Quant II Final Paper

Kelsi Quick

May 16, 2025

Question One

International cooperation is increasingly politicized, as reflected in both mass attitudes and elite behavior (De Vries et al. 2021). In demonstrations of opposition to such international institutions, country leaders have withdrawn their countries from these institutions, such as Trump's withdrawal of the United States from the World Health Organization (WHO) and Orbán's withdrawal of Hungary from the International Criminal Court (ICC) (Saifeddine, 2025; Tasch, 2025). While a global trend, the increasingly politicization of international cooperation is especially apparent in the European Union (EU) and is reflected in the rise of Eurosceptic attitudes in the public and the electoral success of such parties such as *Alternative für Deutschland (AfD)*, *Fidesz*, and *Fratelli d'Italia* (De Vries et al. 2021; Norris and Inglehart 2019; Szczerbiak and Taggart 2024). Intertwined with this growth of Euroscepticism and the politicization of EU integration, the issue of national sovereignty has become a more salient issue (Brack et al. 2021). Especially in the aftermath of repeated crises in the EU, concerns of national sovereignty have increasingly characterized politics in the EU (Bickerton et al. 2022; Kopeček 2019; Venizelos 2016).

While Eurosceptic individuals exist throughout the EU and Eurosceptic parties enjoy success in many of the member states, the object of this study is to examine whether individuals from countries with a post-communist legacy (PCL) evaluate issues of national sovereignty in a different way than their counterparts. Post-communist countries joined the EU under a different historical, political, and economic conditions than the founding or earlier Western member states, often after periods of externally imposed rule. These historical legacies may shape stronger attachments to national sovereignty and greater reluctance to delegate authority away from national governments. More specifically, this study seeks to answer whether individuals from countries with a post-communist legacy are more likely to perceive greater deficits in national sovereignty. I define the National Sovereignty Deficit (NSD), an original concept*, as the perceived gap between the ideal national sovereignty for a country and the national sovereignty that is observed. The benefit of such a measurement is that it offers insight into individual-level perceptions of sovereignty, which further contextualizes the increasing contestation and politicization of national sovereignty. If individuals from post-communist countries are more likely than their counterparts to perceive greater national sovereignty deficits, this would lend credence to the theory that individuals from countries with a post-communist legacy have differentiated experiences with the EU. If there are greater attachments to national sovereignty in post-communist countries, this finding would have important policy implications for the EU, such as expanding opportunities for differentiated integration or slowing the pace on deeper integration in order to prevent rising Euroscepticism and obstructionist approaches to EU decision-making.

*It should be noted that NSD is not introduced for the first time for this paper; it is an original, unpublished idea of mine I've been working on in various papers since 2023.

Question Two

The post-communist countries of the EU, which I will use interchangeably with the term Central and Eastern Europe (CEE)*, joined the European Union with drastically different circumstances than the founding and original 15 member states did. After decades of being under Soviet occupation or influence, the newly

independent states now faced the challenge of having to state-build, democratize, modernize, and nation-build all at once, creating the conditions of a difficult transition (Wolchik and Curry 2018). Though the accession to the EU initially enjoyed popular support among publics in CEE countries (Linden 2018), the asymmetry between the old states and the new grew more pronounced, leading to resentment and scepticism in the East (Wolchik and Curry 2018). Whereas the hope had been for democratization and modernization to enable the East to 'catch up' to the West, persistent economic and developmental inequalities remain, often causing Central and Eastern countries to fall into development traps, which can breed increased Euroscepticism over time (Csath 2023; Rodríguez-Pose et al. 2024; Venizelos 2017).

Post-communist countries already faced significant challenges to EU accession that their Western counterparts did not face, but these challenges were compounded by polycrisis in the EU. In particular, the Migrant Crisis, which began in 2015, has been a trigger for national sovereignty concerns, and especially in Central and Eastern Europe (Krastev 2017; Venizelos 2016). Venizelos (2016) points to the tendency of select CEE countries in the wake of the Migrant Crisis to characterize Brussels' policies as a threat to national sovereignty, taking on 'sovereigntist' roles for themselves. Krastev (2017) argues that the Migrant Crisis has reawakened the East/West divide in the EU with challenges that can be traced back to the post-communist transition. And because post-communist legacies can linger in national memory and nostalgia (Pettai and Pettai 2018), individuals in post-communist countries experience a fundamentally different experience of the EU than their Western counterparts. National sovereignty is a particularly salient issue due to recent polycrisis in the EU, but sovereignty is already complicated in the post-communist countries, as they only had a short time of independence after decades of oppression before surrendering key parts of their sovereignty to a different power center, the European Union (Linden 2018). These crises therefore serve as a catalyst that can invoke perceptions of EU misalignment with national interest, instilling a feeling of a loss of control over key issues such as national identity, culture, and self-determination. Due to the introduction of the NSD measure, this theory seeks to connect the context of historical legacies with perceived gaps between ideal and actual national sovereignty. Specifically, this study aims to test the theory that due to the unique post-communist political heritage of Central and Eastern Europe, individuals in these member states are more likely to perceive national sovereignty deficits. This leads to the following hypothesis:

