UCL DEPARTMENT OF POLITICAL SCIENCE
SCHOOL OF PUBLIC POLICY

Essay Front Page

Please complete this sheet electronically and attach it as the front page
for each essay submitted.

Candidate Number: DWKB3

Essay Title, Question Number/s or Text of Question/s you have answered:
Assessment 1

Essay Number: 1
Module Code: POLS0012

Module Title: Causal Analysis

Submission Date: 14/02/2024

Word Count: 2997

The essay cover sheet does not count towards the word count on your submission.

Please tick this box if you **do not** agree to your essay being used anonymously in future teaching ☐

PART A

Question 1

a)

i) Estimate the first-stage while clustering SEs by municipality. Interpret:

The estimate of -0.249 indicates that for every 1% distance increase between respondents' municipality and the Turkish coast, the probability of the municipality receiving immigrants decreases by -0.00249 percentage points.

ii) Estimate the first-stage F-statistic. NOT necessary to interpret:

The F-statistic is 152.590.

iii) How relevant do you think the instrument is?

The first-stage estimate is non-zero and significant at the 0.1% level. The F-statistic of 152.590 is considerably above 10, the minimum for an instrument to not be considered weak. It is significant at the 0.1% level, rejecting the null hypothesis of no instrument effect. Accordingly, the instrument is highly relevant.

b)

i) Explain the independence/randomisation assumption in the context of this study:

Municipalities' logged km distance from the Turkish coast must be as-if randomly assigned, independent of covariates and potential outcomes. This is necessary to introduce exogenous variation in municipal immigrant arrival and avoid biased causal estimates. Islands are expected to be identical across many observable and unobservable characteristics, supporting independence (p.444). Not all covariates can be perfectly balanced; factors as marine currents and proximity to other islands are likely to be non-randomly distributed among coastal distances while affecting immigrant arrival. To maintain randomisation, importantly imbalanced covariates must be controlled for.

ii) Explain the exclusion restriction in the context of this study, *and whether I believe it is likely to hold.*

The only way a municipality's distance from the Turkish coast affects residents' hostility attitudes is through its impact on the probability of immigrant arrival. This isolates immigrants' impact on attitudes, excluding confounding effects. The exclusion restriction likely holds, as factors which could affect attitudes, like political competition, governance, and public services do not vary with distance, supported by shared administrative systems between islands. Coastal proximity might increase interactions with Turkey, potentially affecting attitudes independently of immigrant arrival. Distances vary widely, from over 300km to less than 5km (Google Maps, 2024). However, the authors found no significant effect for a wide range of placebo outcomes, confirming the exclusion restriction's validity.

c) Replicate the second-stage coefficients and standard error for *score-asylum* and *score-imming (both outcomes)*, reported on pp.449-50 (1^{st} and 7^{th} estimates in the figure), clustering the SEs by municipality.

The estimated LATEs indicate that receiving immigrants in a municipality led to a 0.218-point and 0.212-point increase in hostility of its residents towards asylum seekers and immigrants on the PCA scale, respectively. Both estimates are significant at the 0.1% level (p-values = 8.18e-06 and 1.32e-05) and have standard errors of 0.054 and 0.070 when clustering by municipality, respectively.

d)

i) Explain, in this case, what type of municipality is a complier and what type of municipality is an always-taker:

Compliers are municipalities only receiving immigrants when they are within 40km of the Turkish coast. Always-taker-municipalities receive immigrants regardless of their coastal distance.

ii) Calculate the proportion of compliers and the ITT effect:

The ITT is 0.222.
93.1% of municipalities are compliers.

iii) Use the answer from ii) to calculate the CACE:

The CACE is 0.239.

iv) Calculate the p-value of this CACE estimate:

The p-value is 4.69e-07.

Question 2

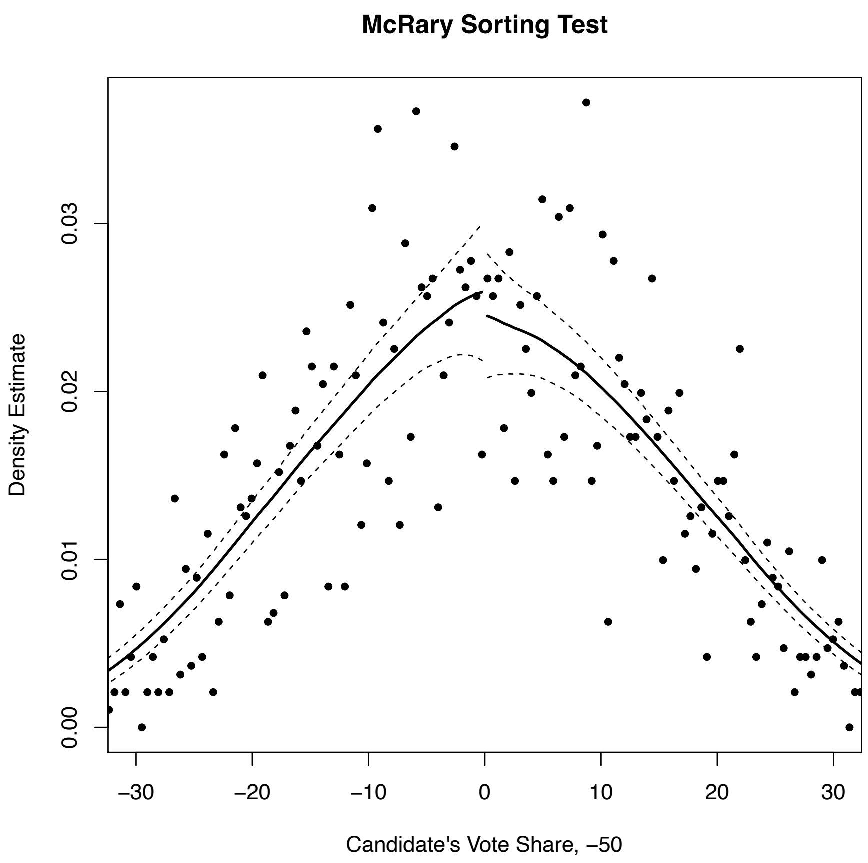
a) Explain why a regression discontinuity design (RDD) could help estimate the causal effect of the election of a Democrat sheriff on the enforcement rate, and why this is likely to be superior to calculating the mean difference between Democratic and Republican sheriffs.

