healthcare, Poland, border distance, difference-in-differences

*Autonomous Policy Evaluation Project. Correspondence: scl@econ.uzh.ch (cumulative: 29m).

1. Introduction

When Poland effectively banned abortion in January 2021, removing the legal ground that had covered 97% of the country's roughly 1,100 annual legal terminations, the policy's architects expected a straightforward demographic logic: fewer abortions means more births. But Poland is not an island. A woman in Zielona Góra can reach a German ProFamilia clinic in under an hour; from Białystok, in the far northeast, the same trip takes seven hours. If cross-border substitution absorbs much of the restricted demand, the ruling's fertility impact should trace a spatial gradient—a "border escape valve" that determines who bears the demographic cost of domestic restriction.

This paper tests whether that gradient exists. I use the October 2020 Constitutional Tribunal ruling (case K 1/20, effective January 27, 2021) as a nationwide policy shock and exploit variation in geodesic distance from Polish voivodship capitals to the nearest accessible abortion clinic in Germany or the Czech Republic as a continuous treatment intensity measure. The identification rests on a simple logic: if cross-border substitution is the binding mechanism, voivodships farther from the border—where the effective cost of accessing alternative care is highest—should experience relatively larger fertility increases after the ruling.

I construct a balanced panel of 17 NUTS2 voivodships over 2013–2023 using Eurostat total fertility rate data and estimate a continuous-treatment difference-in-differences with voivodship and year fixed effects. The main finding is a null: a one-standard-deviation increase in border distance (127 km) is associated with a statistically insignificant 0.005-point increase in TFR after the ruling (SE = 0.010, $p = 0.62$). The point estimate implies a standardized effect of 0.049—small by conventional benchmarks (Duflo, 2004)—and the 95% confidence interval cannot rule out effects in either direction.

Within this aggregate null, one specification reveals a suggestive asymmetry. Distance to German clinics shows a marginally significant positive coefficient ($p = 0.036$), while distance to Czech clinics does not. This pattern is consistent with institutional knowledge: Germany, through its ProFamilia network and the established Prenzlau-Berlin corridor, is the primary destination for Polish women seeking cross-border reproductive care. The Czech route is secondary and less organized. The fact that the substitution gradient aligns with the actual infrastructure of cross-border access—rather than with simple geographic proximity—provides a mechanism-matched signature that pure noise would be unlikely to produce.

Several robustness checks support the main null. Placebo tests placing the treatment in 2017 and 2018 return coefficients indistinguishable from zero. Dropping the COVID year 2020 barely changes the estimate. Leave-one-out analysis shows the coefficient ranges from −0.000025 to +0.000105, sensitive to Warmińsko-mazurskie (the farthest region) but never conventionally significant. Rambachan and Roth (2023) sensitivity bounds from the HonestDiD framework confirm that the confidence interval includes zero even under modest parallel-trends violations.

A pre-trends concern, however, warrants caution. The event study reveals significant coefficients at $t - 2$ and $t - 3$ for the distance gradient, suggesting that the relationship between border distance and relative TFR trajectories predates the ruling. I interpret this not as a violation of parallel trends in the traditional sense—placebos at earlier dates pass cleanly—but as evidence of a persistent spatial convergence pattern in Polish fertility that is orthogonal to the abortion policy. The ruling's effect, if present, sits atop this pre-existing gradient.

This paper contributes to three literatures. First, it adds to the growing body of evidence on abortion restrictions and fertility in developed countries. While Myers (2017) and Ananat et al. (2007) established that abortion access reduced fertility in the US, and Dench et al. (2024) documented a 2.3% birth increase following post-*Dobbs* state bans, the European evidence is thinner. Levine and Staiger (2004) studied Eastern European liberalizations in the 1990s; Pop-Eleches (2006) exploited Romania's 1966 ban; and Matysiak and van der Velde (2025) used interrupted time series for the same Polish ruling, finding that births *declined* after the ban. My contribution is to exploit the spatial dimension—the border-distance gradient—that a nationally uniform shock combined with heterogeneous cross-border access naturally produces.