H: Individuals in EU member states with a post-communist political heritage are more likely to perceive greater national sovereignty deficits.

*I use Central and Eastern Europe with the note that the Baltic states are also included in this categorization.

Question Three

To assess the theory that individuals in countries with post-communist legacies would be more likely to perceive greater national sovereignty deficits, I draw on data from the European Social Survey Round 10 (ESS10), conducted in 2020 (though due to Covid, some countries submitted responses as late as 2022). The ESS10 is a cross-national observational study of European countries, both inside and outside of the EU. The ESS data includes country codes for each respondent, which will aid in determining which respondents are from post-communist countries. Additionally, the rotating module of questions from ESS10 featured two questions related to national sovereignty, which I utilize to build a measure of national sovereignty, which is more detailed in my response to Question Five.

I then utilize a matched research design to better draw comparative conclusions between those from post-communist countries and those who are not. Since random assignment of individuals to countries with a post-communist legacy is impossible, matching on key covariates allows me to approximate an as-if randomized design, comparing individuals who are similar in their background characteristics. The 'treatment' in this case is whether or not individuals reside in post-communist states, indicated by a binary variable where 0 is non-post-communist countries (non PCL) and 1 is post-communist countries (PCL). I then match on several key covariates to account for confounding variables that may independently influence perceptions of national sovereignty deficits. For this study, I focus on 4 key covariates: age, gender, education, and political ideology (left-right self-placement). I chose these covariates because they capture core demographic and political predispositions that may influence perceptions of national sovereignty deficit but aren't caused

by living in a post-communist country. By matching on these covariates, I aim to ensure that the observed differences in national sovereignty deficits is due to the effect of living in a post-communist country rather than by differences in underlying demographics or ideological predispositions.

Question Four

The primary advantage of this research design in aiming to answer whether individuals residing in post-communist countries are more likely to perceive greater national sovereignty deficits is that matching allows for us to handle potentially confounding variables in a way that covariance adjustment does not. Covariance adjustment relies upon model assumptions, such as linearity in the context of a linear regression. For example, if we were to explore the same substantive question but use a covariance adjustment approach with linear regression, we would be assuming that the confounding variables influence the outcome (NSD) linearly and additively. By utilizing a matching design instead, we can avoid these assumptions. Instead, we create groups of comparable subjects *prior* to modeling the relationship between the independent variable (post-communist or not) and the outcome (NSD). This approach allows us to utilize elements of randomized, experimental studies, even in the context of an observational study such as this. We create an 'as-if randomized' design that allows us to observe, among individuals of similar age, gender, education and ideology, whether living in a post-communist country still has an effect on perceptions of national sovereignty deficits.

Despite its advantages, this research design still has limitations and disadvantages. First, even after matching, it is unlikely that we would remove all confounding from other factors, such as historical differences in the way communism was implemented in different countries, or the effects of different cultural and/or subnational identities. Furthermore, the cross-sectional data, given it is from only one year, cannot lend insight into temporal effects or show a causal process over time. This research design is therefore more associational rather than causal per se, but the associational claims we can make are still strengthened due to the as-if randomized matching design. Furthermore, the design may be limited in its generalizability, reflecting a certain time (the time of the survey) and the specific historical context of European post-communist countries. While this approach is helpful to understanding dynamics of national sovereignty in the EU, this finding may not be generalizable to other post-communist states, even the other post-soviet states in Europe but not part of the European Union.