Differences in mean enforcement rates between counties with Democratic versus Republican sheriffs could reflect pre-existing socioeconomic and political factors, not the sheriff's party affiliation. Regression discontinuity (RD) addresses this: within-county conditions are nearly identical at the cutoff, approximating random assignment and isolating the sheriff's party effect.

b) Briefly, discuss two potential violations of the RDD assumptions that could affect the inferences we make: how plausible are these violations, in your view? Provide evidence from the data to evaluate **one** of them quantitatively.

Considering violations of randomness, composite treatments could arise if change at the electoral threshold is not limited to the sheriff's party affiliation, including variation in rules or policies influencing enforcement rates. Examples include changes in law enforcement policies or funding, possibly confounding the treatment effect. This is implausible, as persistent partisan-related changes would undermine implementation and the system's credibility. Sorting could occur if sheriffs engaged in electoral fraud. Investigating discontinuous jumps in the density of the running variable near the threshold with a McRary test shows a slight discontinuity (see Figure 1). However, it is insignificant (p-value = 0.454), failing to reject the null hypothesis of no density jump, indicating no evidence of sorting or randomness violation.

Figure 1: McRary Sorting Test for Electoral Fraud



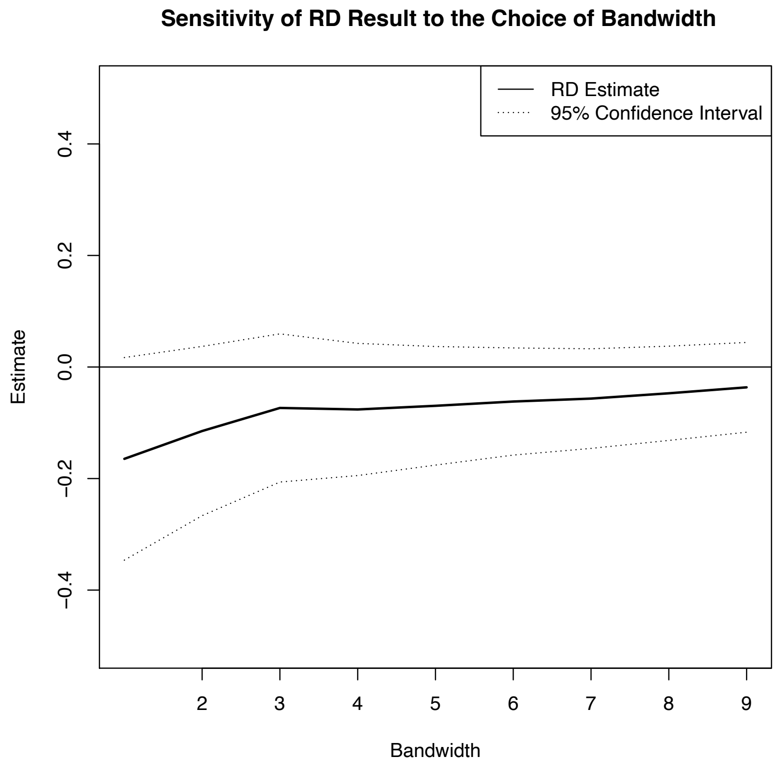
c) Estimate the LATE of electing a Democratic sheriff on enforcement using RDD analysis with the "optimal" bandwidth. Report your results.

The estimated LATE is -0.150, significant at the 10% level (p-value = 0.091). There is no strong evidence of a treatment effect at the cut-off, aligning with the paper.

d) Re-estimate the effects in question (c) using all whole-number bandwidths from 2 to 10. Use a figure to illustrate how sensitive your results are to these changes. Briefly, interpret your results.

Increasing the sample size narrows the 95% confidence interval, illustrating the bias-variance trade-off. Variance is inversely related to the sample size. However, larger samples increase the risk of including outliers when estimating patterns at the cut-point, inflating the probability of bias. As the bandwidth increases, estimates tend towards zero, indicating model dependence. Weak sensitivity suggests that the results obtained at the optimal bandwidth are unlikely to be due to chance.

Figure 2: Sensitivity of Results to the Bandwidth



e) An alternative way to evaluate the hypothesis that the election of a Democratic sheriff has a causal effect on enforcement of ICE detention requests is to use a fixed effects design. Implement an appropriate model to estimate this and **report** your result.

The estimated LATE of -0.002 is insignificant (p-value = 0.129).