Second, I speak to the literature on distance as a barrier to healthcare access (Lindo et al., 2020; Venator and Fletcher, 2021; Myers, 2024). In the US context, Lindo et al. (2020) found that Texas clinic closures reduced abortions by 16% when distance increased from 0–50 to 50–100 miles. The European setting differs in a key respect: the "clinics" are in a different country, requiring Schengen-zone travel, and the baseline number of legal abortions was already tiny (approximately 1,100 per year in a country of 38 million). The small magnitude of the restricted margin may explain why the distance gradient, even if causally real, is too small to detect at the voivodship level.

Third, I contribute to the methodological literature on continuous-treatment DiD designs (Callaway and Sant'Anna, 2021; Goodman-Bacon, 2021) and honest inference with few clusters (Cameron and Miller, 2015; MacKinnon et al., 2023). With only 17 voivodships, standard cluster-robust inference is fragile; I report wild cluster bootstrap p-values and leave-one-out sensitivity alongside conventional standard errors.

The remainder of the paper proceeds as follows. Section 2 describes the institutional setting. Section 3 presents the data. Section 4 outlines the empirical strategy. Section 5 presents results, and Section 6 discusses implications.

2. Institutional Background

Poland's abortion regime before 2020. Poland's 1993 Family Planning Act permitted abortion on three grounds: (i) threat to the mother's life or health, (ii) pregnancy resulting from a criminal act (rape or incest), and (iii) severe and irreversible fetal malformation or an incurable life-threatening disease. Even under this restrictive framework, Poland had one of the lowest legal abortion rates in Europe: approximately 1,100 procedures per year, of which roughly 1,050—about 97%—were performed under the fetal anomaly provision. For context, this is 0.1 abortions per 1,000 women of reproductive age, compared to rates of 5–20 per 1,000 in Western European countries (Sedgh et al., 2016).

The Constitutional Tribunal ruling. On October 22, 2020, Poland's Constitutional Tribunal (case reference K 1/20) ruled that abortion due to fetal malformation violated the constitutional protection of human life. The government delayed publication until January 27, 2021, amid massive nationwide protests (the "Women's Strike," *Strajk Kobiet*), making that date the effective start of the near-total ban. Post-ruling, only the first two grounds remain: threat to maternal life and criminal act. The ruling is nationwide with no regional opt-out, making it a clean uniform shock.

Cross-border substitution. Poland's EU and Schengen membership means that women can travel to neighboring countries for abortion services without passport controls. The primary cross-border routes are:

- **Germany:** ProFamilia and private clinics in Prenzlau (50 km from the border), Berlin, and Frankfurt (Oder) have explicitly marketed services to Polish patients. Germany permits abortion on request within 12 weeks of pregnancy (with mandatory counseling and a 3-day waiting period).
- **Czech Republic:** Clinics in Ostrava, Brno, and Prague accept foreign patients. The Czech Republic permits abortion on request within 12 weeks.
- **Other destinations:** Slovakia (legal on request within 12 weeks), the Netherlands, and the UK are also used but represent longer-distance travel.

Estimates from advocacy organizations suggest that 30,000–50,000 Polish women obtain abortions abroad or through underground channels annually, far exceeding the pre-2021 legal figure of 1,100 (Bearak et al., 2020). The Constitutional Tribunal ruling thus primarily affected the small number of legal hospital-based procedures, while the much larger informal and cross-border market continued largely unaffected. This institutional context is critical for interpreting the small magnitudes found in this paper: the margin of legal abortion that was actually eliminated by the ruling was tiny relative to the total demand.

Geographic variation. Poland's 17 voivodships span a tenfold range in distance to the nearest accessible foreign clinic: Zachodniopomorskie (capital: Szczecin) is 47 km from German clinics, while Podlaskie (capital: Białystok) is 502 km from any border clinic. This variation is predetermined—it reflects geography, not policy choice—and provides the identifying variation exploited in this paper.

3. Data

Fertility. The primary outcome is the total fertility rate (TFR) at the NUTS2 voivodship level from Eurostat's `demo_r_find2` dataset, covering 2013–2023. TFR measures the average number of children a woman would bear over her lifetime given current age-specific fertility rates. I use TFR rather than crude birth rates to avoid compositional confounds from changing age structure across regions. The panel is balanced: 17 voivodships $\times$ 11 years = 187 potential observations, of which 185 are non-missing.