Question Five

Independent Variable: The independent variable in this study is whether an individual respondent is residing in a post-communist EU member state or not. Countries are coded as post-communist if they were part of the Soviet Union, were a Soveit satellite state, or in the case of Slovenia, were part of the once-communist Yugoslavia. This variable is binary, where the respondent will be assigned a 0 if not from a post-communist country, or a 1 if they are from a post-communist country. Countries in this study and that are EU member states, have available data that are considered post-communist are: Bulgaria, Czechia, Estonia, Croatia, Hungary, Lithuania, Slovenia, and Slovakia. Countries in this study who are non-post-communist, have available data, and are EU member states are: Belgium, Finland, France, Greece, Ireland, Italy, the Netherlands, and Portugal.

Dependent Variable: The dependent variable in this study is perception of a national sovereignty deficit, or the gap between the ideal level of national sovereignty and observed national sovereignty. Utilizing two questions from the ESS10, I create a construct of National Sovereignty Deficit , wherein the respondent perceives a gap between what they consider the ideal level of national sovereignty for their country in the context of the EU and what they observe to 'actual' level of national sovereignty exercised by their national government.Two questions from the ESS10 ask specifically about the importance of national sovereignty. As can be seen in Question D12 (below), the interviewee is asked to think about the context of their own country and then evaluate how important they think it is for democracy (on a scale from 0-10) that key decisions are made by national governments rather than the European Union. While this question is using the language of democracy (invoking associations with legitimacy and popular sovereignty), it is ultimately

asking about the individual's preferred level of national sovereignty over key decisions. As such, I use this question as a measure for an individual's perception of ideal sovereignty. Question D24 works in the same way as the question for ideal sovereignty except that it instead asks the respondent to evaluate the extent to which the condition of key decisions being made by the national government rather than the EU applies to their own country. As such, I use this question as a measure of observed sovereignty to reflect the lived experiences of national sovereignty from the individual's perspective. The question wording follows:

Question D12: "And still thinking generally rather than about [country], how important do you think it is for democracy in general that key decisions are made by national governments rather than the European Union?"

Question D24: "To what extent do you think each of the following statements applies in [country]?... Key decisions in [country] are made by the national government rather than the European Union." Because these two questions reflect an individual's conception of ideal sovereignty and observed sovereignty, I created a measure to determine the perceived national sovereignty deficit wherein results from D24 (observed sovereignty) are subtracted from D12 (ideal sovereignty) to yield a 'deficit.' This is a constructed difference score; positive values indicate that a respondent perceives a deficit in national sovereignty, while values closer to 0 indicate the lack of a deficit.

Matched Covariates: I match on 4 key covariates: Age, Gender, Education, and Ideology. Age is measured in years. Gender is coded as a binary variable, where 1 is male and 2 is female. Education is measured by years of full-time education completed. I chose to use years of full-time education completed in order to make better, closer matches between individuals across different countries instead of using the categorical version of the variable. Ideology is measured on a left right scale from 0-10, where 0 is left and 10 is right (self-assessed).

Question Six

6.1 This study employs a matched observational design to isolate the effect of residing in post-communist countries on perceptions of a national sovereignty deficit. I utilize matching to create comparable sets of individuals that differ in their treatment; that is, whether they reside in a post-communist country or not. Given that individuals cannot be randomly assigned where they live, I use full matching with Mahalanobis distances to approximate an as-if random design. This approach creates groups similar in their key covariates (age, gender, education, and political ideology) to reduce the confounding effects of these variables on the relationship between post-communist context and national sovereignty deficits. I matched individuals exactly on gender and added calipers of 1.5 to age, education, and political ideology. This strategy created sets of individuals from post-communist and non-post-communist countries that were of the same gender, had within 1.5 years of education of each other, and were within 1.5 units on a left-right scale.

The identification strategy aims to demonstrate that what we observe is the effect of residing in a post-communist country on perceptions of a national sovereignty deficit. In order to make this argument, potentially confounding variables such as age, gender, education, and political ideology are used as covariates to match on; doing so creates comparable groups of individuals similar in their background characteristics, similar to experiments. The matching process therefore approximates random assignment and allows us greater leverage to make the claim that the relationship we observe is really the effect of residing in a post-communist country on national sovereignty deficit perceptions rather than through other variables/characterisitcs.