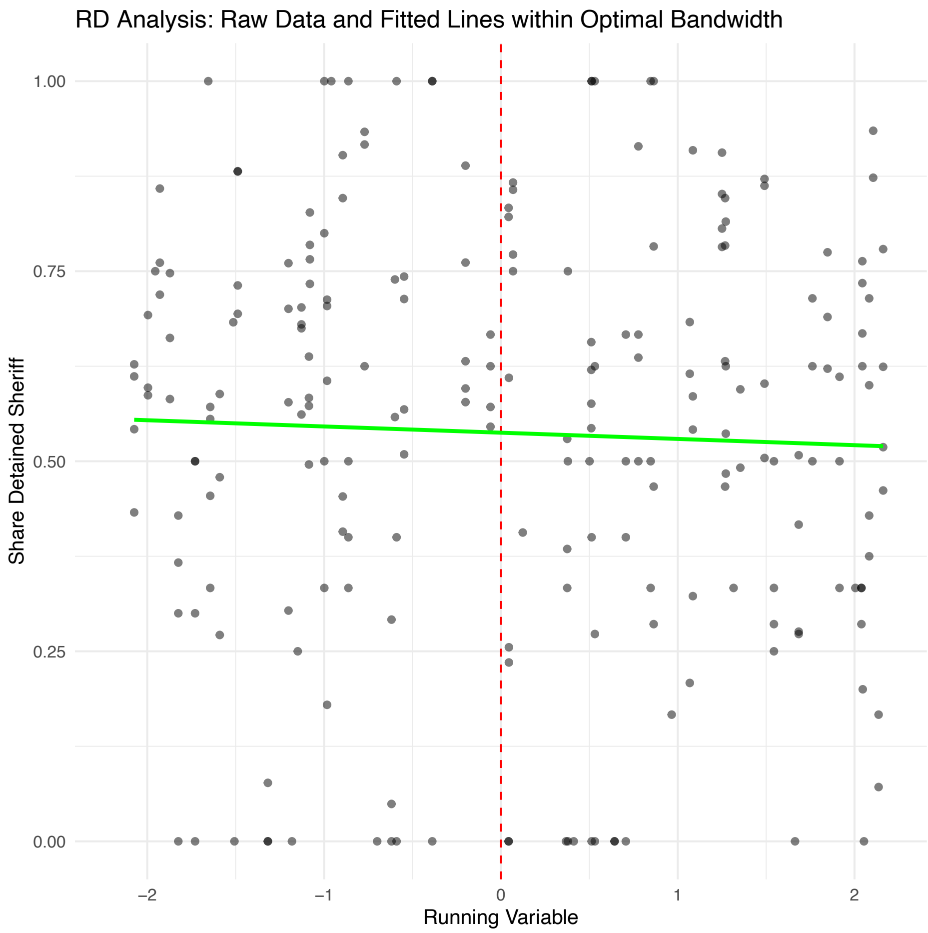
f) Which of the two approaches in this question – regression discontinuity or fixed effects – do you find more convincing as an estimate of the causal effect? Why?

RD offers a more compelling estimate due to its stronger internal validity. Fixed-effects models, while having broader external validity, face significant challenges:

They assume parallel trends, i.e., that counties would have followed the same enforcement rate trend prior to the election of a Democrat sheriff, challenging to prove due to the lack of a baseline when all counties were untreated. Subsetting the trend analysis to counties with an initial Republican sheriff may not accurately represent those with an initial Democrat sheriff. Additionally, counties experience treatments at different times, depending on sheriff elections. Analysing over multiple mandates increases the likelihood of the assumption holding by providing more data points to establish a trend but reduces the number of initially untreated counties. Data imbalance caused by missing values further complicate trend analysis, and potential time-varying confounders, unaccounted for due to data limitations, may falsely suggest parallel trends. Although fixed-effects have a higher external validity by including more counties than RDD, these observations compromise their internal validity.

Conversely, RD's limitations, such as sparse data within the optimal bandwidth (219 observations) and the inability to perform placebo regressions due to the lack of covariates, are outweighed by other considerations. Possible violations of randomness, through composite treatments or sorting, were refuted in (b). The implausibility of sorting is further strengthened, considering that high scrutiny during elections reduces the potential for fraud. Moreover, I proved in (d) that model dependence is marginal, suggesting a robust functional form. Estimating RD effects at placebo cut-offs within the optimal bandwidth reinforces this conclusion, with no significant effect at -2,-1, -0.5, 0.5, and 2. The estimate at 1 is only significant at the 10% level, indicating no important discontinuity between the running variable and outcome. The results should not be driven by chance alone. Figure 3 illustrates this and the low estimated RD effect.

<u>Figure 3: Raw Data and Fitted RD Regression</u>



Accordingly, RD is more convincing, particularly when sheriffs are elected by a narrow margin.

PART B

Study background and objectives:

Residential solar panels (SP) contribute to approximately 45% of Switzerland's total solar energy production and 85% of the sector's growth (IEA, 2021). Understanding SPs' impact on energy household consumption is key for the country's *Energy Strategy 2050*, balancing increasing demand with environmental concerns. The solar rebound effect (SRE), where solar panel efficiency gains are partly offset by increased energy consumption, poses a challenge.

Evidence on SREs varies across regions. Beppler, Matisoff, and Oliver's (2021) study identified a 28.5% SRE in the US, while research on Dutch households reported an SRE of 7.7% (Aydın, Brounen and Ergün, 2023). Zimmerman (2021) noted economy-wide energy-technology rebound effects in Switzerland; Filippini and Obrist (2022) raised concerns about the actual energy savings in green-certified Swiss houses. These findings highlight a gap in the understanding of SREs and questions the current framework's effectiveness.