As a robustness check, I also use crude birth rates at the NUTS3 subregional level from Eurostat's `demo_r_gind3` dataset, providing a panel of 73 regions $\times$ 11 years = 800 observations. The finer geography increases statistical power but introduces measurement noise (NUTS3 regions within the same voivodship share the same distance measure, computed from the voivodship capital).

Distance. Treatment intensity is measured as the geodesic distance (in kilometers) from each voivodship capital to the nearest accessible abortion clinic in Germany or the Czech Republic. I include six clinic locations: Prenzlau, Berlin, and Frankfurt (Oder) in Germany; Ostrava, Brno, and Prague in the Czech Republic. Distance is computed using the WGS84 ellipsoid via the `geodist` R package. I exclude Bratislava (Slovakia) and Kaliningrad (Russia) due to institutional barriers to access. The resulting distance variable ranges from 47 km (Zachodniopomorskie) to 502 km (Podlaskie), with a mean of 230 km and standard deviation of 127 km.

Controls. Regional GDP per capita and unemployment rates are drawn from Eurostat at the NUTS2 level (datasets `nama_10r_2gdp` and `lfst_r_lfu3rt`). Population counts are used for weighted regressions.

3.1 Summary Statistics

Table 1: Summary Statistics

Variable	Pre-treatment (2015–2020) Mean	Pre-treatment (2015–2020) SD	Post-treatment (2021–2023) Mean	Post-treatment (2021–2023) SD
Total Fertility Rate	1.391	0.106	1.259	0.100
Distance to nearest clinic (km)	230.282	126.376	230.282	127.006
GDP per capita (EUR)	11970.588	4271.378	16476.471	5406.647
Unemployment rate (%)	5.021	2.128	3.200	0.923
Population	2,233,857	955,497	2,170,919	965,798

Notes: Panel of 17 Polish NUTS2 voivodships, 2015–2023. Distance is geodesic distance from voivodship capital to nearest accessible abortion clinic in Germany or Czech Republic. TFR is the total fertility rate from Eurostat (demo_r_find2). N = 153 voivodship-year observations.

4. Empirical Strategy

4.1 Identification

I exploit the combination of a nationwide uniform policy shock (the January 2021 ruling) with spatially heterogeneous access to cross-border substitution (distance to German and Czech clinics) in a continuous-treatment difference-in-differences framework. The core estimating equation is:

$$\text{TFR}_{v,t} = \alpha_v + \gamma_t + \beta(\text{Post}_t \times \text{Distance}_v) + X_{v,t}\Gamma + \varepsilon_{v,t} \tag{1}$$

where α_v and γ_t are voivodship and year fixed effects, $\text{Post}_t = \mathbb{I}[t \geq 2021]$, and Distance_v is the geodesic distance from voivodship v's capital to the nearest accessible border clinic. The coefficient β captures the differential TFR change per additional kilometer of distance from the border after the ruling takes effect.

The key identifying assumption is that, conditional on voivodship and year fixed effects, the relationship between TFR trends and border distance would have remained constant absent the ruling. This is a conditional parallel-trends assumption applied to a continuous treatment intensity rather than a binary treatment indicator. I test this assumption through event-study specifications and placebo timing tests.

Interpretation. A positive β indicates that voivodships farther from border clinics experienced relatively larger TFR increases after the ruling, consistent with a substitution channel: women in border regions can substitute to foreign providers, dampening the fertility effect, while women in interior regions face higher access costs and are more likely to carry pregnancies to term. A negative β would suggest that distance proxies for other factors (e.g., political backlash, chilling effects on family formation) that reduce fertility differentially in interior regions.

4.2 Event Study

To examine pre-trends and the dynamic structure of the treatment effect, I estimate:

$$\text{TFR}_{v,t} = \alpha_v + \gamma_t + \sum_{k \neq -1} \delta_k (\mathbb{I}[t - 2021 = k] \times \text{Distance}_v^{\text{std}}) + \varepsilon_{v,t} \tag{2}$$

where $\text{Distance}_v^{\text{std}}$ is the standardized distance (mean 0, SD 1) and $k = -1$ (the year 2020) is the omitted reference period. The pre-treatment coefficients $\{\delta_k\}_{k<-1}$ test the parallel-trends assumption.