6.2 To demonstrate that this design creates interpretable comparisons, I conducted balance diagnositcs using both standardized mean differences and and formal hypothesis tests for covariate balance. A limitation to note here is that my matching design created approximately 5000 informative sets (sets that included both the treated and control groups), but standard tools like BalanceTest demonstrated an inability to compute balance for such a large sample size, despite various troubleshooting methods such as running each covariate independently. Therefore, I took a sample of 500 informative matched sets and ran a balance test on them. Results from the balance test demonstrate that after matching, age, gender and political ideology are well-balanced (standardized mean differences less than .1, and non-significant p-values). However, another limitation is that education remained somewhat imbalanced. For an ideal design, I would continue with

different matching procedures to create balance for education. But to focus on what we would ideally *want* to see, we would also want to education to have a SMD below .1 and a non-significant p-value. While some imbalance reamins for education, this limitation is mitigated by the strong overall balance of the matched sample, as demonstrated by the non-significant chi-sqaure test of joint covariate balance post-matching (p= 0.16). This suggests that, collectively, the covariates do not demonstrate significant imbalance after matching. The results of the balance test can be seen in the two figures below; Table X shows the covariate balance after matching and Table Y shows the overall balance test for the covariates collectively after matching.

Table 1: Table X. Covariate Balance After Matching (500 Informative Sets)

Covariate	Treatment.Mean	Control.Mean	SMD	z.Statistic	p.Value
Age	51.10	51.10	0.00	0.54	0.59
Gender (Female)	1.53	1.51	0.03	0.96	0.34
Education (Years)	13.00	13.40	-0.11	2.10	0.04 *
Ideology (LR Scale)	5.39	5.29	0.05	0.94	0.35

Table 2: Table Y. Overall Balance Test (Chi-Square)

Test	Chi.Square	df	p.Value
Unadjusted	5.53	4	0.237
Adjusted	6.66	4	0.155

Note. The overall balance test examines whether covariates jointly remain imbalanced after matching. Both tests indicate no significant joint imbalance (p > .05).

Question Seven

To handle missing data, I first recoded responses as missing following the direction of the ESS codebook. After recoding, I then assessed the extent of the missingness after filtering the dataset to create a new dataset of EU member states, which is the focus of this study. I found that missingness was minimal for most covariates, though highest for ideology. Out of a sample size of 29,918 respondents, about 4000 were missing ideology. I then looked at how many respondents would be left if I excluded anyone with NAs to see if matching would still be feasible. I found that after excluding all NAs, I would still have comparable sizes for post-communist countries and non-post-communist countries, with sample sizes of 11584 and 12355, respectively. Given this pre-matching balance, I chose to continue with the strategy of using listwise deletion given the original low or moderate missingness and then the ample and comparable sample sizes of the remaining data for matching.

I then checked for extreme data and outliers. I checked summary statistics as well as boxplot visualizations of the data distributions. Most variables fell within expected and realistic bounds, without outliers. Responses for 'Ideal Sovereignty' did have some outliers on the lower end, but as these responses still fall within the 0-10 scale and are interpretable and justified political opinions, I chose to retain them. Years of education, however, had some issues, with some respondents reporting more than 30 or as much as 65 years of education. Given that these are implausible and most likely coding issues or false data, I looked at how many responses fell above a relatively reasonable cut-off of 25 years, factoring in long PhD degrees. 95 responses out of the over 23,000 fell above 25 years, so I made the decision to drop these observations. In creating a clean dataset for matching, I want to emphasize reliable comparability and given the large sample size, I chose to drop the observations rather than winsorize.

Question Eight

To assess whether residing in a post-communist country is associated with higher levels of a national sovereignty deficit, I will utilize Ordinary Least Squares Regression (OLS) on my matched sample. The statistical tests I plan to use are embedded within OLS regression framework: I will construct a confidence interval as well as a hypothesis test on the regression coefficient to assess whether the null hypothesis of no effect (of residing in a post-communist country) can be rejected. Because the confidence interval and the p-value (from the hypothesis test) come from the same test, meaning that they are not independently testing the hypothesis but rather derived from the same estimation procedure, I will not correct for multiple testing. Adjusting for multiple testing, such as through the Bonferroni correction of False Discovery Rate, would be more appropriate for a design that conducts several repeated, independent hypothesis tests.

I chose to conduct a hypothesis test on the OLS regression coefficient because it allows us to assess whether an observed regression coefficient $\hat{\beta}$ is different from zero; that is, how extreme the observation of $\hat{\beta}$ would be under the assumption of *no* effect (that the null hypothesis is true). This is complemented by the confidence interval, which provides a range of plausible values for the true effect at a set significance level, usually 95% ($\alpha= 0.05$); in other words, if we were to repeat the confidence interval construction on different samples, then these intervals would contain the true value approximately 95% of the time. I will interpret the results of these tests using the conventional threshold of $\alpha= 0.05$ for the p-value. If the p-value is less than the significance threshold of 0.05, then I will conclude that there is sufficient evidence to reject the null hypothesis of no effect. However, should the p-value be greater than 0.05, then we fail to reject the null hypothesis. Similarly, if the confidence interval excludes 0, then I will infer that the true effect is unlikely to be zero. However, if the confidence interval includes 0, I will infer that the null hypothesis cannot be rejected. Beyond these criteria, I will also interpret substantive significance of the results; that is, whether the differences in national sovereignty deficits are meaningful and theoretically informative.