I aim to fill this gap by using instrumental variables (IV), providing an unbiased estimate of SP's influence on Swiss households' average yearly energy consumption. I use the suitability for SP of single-family houses, owned by residents, as my instrument. This creates exogenous variation in SP adoption (treatment), mitigating biases from direct consumption comparisons between homes with and without SP, addressing issues of self-selection and unobserved error. My findings are expected to refine Switzerland's energy strategy, better aligning with household energy consumption.

Data:

Instrument:
I identify single-family homes using the *Federal Register of Buildings and Dwellings*, classifying the type of every building nationwide (FSO, 2024). Suitability for SPs is assessed using data from the Swiss Geo-Portal. This involves averaging the solar potential of buildings' facades and roofs, categorised as low, medium, high, very high, or excellent based on government estimations. These classifications consider solar energy exposure, shading, and the pitch and orientation of facades and roofs. They assume a 20% SP efficiency, the market standard.

Treatment:
Treatment data, sourced from SFOE, covers all photovoltaic installations nationwide with a production potential of 2kW or more, provided they are registered with a guarantee of origin or received government subsidies (SFOE, 2024). I assume that all residential SP buyers are registered, given 1-to-2-bedroom houses consume 2.5kW on average, people's inclination to save money, the availability of subsidies with no conditions, and the necessity of a guarantee of origin to sell surplus production as premium "green energy" (Rawal, 2024; Pronovo, 2023). This aligns with the Federal Energy Department's estimation of Switzerland's solar energy production (SFOE, 2022). The SFOE data provides information on SPs' location, capacity, and operational status. It also reduces the risk of self-reporting errors or biases occurring when asking people if and since when they have SPs.

Outcome:
The outcome is average household annual energy consumption in kW, due to seasonal demand fluctuations. I exclude households residing in a location for less than a year, and only consider complete years of energy consumption and SP installation. Data is collected through a national-level survey using random sampling, with differential incentives for participation. Quotas are set for each instrument level, and the survey is conducted via mail, online, and in-person methods to mitigate non-response and ensure representation. The public land register specifies building ownership, facilitating an initial exclusion of single-family houses owned by corporate landlords (Cadastre, n.d.). Resident status is confirmed during the survey.

Model Specification:

The two-stage least squares (2SLS) estimator is employed because the instrument is non-binary. The LATE quantifies the average SP installation effect on mean yearly energy consumption among compliers, those influenced by their property's solar potential. Additionally, 2SLS allows for the inclusion of covariates if imbalances are detected.

Assumptions:

- **Exclusion restriction:**

Houses’ solar potential should only affect residents’ energy consumption through its influence on SP installation.

A violation arises if solar potential prompts residents to research energy consumption optimisation, changing energy use independently of SPs. This seems unlikely. Most homeowners are probably unaware of their solar potential unless considering SP installation. Furthermore, awareness of solar potential alone is not expected to prompt most individuals to completely rethink their energy consumption habits or undertake extensive home renovations, such as improving insulation, which significantly impacts energy use. Instead, sustainability and financial concerns driving most energy consumption changes, likely stem from a broad range of reasons (American Energy Department, n.d.).

Another violation relates to solar potential's influence on energy consumption via its effect on indoor temperature. Examination suggests that any such effect is minimal. Firstly, in winter, reduced solar energy from the sun's lower position in the sky and shorter days diminishes temperature effects caused by sunlight, shade, pitch, and orientation. Secondly, in mountainous regions, snow accumulation on roofs acts as insulation, further reducing the variability in indoor temperature induced by solar potential (Flame Furnace, 2019). Thirdly, factors like building material and age, maintenance, ventilation, insulation, colour, weather patterns, and technological add-ons also play significant roles in house temperature, mitigating solar potential’s individual impact (Mikku & Sons, 2023). Finally, only 5% of Swiss households have air conditioning due to strict regulations, making it challenging to address potential temperature increases caused by a higher solar potential (Randazzo et al, 2020, Bachmann, 2022).

To check for violations, I will compare IV and OLS estimates. Larger IV estimates may signal a violation, as they are less susceptible to the selection bias likely inflating OLS estimates. This does not confirm a violation and could result from finite sample bias.

- **Independence**

Solar potential should be as-if randomly assigned, unrelated to residents' potential outcomes and covariates. I conduct balance tests to assess plausibility, regressing suspected imbalanced covariates on the instrument. If coefficients are significant at the 5% alpha level, I reject the null hypothesis of no correlation, indicating a violation. When the magnitude is large, imbalanced covariates are controlled for and estimates with and without the controls presented. The following are investigated:

Building characteristics
Poor energy efficiency leads to higher consumption and expenses, instigating consideration of grid-energy-saving options like SPs. Accordingly, variables impacting energy consumption are probably confounders. These variables include those highlighted in the third justification defending the exclusion restriction. Additionally, the *Federal Register of Buildings and Dwellings* contains information on residential units nationwide, including area, heating system, and the number of floors and rooms, all impacting energy consumption.

Weather characteristics
Wind speed and exterior temperature impact energy consumption, regardless of building characteristics. I assess their balance at the postcode level using data from MeteoSwiss.

Household-specific characteristics
Household variables as income, wealth, electricity prices, and municipal and cantonal solar subsidies influence SP affordability and return on investment, and thus installation. The survey investigates political leaning and environmental consciousness, providing insights into potential outcomes and individuals' reaction to the instrument.

Instrument spatial distribution
Certain house types, associated with specific instrument values, may be more prevalent in particular areas. Chalets often feature steep roofs in mountain regions, while modern houses with flat roofs are common in urban areas. In reality, the impact of roof pitch on solar potential is limited: SPs on flat roofs are typically mounted on supports for optimal inclination, and most steep roofs fall within a range (13% to 45%) with less than a 10% impact (Fouché, 2024). However, flat roof structures are not accounted for by the instrument. Nonetheless, roof pitch distribution is relatively balanced, with 48% pitched and the remainder flat (Sager, 2018). The instrument addresses this concern by neutralising potential trends associated with building attributes through the randomness of other factors affecting solar suitability, including weather conditions, tree coverage, and street width.