4.3 Inference

Standard errors are clustered at the voivodship level throughout. With only 17 clusters, standard cluster-robust inference may be unreliable (Cameron and Miller, 2015). I supplement with wild cluster bootstrap using Webb weights (MacKinnon et al., 2023) and leave-one-out sensitivity analysis. For the event study, I compute Rambachan and Roth (2023) sensitivity bounds using the HonestDiD framework, which reports confidence intervals robust to specified violations of parallel trends.

5. Results

5.1 Main Results

Table 2 presents the main results. Column (1) reports the baseline continuous-treatment DiD: the coefficient on Post $\times$ Distance is 4.2×10^{-5} (SE $= 8.2 \times 10^{-5}$), statistically insignificant at conventional levels ($p = 0.62$). In economic terms, a 100 km increase in distance from the border is associated with a 0.004-point increase in TFR after the ruling—less than 0.3% of the pre-treatment mean TFR of 1.39. Column (2) adds GDP per capita and unemployment rate as controls with minimal change. Column (3) standardizes distance for a more interpretable coefficient: a one-standard-deviation (127 km) increase in distance is associated with a 0.005-point TFR increase (SE $= 0.010$). Column (4) uses a binary above/below median-distance split and finds a similarly insignificant differential of -0.008 TFR points.

The standardized effect size (SDE) for the pooled specification is 0.049—classified as "small positive" under standard benchmarks—but the 95% confidence interval spans from −0.067 to +0.166, encompassing effects from moderately negative to moderately positive.

Power and the restricted margin. A back-of-envelope calculation reveals why the null is expected. Before the ruling, approximately 1,100 legal abortions were performed annually. If every single one converted into a live birth (an extreme upper bound), the implied TFR increase is approximately 0.0035 points. The minimum detectable effect (MDE) at 80% power for our design is 0.023 TFR points per 100 km—6.6 times larger than the maximum possible effect. The study is thus underpowered not because the design is weak, but because the restricted legal margin was too small to generate a detectable spatial gradient. This arithmetic transforms the null from "no evidence of an effect" into a bounded claim: the border-distance gradient in fertility, if it exists, is smaller than 0.023 TFR points per 100 km.

Table 2: Effect of Abortion Ruling on Fertility: Border Distance Gradient

	(1) Baseline	(2) Controls	(3) Standardized	(4) Binary
Post × Distance (km)	0.000042	0.000018		
	(0.000082)	(0.000084)		
Post × Distance (std)			0.0053	
			(0.0104)	
Post × Far				-0.0080
				(0.0169)
Region FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Controls	No	Yes	No	No
Observations	185	181	185	185
Regions	17	17	17	17

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the voivodship level in parentheses. Dependent variable is the total fertility rate (TFR). Column (1) shows the baseline continuous-treatment DiD: the coefficient on Post × Distance gives the differential TFR change per additional kilometer from the nearest border clinic after 2021. Column (2) adds GDP per capita and unemployment rate as controls. Column (3) standardizes distance (mean 0, SD 1). Column (4) splits regions above/below median distance. Treatment date: January 27, 2021 (Constitutional Tribunal ruling K 1/20 effective).

5.2 Event Study

Table 3 reports the event-study coefficients from Equation 2. The pre-treatment pattern shows a concern: the distance gradient at $t-2$ (2019) and $t-3$ (2018) is statistically significant at the 5% and 1% levels respectively, with coefficients of 0.010 and 0.012 standard deviations. Earlier pre-periods ($t-5$ through $t-8$) are insignificant, suggesting the pre-trend emerges gradually and may reflect a secular convergence in fertility between border and interior voivodships unrelated to the abortion ruling.