Potential outcomes' spatial distribution
Potential outcomes may vary spatially because of culture; comparing individuals across regions likely violates independence. Even with balanced covariates, residents facing the same instrument level may perceive it differently. Accordingly, we control for cantons. The analysis remains limited by residents' nationality, another cultural marker. Balance tests are unnecessary as Swiss nationals, followed by European nationals outnumber other nationalities (FSO, 2022), and nationalities are geographically clustered (Nguyen, 2017). Despite the imbalance, nationalities are not controlled for, avoiding introducing too many controls and weakening the instrument.

- **Relevance**

Solar potential should strongly influence SP installation. Weak instruments lead to small or imprecise first-stage estimates, causing 2SLS estimates to resemble OLS estimates of the treatment-outcome relationship. 2SLS estimates then become vulnerable to finite sample bias, potentially overestimating treatment effects.

Pre-study
Finite sample bias is unlikely when a single instrument is used. Additionally, the assumption is supported by solar potential's direct correlation with economic benefits and carbon footprint reduction, key motivations for SP adoption. Furthermore, the proportion of compliers should be greater than zero, indicating that some individuals install SPs due to their solar potential. This is supported by the installed solar capacity's 50% annual growth rate, achievable only through compliers and always-takers (assuming monotonicity). Given always-takers seem unlikely, installing solar panels no matter what, attributing this increase uniquely to their proliferation appears excessive (IEA, 2021).

During-study
The first-stage coefficient must be non-zero and its F-statistic exceed 10, enabling the rejection of the null hypothesis of no instrument effect. Large and significant reduced-form estimates, unaffected by finite sample bias, further suggest a robust instrument. Comparing 2SLS to Limited Information Maximum Likelihood (LIML) estimates provides insights into the robustness of 2SLS estimates. LIML estimates, less prone to bias from weak instruments, should align with the 2SLS estimates.

- **Monotonicity:**

Higher solar potential should have no effect or increase the likelihood of SP installation. It is improbable that some individuals have a rising propensity to abstain from installing solar panels as their house's solar potential augments, particularly given most people are likely unaware of their solar potential unless having actively researched SPs. An extreme case would be somebody installing SPs despite having a solar potential of zero, making them a defier or always-taker. However, no such house exists, every building being exposed to some solar radiation. To test monotonicity, I assess whether the proportion of people installing SPs increases with solar potential. I also run robustness checks and sensitivity tests.

- **Non-interference:**

A house' solar potential or SPs should not influence others' SP installation or energy consumption. Violations are improbable and challenging to verify since people encounter solar panels beyond the direct proximity of their house. Tracking them to infer SP exposure and its impact raises technical and ethical considerations.

Limitations:

- **Attrition**:

Attrition is not a concern, using retrospective longitudinal data within a cross-sectional survey design.

- **External validity**:

Natural experiment results are mostly context specific. The exclusion restriction's validity depends on air conditioning regulation. Additionally, the influence of instrument components (roof pitch, orientation, etc.) on solar energy production varies with latitude. Results will be most applicable to houses located at latitudes similar to Switzerland. The aim to inform the Swiss Energy 2050 Strategy and consistent national latitude mitigate these concerns.

This study only considers single-family homes owned by their resident. Results are expected to differ for tenants, as financial returns are typically diminished; with energy surplus profits going to the landlord, while tenants benefit from cheaper electricity (Marshall, 2022). In buildings with multiple residential units, SPs' impact is divided among units, and conflicting preferences among residents may arise.

- **Internal validity:**

Imputing missing outcome and covariates helps address non-response. Survey participation incentives may introduce bias by attracting those benefitting the most. Although differential incentives mitigate this, some non-randomness may persist.

Solar companies or municipalities may engage in targeted SP advertising, favouring neighbourhoods with greater sunlight exposure, specific socioeconomic backgrounds, and high energy prices. This could bias first-stage estimates, requiring further investigation and potential control for advertising campaigns.

Households could instal SPs next to their house, omitted in this study. This is expected to be rare due to visual concerns.

Averaging facade and wall solar potential may oversimplify the instrument. Roof panels are likely preferred for aesthetic reasons and offer higher efficiency than facades, potentially influencing installation decisions more significantly.

Conclusion:

This study uses IV to provide unbiased estimates of the impact of residential SPs on households' average yearly energy consumption, contributing to Switzerland's Energy Strategy 2050. Focusing on single-family homes owned by residents ensures a strong analysis framework. While the LATE provides estimates at the national level, further research could investigate cantonal level effects, contributing to better define regional green-energy subsidies. Results are expected to facilitate decision-making and the achievement of national sustainable energy goals.

Reference list

American Energy Department (n.d.). *Reducing Electricity Use and Costs*. [online] Energy.gov. Available at: https://www.energy.gov/energysaver/reducing-electricity-use-and-costs#:~:text=Reducing%20energy%20use%20in%20your [Accessed 7 Feb. 2024].

Aydın, E., Brounen, D. and Ergün, A. (2023). The rebound effect of solar panel adoption: Evidence from Dutch households. *Energy Economics*, 120(ISSN 0140-9883), p.106645. doi:https://doi.org/10.1016/j.eneco.2023.106645.