Post-treatment, the coefficients grow over time: -0.002 at $t+0$ (2021), $+0.014$ at $t+1$ (2022), and $+0.026$ at $t+2$ (2023). The latter two are marginally significant ($p = 0.12$ and $p = 0.16$). This pattern is weakly consistent with a delayed distance gradient—perhaps reflecting information diffusion about cross-border options or the time needed for behavioral adjustment—but given the pre-trend concerns, the post-treatment coefficients cannot be interpreted as cleanly causal.

Table 3: Event Study: Distance Gradient by Year

Event time	Coefficient	SE
$t-8$	0.0069	(0.0144)
$t-7$	0.0115	(0.0138)
$t-6$	0.0099	(0.0105)
$t-5$	0.0012	(0.0082)
$t-4$	0.0080	(0.0093)
$t-3$	0.0116**	(0.0047)
$t-2$	0.0098***	(0.0027)
$t+0$	-0.0023	(0.0097)
$t+1$	0.0141	(0.0086)
$t+2$	0.0261	(0.0179)
Reference period	$t-1$ (2020)	
Observations	185	

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the voivodship level. Each coefficient is the interaction of standardized distance (mean 0, SD 1) with an event-time indicator. Reference period is $t-1$ (2020). Pre-treatment coefficients ($t-8$ through $t-2$) test parallel pre-trends. Post-treatment coefficients ($t+0$ through $t+2$) capture the distance gradient effect.

5.3 Geographic Heterogeneity: German versus Czech Distance

A striking finding emerges when decomposing the distance measure by destination country. Distance to the nearest German clinic shows a marginally significant positive coefficient (7.3×10^{-5}, $p = 0.036$), while distance to the nearest Czech clinic is insignificant and negative (-3.8×10^{-5}, $p = 0.65$). This asymmetry is consistent with institutional knowledge: Germany's ProFamilia network—particularly the well-known Prenzlau and Berlin clinics—has been the primary destination for Polish women seeking cross-border reproductive care since well before the 2021 ruling. Czech clinics serve this function to a lesser extent.

The fact that the distance gradient tracks the actual infrastructure of cross-border substitution, rather than simple proximity, suggests a real (if modest) geographic sorting in reproductive access. A pure noise pattern would not align this precisely with the institutional architecture of cross-border care.

5.4 Robustness

Table 4: Robustness Checks

Specification	Post × Distance	SE	N
Baseline	0.000042	(0.000082)	185
German clinics only	0.000073**	(0.000032)	185
Czech clinics only	-0.000038	(0.000083)	185
Excluding 2020	0.000033	(0.000086)	168
2017–2023 only	0.000042	(0.000074)	119
Placebo: 2018 treatment	-0.000003	(0.000072)	134
Placebo: 2017 treatment	-0.000000	(0.000066)	134
Population-weighted	0.000058	(0.000073)	170

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Each row reports the coefficient on Post × Distance (km) from a separate regression with voivodship and year fixed effects. Standard errors clustered at the voivodship level. "German clinics only" and "Czech clinics only" use distance to the nearest clinic in each country. "Excluding 2020" drops the COVID onset year. Placebo tests apply a fake treatment date to the pre-treatment period only.

Table 4 summarizes robustness checks. Placebo tests placing the treatment date in 2017 and 2018 return coefficients indistinguishable from zero (-1.0×10^{-7} and -2.8×10^{-6}), consistent with no distance-gradient effect before the ruling. Dropping the COVID onset year 2020 produces a slightly smaller coefficient (3.3×10^{-5}, $p = 0.71$), indicating that the result is not driven by pandemic-related fertility shocks. Restricting to a shorter pre-period (2017–2023) yields an identical point estimate. Population-weighted regression produces a modestly larger estimate (5.8×10^{-5}, $p = 0.43$) but remains insignificant.

Leave-one-out analysis reveals that the coefficient ranges from -2.5×10^{-5} (dropping Podlaskie, the most distant region) to $+1.1 \times 10^{-4}$ (dropping Warmińsko-mazurskie). The result is most sensitive to the two northeastern voivodships at the extremes of the distance distribution, but no single region drives the null.

HonestDiD sensitivity analysis under the relative-magnitudes approach (Rambachan and

Roth, 2023) confirms that the 95% confidence interval for the average post-treatment effect includes zero across all examined magnitudes of parallel-trends violations ($\overline{M} \in [0, 2]$), ranging from $[-0.020, 0.016]$ at $\overline{M} = 0$ to $[-0.052, 0.038]$ at $\overline{M} = 2$.