Bachmann, H. (2022). *Why getting permission for air conditioners is so hard in Switzerland*. [online] The Local CH. Available at: https://www.thelocal.ch/20220712/why-getting-permission-for-air-conditioners-is-so-hard-in-switzerland#:~:text=Geneva%27s%20rules%20are%20among%20the,your%20health%20condition%20requires%20it. [Accessed 13 Jan. 2024].

Beppler, R.C., Matisoff, D.C. and Oliver, M.E. (2021). Electricity consumption changes following solar adoption: Testing for a solar rebound. *Economic Inquiry*, [online] 61(1). doi:https://doi.org/10.1111/ecin.13031.

Cadastre (n.d.). *Objectives & organisation*. [online] cadastre.ch – The Swiss cadastral surveying. Available at: https://www.cadastre.ch/en/gb/objective.html [Accessed 3 Feb. 2024].

Filippini, M. and Obrist, A. (2022). Are households living in green certified buildings consuming less energy? Evidence from Switzerland. *Energy Policy*, 161, p.112724. doi:https://doi.org/10.1016/j.enpol.2021.112724.

Flame Furnace (2019). *What Snow on Your Roof Says About Your Home - FLAME Furnace*. [online] flamefurnace.com. Available at: https://flamefurnace.com/blog/what-snow-on-your-roof-says-about-your-home/#:~:text=BONUS%3A%20Snow%20on%20your%20Roof%20Means%20Extra%20Insulation&text=Paul%20Fisette%2C%20wood%20technologist%2C%20explains [Accessed 7 Feb. 2024].

Fouché, L. (2024). *Quelle inclinaison et orientation pour mes panneaux solaires ? | Choisir.com*. [online] www.choisir.com. Available at: https://www.choisir.com/energie/articles/137322/choisir-lorientation-et-linclinaison-de-panneaux-solaires-le-guide-complet [Accessed 7 Feb. 2024].

FSO (n.d.). *Federal Register of Buildings and Dwellings*. [online] www.bfs.admin.ch. Available at: https://www.bfs.admin.ch/bfs/en/home/registers/federal-register-buildings-dwellings.html [Accessed 3 Feb. 2024].

FSO (2022). *Population by Migratory Status*. [online] www.bfs.admin.ch. Available at: https://www.bfs.admin.ch/bfs/fr/home/statistiques/population/migration-integration/selon-statut-migratoire.html#:~:text=Pr [Accessed 7 Feb. 2024].

HANGARTNER, D., DINAS, E., MARBACH, M., MATAKOS, K. and XEFTERIS, D. (2018). Does Exposure to the Refugee Crisis Make Natives More Hostile? *American Political Science Review*, [online] 113(2), pp.1–14. doi:https://doi.org/10.1017/s0003055418000813.

IEA (2021). *National Survey Report of PV Power Applications in SWITZERLAND 2021 PVPS Task 1 Strategic PV Analysis and Outreach*. [online] *International Energy Agency*. International Energy Agency. Available at: https://iea-pvps.org/wp-content/uploads/2023/02/IEA-PVPS-National-Survey-Report-Switzerland-2021.pdf [Accessed 17 Jan. 2024].

Marshall, E. (2022). *Rooftop solar: A win/win for landlords and tenants?* [online] Walker Morris. Available at: https://www.walkermorris.co.uk/in-brief/rooftop-solar-a-win-win/ [Accessed 30 Jan. 2024].

MeteoSwiss (2024). *Local forecasts 1205 Geneva - MeteoSwiss*. [online] www.meteoswiss.admin.ch. Available at: https://www.meteoswiss.admin.ch/local-forecasts/geneva/1205.html#forecast-tab=detail-view [Accessed 7 Feb. 2024].

Mikku & Sons (2023). *Understanding the Impact of Roof Shape on Energy Efficiency*. [online] Mike and Sons. Available at: https://www.mikkuandsons.com/impact-of-roof-shape-on-energy-efficiency/ [Accessed 13 Jan. 2024].

Nguyen, D.-Q. (2017). *Defining the 25% foreign population in Switzerland*. [online] SWI swissinfo.ch. Available at: https://www.swissinfo.ch/eng/society/migration-series-part-1-_who-are-the-25-foreign-population-in-switzerland/42412156.

OpenAI (2024). *ChatGPT*. [online] chat.openai.com. Available at: https://chat.openai.com Used to verify if there was any typo in the submitted assessment.

Pronovo (2023). *Garanties d'Origine – Pronovo AG*. [online] Pronovo. Available at: https://pronovo.ch/fr/garanties-dorigine/ [Accessed 17 Jan. 2024].

Qualtrics (2024). *Market Research Software Platform & Tool Set*. [online] Qualtrics. Available at: https://www.qualtrics.com/uk/market-research/.

Randazzo et al. (2020). Air conditioning and electricity expenditure: The role of climate in temperate countries. *Economic Modelling*, [online] 90(ISSN 0264-9993), pp.273–287. doi:https://doi.org/10.1016/j.econmod.2020.05.001.

Rawal, A. (2024). *How Many Solar Panels Do I Need in the UK? | GreenMatch*. [online] www.greenmatch.co.uk. Available at: https://www.greenmatch.co.uk/blog/how-many-solar-panels-do-i-need [Accessed 30 Jan. 2024].

Sager (2018). *Les Suisses adorent les toitures en pente, mais...* [online] Sager SA. Available at: https://www.sager.ch/bulletin-contact-1/kontakt-2018-01-fr.pdf [Accessed 7 Feb. 2024].

SFOE (2024). *Elektrizitätsproduktionsanlagen in der Schweiz*. [online] Elektrizitätsproduktionsanlagen in der Schweiz. Available at: https://www.uvek-gis.admin.ch/BFE/storymaps/EE_Elektrizitaetsproduktionsanlagen/?lang=fr [Accessed 17 Jan. 2024].

Swiss Confederation (2023). *Swiss Geoportal*. [online] geo.admin.ch. Available at: https://map.geo.admin.ch/?lang=en&topic=energie&bgLayer=ch.swisstopo.swissimage&catalogNodes=2419 [Accessed 15 Jan. 2024].

Swiss Federal Office of Energy (SFOE) (2023). *Suitability of roofs for use of solar energy*. [online] www.bfe.admin.ch. Available at: https://www.bfe.admin.ch/bfe/en/home/supply/statistics-and-geodata/geoinformation/geodata/solar-energy/suitability-of-roofs-for-use-of-solar-energy.html [Accessed 13 Jan. 2024].

Swiss Geographic Catalogue (2024). *geocat.ch*. [online] www.geocat.ch. Available at: https://www.geocat.ch/geonetwork/srv/eng/catalog.search#/metadata/b614de5c-2f12-4355-b2c9-7aef2c363ad6 [Accessed 15 Jan. 2024].

THOMPSON, D.M. (2019). How Partisan Is Local Law Enforcement? Evidence from Sheriff Cooperation with Immigration Authorities. *American Political Science Review*, 114(1), pp.222–236. doi:https://doi.org/10.1017/s0003055419000613.

Zimmermann, M., Vöhringer, F., Thalmann, P. and Moreau, V. (2021). Do rebound effects matter for Switzerland? Assessing the effectiveness of industrial energy efficiency improvements. *Energy Economics*, 104, p.105703. doi:https://doi.org/10.1016/j.eneco.2021.105703.

Appendix

```
# Set working directory
setwd("/nfs/cfs/home2/zctq/zctqdor/Causal final exam part A")

# Load data for both questions
load("/nfs/cfs/home2/zctq/zctqdor/Causal final exam part A/2023essay_q1.Rda")
load("/nfs/cfs/home2/zctq/zctqdor/Causal final exam part A/2023essay_q2.Rda")

# Load all packages
library(AER) # For IV 2SLS
library(fixest)
library(lmtest)
library(multiwayvcov)
library(Synth)
library(rdrobust)
library(rdd)
library(ggplot2)


# Question 1

# Variables:
# munid:ID number of the municipality where the respondent lives (typically, an island) -->
municipalities to cluster
# treatment: =1 if the respondent's municipality received immigrants, 0 otherwise --> the
treatment
# logdistance: Distance of the municipality from the Turkish coastline in kilometres, logged -
-> the primary instrument
# low_distance: =1 if municipality is < 40km from the Turkish coastline, 0 otherwise -->
alternative instrument for d
# score_asylum: Summary scale (from PCA), hostility towards asylum seekers (higher values
= more hostile) --> outcome 1
# score_immig: Summary scale (from PCA), hostility towards immigrants (higher values =
more hostile) --> outcome 2


# a) The objective is to examine the relevance of the continuous instrument, logdistance


#i) Estimate the first stage of the authors' two-stage least-squares estimation, clustering the
standard errors by municipality.
first_stage <- lm(treatment ~ logdistance,data = g)
summary(first_stage) # Estimate: -0.249056, significant at the 0.1% level (p < 2e-16)
```

# ii) Estimate the first stage F-statistic:
waldtest(first_stage, vcov=cluster.vcov(first_stage,g$munid))# F-statistic = 152.59, significant at the 0.1% level (p < 2.2e-16)

iii) not coded.

b) not coded.

c) Replicate the second-stage coefficients and standard errors for score-asylum and score-immig (both outcomes), reported on pp.449-50, clustering SEs by municipality.

# Calculating for asylum
TWO_SLS_asylum <- ivreg(score_asylum ~ treatment|logdistance, data = g)
summary(TWO_SLS_asylum) # estimate 0.21762, p = 8.18e-06

coeftest(TWO_SLS_asylum, cluster.vcov(TWO_SLS_asylum,g$munid)) # SE = 0.054465

# Calculating for immigrants
TWO_SLS_immig <- ivreg(score_immig ~ treatment|logdistance, data = g)
summary(TWO_SLS_immig) # estimate 0.21240, p = 1.32e-05

coeftest(TWO_SLS_immig, cluster.vcov(TWO_SLS_immig,g$munid)) # SE = 0.070409

d) Use the binary instrument low_distance and the outcome variable score_asylum to re-do the paper's analysis using a Wald estimator.

i) no code

ii) # Calculate the proportion of compliers and ITT --> two-sided in principle, as islands could be too far from the coast to be in the treatment group and still receive migrants, or be close enough to be in the treatment group but not receive migrants for some reason.

itt <- mean(g$score_asylum[g$low_distance==1]) -
mean(g$score_asylu[g$low_distance==0]) # 0.2224224

prop.c <- sum(g$treatment[g$low_distance==1])/length(g$treatment[g$low_distance==1]) -

```
  sum(g$treatment[g$low_distance==0])/length(g$treatment[g$low_distance==0]) # 
0.9310556


# iii)

cace <- itt/prop.c # 0.239 --> quite close to 2SLS LATE using the continuous instrument


# iv) Calculate the p-value of the CACE
binary_instrument_model <- ivreg(score_asylum ~ treatment|low_distance, data = g)
summary(binary_instrument_model) # estimate: 0.239 (identical to when calculated 
manually), p = 4.69e-07
```