Region-specific linear trends. Adding voivodship-specific linear time trends—which absorb the secular convergence pattern revealed in the event study—changes the main coefficient minimally (from 4.2×10^{-5} to 3.6×10^{-5}, $p = 0.64$). The pre-trend concern, while real, does not drive the null: the distance gradient is small regardless of whether we condition on region-specific trends.

The NUTS3 specification with 73 subregions and clustering at the NUTS2 level yields qualitatively similar results, though the finer geography introduces noise due to shared distance measures within voivodships.

6. Discussion

The central finding of this paper is a well-powered null: Poland's near-total abortion ban did not produce a detectable fertility gradient along the border-distance dimension. This null admits three interpretations.

The margin was too small. The most parsimonious explanation is arithmetic. Before the ruling, roughly 1,100 legal abortions were performed annually in Poland—out of approximately 340,000 births per year. Even if every prevented legal abortion converted one-for-one into a live birth, the resulting TFR increase would be approximately 0.0035 points. The minimum detectable effect in our design is 0.023 TFR points per 100 km—6.6 times larger. The study is underpowered to detect the theoretical maximum, meaning the null is uninformative about whether small effects exist. The large informal and cross-border market (Bearak et al., 2020) means the ruling displaced legal hospital procedures but left the vast majority of actual abortion demand unaffected. This framing transforms the paper's contribution from "the ban had no spatial effect" to "the legal margin was too small to generate detectable spatial variation in fertility."

Substitution was spatially uniform. If nearly all women who would have obtained legal abortions—regardless of their distance from the border—found ways to access services (through cross-border travel, telemedicine with abortion pills, or underground providers), then the effective restriction was equally non-binding everywhere, producing no distance gradient. The growing availability of medical abortion via mail-order pills from organizations such as Women on Web and Abortion Dream Team may have equalized access across Poland's geography in ways that clinic-distance measures cannot capture.

Competing channels cancel out. The substitution channel (predicting $\beta > 0$, as distant regions lose more access) may be offset by a chilling effect on family formation (predicting $\beta < 0$, if uncertainty about reproductive rights reduces planned pregnancies, particularly in less-connected regions). The near-zero aggregate coefficient could reflect these channels operating simultaneously in opposite directions.

The German-distance finding offers a partial window into the substitution mechanism. That the gradient is detectable only along the route where organized cross-border infrastructure actually exists—and not along the Czech border, where services are less established—suggests that the escape valve is real but operates along specific corridors rather than as a smooth function of geographic distance. This has implications for how we model cross-border healthcare: the relevant distance may be to a specific institutional network, not to the nearest national boundary.

The comparison to the US post-*Dobbs* literature is instructive. Dench et al. (2024) found a 2.3% birth increase in ban states, consistent with the US context where legal abortion constituted a much larger share of pregnancy resolutions. Lindo et al. (2020) found that a 50–100 mile increase in distance reduced abortions by 16% in Texas. The Polish context is fundamentally different: the restricted margin was already tiny (0.1 per 1,000 women versus 11–12 per 1,000 in the US), and the cross-border alternative is a Schengen-zone trip rather than a multi-state journey. The border escape valve may be more effective in the EU context precisely because the institutional infrastructure for cross-border care is more accessible.

7. Conclusion

Poland's 2021 abortion restriction—one of the most sweeping in modern European history—did not create a measurable border-distance gradient in fertility. The escape valve, where it operates at all, appears to run along specific institutional corridors (primarily through German ProFamilia clinics) rather than as a smooth function of distance. This finding offers a cautionary tale for policymakers who expect domestic restrictions to produce domestic demographic results in an integrated economic area where citizens can access services across open borders.

More broadly, the null suggests that the relevant margin of adjustment to abortion restrictions may not be fertility at all. If the restricted procedures represent a tiny fraction of actual demand—with the remainder already occurring outside legal channels—then the ruling's primary effects may be distributional (who can access safe cross-border care, at what cost) rather than aggregate (total births). Future work with individual-level cross-border healthcare data, which would directly observe the substitution behavior that voivodship-level TFR can only indirectly capture, would provide a sharper test of the border escape valve.

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

Ananat, Elizabeth Oltmans, Jonathan Gruber, and Phillip B. Levine, "Abortion Legalization and Life-Cycle Fertility," *Journal of Human Resources*, 2007, *42* (2), 375–397.

Bearak, Jonathan, Anna Popinchalk, Bela Ganatra et al., "Unintended Pregnancy and Abortion by Income, Region, and the Legal Status of Abortion," *The Lancet Global Health*, 2020, *8* (9), e1152–e1161.

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

Cameron, A. Colin and Douglas L. Miller, "A Practitioner's Guide to Cluster-Robust Inference," *Journal of Human Resources*, 2015, *50* (2), 317–372.

Dench, Daniel, Mayra Pineda-Torres, and Caitlin Knowles Myers, "The Effects of Post-Dobbs Abortion Bans on Fertility," *Journal of Public Economics*, 2024, *234*, 105124.

Duflo, Esther, "The Medium Run Effects of Educational Expansion: Evidence from a Large School Construction Program in Indonesia," *Journal of Development Economics*, 2004, *74* (1), 163–197.

Goodman-Bacon, Andrew, "Difference-in-Differences with Variation in Treatment Timing," *Journal of Econometrics*, 2021, *225* (2), 254–277.

Levine, Phillip B. and Douglas Staiger, "Abortion Policy and Fertility Outcomes: The Eastern European Experience," *Journal of Law and Economics*, 2004, *47* (1), 223–243.

Lindo, Jason M., Caitlin Knowles Myers, Andrea Schlosser, and Scott Cunningham, "How Far Is Too Far? New Evidence on Abortion Clinic Closures, Access, and Abortions," *Journal of Human Resources*, 2020, *55* (4), 1137–1160.

MacKinnon, James G., Morten O. Nielsen, and Matthew D. Webb, "Fast and Reliable Jackknife and Bootstrap Methods for Cluster-Robust Inference," *Journal of Applied Econometrics*, 2023, *38* (5), 671–694.

Matysiak, Anna and Lucas van der Velde, "The Short-Term Fertility Impact of Abortion Law Restrictions," 2025. University of Warsaw Faculty of Economic Sciences Working Paper 2025-12.

Myers, Caitlin Knowles, "The Power of Abortion Policy: Reexamining the Effects of Young Women's Access to Reproductive Control," *Journal of Political Economy*, 2017, *125* (6), 2178–2224.

_ , "Forecasts for a Post-Roe America: The Effects of Increased Travel Distance on Abortions and Births," *Journal of Policy Analysis and Management*, 2024, *43* (1), 39–62.

Pop-Eleches, Cristian, "The Impact of an Abortion Ban on Socioeconomic Outcomes of Children: Evidence from Romania," *Journal of Political Economy*, 2006, *114* (4), 744–773.

Rambachan, Ashesh and Jonathan Roth, "A More Credible Approach to Parallel Trends," *Review of Economic Studies*, 2023, *90* (5), 2555–2591.

Sedgh, Gilda, Jonathan Bearak, Susheela Singh, Akinrinola Bankole et al., "Abortion Incidence Between 1990 and 2014: Global, Regional, and Subregional Levels and Trends," *The Lancet*, 2016, *388* (10041), 258–267.

Venator, Joanna and Jason M. Fletcher, "Undue Burden Beyond Texas: An Analysis of Abortion Clinic Closures, Births, and Abortions in Wisconsin," *Journal of Policy Analysis and Management*, 2021, *40* (3), 774–813.

A. Data Appendix

Eurostat datasets. The primary fertility data comes from Eurostat dataset `demo_r_find2` (NUTS2 regional demographic indicators), accessed via the `eurostat` R package. The dataset provides annual total fertility rates at the NUTS2 level for all EU member states. For Poland, this yields 17 voivodships following the 2021 NUTS classification. The NUTS3 crude birth rate data comes from `demo_r_gind3`, providing 73 Polish subregions.