Question 2

View(sheriff)

Variables:
Treat = =1 if the Democratic candidate was elected, 0 otherwise
dem_vote_share = The Democratic candidate's share of the total votes
share_detained_sheriff = The share of ICE detention requests enforced by the Sheriff (from 0 to 1)
runnning_var = The democratic candidate's vote share, minus 50 --> basically be how much over majority the candidate won
year = year of observation
county_id = unique ID for each county

a) no code

b) Assessing sorting

Are there any discontinuities in the density of the running variable at the threshold?: McRary Sorting Test

```
pdf("density_plot.pdf")
DCdensity(sheriff$running_var,verbose=T)

title("McRary Sorting Test")

# add a legend for the y-axis
mtext("Density Estimate", side=2, line=3, cex.lab=1,las=0 )

# Add a legend for the x-axis
mtext("Candidate's Vote Share, -50", side = 1, line = 3)

dev.off()
```

There is a slight visual discontinuity but it is insignificant at any conventional level. We fail to reject the null hypothesis of no jumps in the density of the running variable at the threshold. There is no clear evidence of sorting.

c) Estimating the LATE with the optimal bandwidth

RDest <- RDestimate(share_detained_sheriff ~ running_var,data = sheriff)
summary(RDest)

# Note, this is the same as doing:
summary(RDestimate(share_detained_sheriff~running_var,cut-point= 0,data=sheriff))

The estimated LATE is -0.14979. It is not statistically significant at the 5% level (p = 0.09067)--> there is no strong evidence to support the presence of a treatment effect when considering the running variable.

d) Re-estimate the effects in question (c) using all whole-number bandwidths from 2 to 10. Use a figure to illustrate how sensitive your results are to these changes.

```
rdests <- rdci.up <- rdci.down <- c()
thresholds <- seq(from=2, to=10, by=1)

for(i in 1:length(thresholds)) {
  rdest <- RDestimate(share_detained_sheriff ~ running_var, bw = thresholds[i], data =
sheriff)
  rdests[i] <- rdest$est[1]
  rdci.up[i] <- rdests[i] + 1.96 * rdest$se[1]
  rdci.down[i] <- rdests[i] - 1.96 * rdest$se[1]
}

# pdf("Bandwidths.pdf")
plot(rdests, type = "l", lwd = 2, ylim = c(-0.5, 0.5), xaxt = "n", xlab = "Bandwidth", ylab =
"Estimate",  main=" Sensitivity of RD Result to the Choice of Bandwidth")
axis(1, at = c(2, 3, 4, 5, 6, 7, 8, 9, 10), labels = c(2, 3, 4, 5, 6, 7, 8, 9, 10))
abline(h = 0)
lines(rdci.up, lty = 3)
lines(rdci.down, lty = 3)
legend("topright", c("RD Estimate", "95% Confidence Interval"), lty = c(1, 3))
dev.off()
```

e) Fixed effects model for the effect of the election of a Democratic sheriff on enforcement of ICE detention requests

Clustering SEs by county
summary(feols(share_detained_sheriff ~ dem_vote_share | year + county_id, data= sheriff, cluster=~county_id)) # estimate: -0.002266, p=0.12894 > 0.05 --> not significant

```
# f)

# Investigate RD effects at placebo cut-offs within the optimal bandwidth --> CHECKING
THE FUNCTIONAL FORM

# Calculate the optimal bandwidth:
IKbandwidth(sheriff$running_var, sheriff$share_detained_sheriff) #the optimal bandwidth is
2.190

# Evalutate RD effects at placebo cutoffs within the optimal bandwidth
summary(RDestimate(share_detained_sheriff~running_var,cutpoint=-0.5,data=sheriff)) #
LATE = 0.014922, p = 0.8647
summary(RDestimate(share_detained_sheriff~running_var,cutpoint=-1,data=sheriff)) #
LATE = 0.06584, p = 0.3688
summary(RDestimate(share_detained_sheriff~running_var,cutpoint=-2,data=sheriff)) #
LATE = 0.02748, p = 0.6884
summary(RDestimate(share_detained_sheriff~running_var,cutpoint=0.5,data=sheriff)) #
LATE = 0.08911, p = 0.3311
summary(RDestimate(share_detained_sheriff~running_var,cutpoint= 1,data=sheriff)) #
LATE = 0.15070, p = 0.07364
summary(RDestimate(share_detained_sheriff~running_var,cutpoint= 2,data=sheriff)) #
LATE = -0.07928, p = 0.2195

# Within the optimal bandwidth, there is a slight discontinuity at 1, significant at the 10%
level.




# Illustration of the overall lack of discontinuity, within the bandwidth and at the cutoff - also
including the regression that we do when calculating RD at optimal bandwidth

# pdf("Within Bandwidth LM.pdf")
ggplot(sheriff, aes(x = running_var, y = share_detained_sheriff)) +
  geom_point(alpha = 0.5) +  # Plot the raw data points
  geom_vline(xintercept = 0, color = "red", linetype = "dashed") +  # Cutoff line at 0
  geom_smooth(data = subset(sheriff, running_var >= -2.19 & running_var <= 2.19), method
= "lm", se = FALSE, color = "green", aes(x = running_var, y = share_detained_sheriff)) +
  xlim(-2.19, 2.19) + # Limiting the x-axis to the optimal bandwidth
  theme_minimal() +
  labs(title = "RD Analysis: Raw Data and Fitted Lines within Optimal Bandwidth",
       x = "Running Variable",
       y = "Share Detained Sheriff")
dev.off()
```