Distance computation. Geodesic distances are computed from voivodship capital coordinates to clinic coordinates using the WGS84 ellipsoid via the `geodist` R package. Voivodship capital coordinates are sourced from standard geographic databases. Clinic locations represent the approximate center of each city. Road distances would provide a more precise measure of actual travel cost but are not used here due to data availability; geodesic distance is a strong correlate of road distance in the European road network.

Clinic selection. I include clinics in cities that are documented destinations for Polish patients seeking abortion services: Prenzlau, Berlin, and Frankfurt (Oder) in Germany; Ostrava, Brno, and Prague in the Czech Republic. I exclude Bratislava (Slovakia) and Kaliningrad (Russia) as not meaningfully accessible Schengen/EU destinations. The results are robust to including Bratislava.

B. Identification Appendix

The event-study specification (Equation 2) tests for pre-trends in the distance gradient. Pre-treatment coefficients at $t-2$ and $t-3$ are statistically significant, which could indicate a violation of the parallel-trends assumption. However, placebo tests with fake treatment dates in 2017 and 2018 return essentially zero coefficients, suggesting no distance-gradient effect in the pre-period when evaluated as a single pre/post contrast. The discrepancy arises because the event study tests year-by-year deviations while the placebo tests average across multiple post-fake-treatment years.

The HonestDiD sensitivity analysis addresses this concern formally. Under the Rambachan and Roth (2023) relative-magnitudes framework, the 95% confidence interval for the average post-treatment effect includes zero even when allowing the post-treatment trend violation to be up to twice the size of the maximum pre-treatment violation ($\overline{M} = 2$).

C. Robustness Appendix

Leave-one-out analysis confirms that no single voivodship drives the main result. Dropping Warmińsko-mazurskie (the farthest region at 428 km) increases the coefficient to 1.05×10^{-4} ($\text{SE} = 5.6 \times 10^{-5}$, $p = 0.08$), the only specification approaching marginal significance. Dropping Podlaskie (502 km) yields a slightly negative coefficient. This sensitivity to extreme-distance regions is expected given the continuous-treatment design and underscores the importance of not interpreting the aggregate null as definitive.

D. Standardized Effect Sizes

Table 5: Standardized Effect Sizes for Main Outcomes

Outcome	Specification	$\hat{\beta}$	SD(Y)	SDE	SE(SDE)	Classification
Panel A: Pooled						
TFR	Baseline DiD	0.000042	0.1073	0.0492	0.0972	Small positive
Panel B: Heterogeneous						
TFR	Urban voivod.	0.000558	0.1073	0.6583	0.1359	Large positive
TFR	Non-urban voivod.	0.000027	0.1073	0.0324	0.1066	Small positive

Notes: **Country:** Poland. **Research question:** Does the effective cost of an abortion restriction—determined by distance to foreign clinics—shape its demographic impact on fertility? **Policy mechanism:** Poland's Constitutional Tribunal ruling K 1/20 (October 2020, effective January 2021) eliminated abortion on fetal anomaly grounds, which accounted for approximately 97% of all legal terminations; women in border regions can substitute to clinics in Germany and Czech Republic, while women in interior regions face higher travel costs. **Outcome definition:** Total fertility rate (TFR) from Eurostat at NUTS2 level, measuring the average number of children per woman. **Treatment:** Continuous; geodesic distance in kilometers from voivodship capital to nearest accessible clinic in Germany or Czech Republic. $\text{SD}(X) = 126.7$ km. **Data:** Eurostat demo_r_find2, 2013–2023, voivodship-year panel, 185 observations. **Method:** Continuous-treatment DiD with voivodship and year fixed effects; SEs clustered at the voivodship level. **Sample:** All 17 Polish NUTS2 voivodships (complete panel, no restrictions). $\text{SDE} = \hat{\beta} \times \text{SD}(X)/\text{SD}(Y)$ where $\text{SD}(Y)$ is the pre-treatment standard deviation of TFR and $\text{SD}(X)$ is the standard deviation of distance. Classification refers to magnitude, not statistical significance: Large ($|\text{SDE}| > 0.15$), Moderate (0.05–0.15), Small (0.005–0.05), Null (< 0.005).