Question Nine

To judge the performance of the hypothesis test, I will examine its false positive rate and evaluate its power. Evaluating the hypothesis test's false positive rate will help determine the Type I error rate, or how often the test incorrectly rejects the null hypothesis (i.e. claiming a 'significance' of p<0.05 when no real effect is present). We do this by simulating data according to the null hypothesis, creating a null distribution where the real effect is that there is *no* effect of residing in a post-communist country on perceptions of a national sovereignty deficit. We then run the full analysis again, and the resulting proportion of simulations that register a significant effect (p-value > 0.05) is the estimated false positive rate. We then complement this by simulating data under a *known* treatment effect to determine the power of the test, determined by how often the test correctly rejects the null hypothesis of no effect. High power is therefore a sign of a small Type II error rate (the probability of failing to reject the null hypothesis when it is actually false). Because this research design is only interested in examining the outcome of one hypothesis test, I do not include family-wise error rate or false discovery rate. Such an approach is more appropriate for testing multiple hypotheses on the same dataset. For example, if I several sub-hypotheses examining the effect of residing in a post-communist country on EU trust, Euroscepticism, anti-immigrant attitudes, etc., then I would want to include family-wise error rate correction because with multiple testing, the probability of falsely finding a significant result increases. However, because I am only interested in the one hypothesis test here, examining false positive rate and power is sufficient.

Question Ten

False Positive Rate: To assess how my hypothesis test performs, I first evaluate its false positive rate. A false positive rate determines how often a test incorrectly rejects the null hypothesis (i.e. indicating a true effect when none exists; Type I Error). To do this, I performed a permutation-based simulation to estimate the false positive rate under the sharp null hypothesis of no effect. That is, I define a 'true' treatment effect of 0, that residing in a post-communist country has no effect on perceptions of a national sovereignty deficit.

This simulation design respected the existing matched structure of the data by permuting within matched sets, randomly re-assigning treatment within groups. This process simulates the absence of a treatment effect and allows us to see what we would observe if residing in a post-communist country were randomly assigned.

For each of the 1000 simulations, I re-estimated the treatment effect using the same OLS regression model, storing the p-values in the process. I then calculated the proportion of p-values below the conventional significance level of $\alpha = 0.05$. The result, listed below, was an estimated false positive rate of 25.8%, much higher than the 5% threshold. This suggests that my hypothesis test does *not* perform well and is prone to indicating a treatment effect when there is in fact not one. This high Type I Error indicates that my hypothesis test is too liberal (i.e. rejects the null hypothesis too often) under my current research design and is not a reliable test. Some of this may come from remaining covariate imbalance (which I noted earlier for education), but it could also come from issues with the OLS model I chose to analyze, such as not employing robust standard errors. Additionally, this issue could also arise from unmeasured confounding from variables or factors not included in the current research design. This assessment therefore leads me to conclude that the test is not very reliable under the current research design and demonstrates the importance of checking for false positive rate.

Estimated False Positive Rate (Type I Error): 0.258

Power: Next, I assess the power of the hypothesis test. Having high power for a test is desirable because it reduces the probability of Type II errors (failing to reject the null when there is actually a treatment effect). Similar to evaluating the false positive rate, we designate a *known* treatment effect for this procedure; however, this time we are looking at non-zero effects. I specify a series of effect sizes: 0 (which cross-validates my earlier simulation for FPR), 0.5, 1, 2, and 3. Substantively, this means I'm looking at various situations where the 'true' effect of treatment is that those residing in post-communist countries report NSDs 0.5, 1, 2, or 3 points higher than their counterparts. I follow the same procedure of permuting within matched sets, randomly assigning treatment. I then run 1000 simulations of the OLS regression on each effect size and store the p-values. I then examine the proportion of p-values below 0.05 for each effect size to determine the power. Figure 1 displays a curve demonstrating how power changes as the true treatment effect changes. What this figure shows is that the hypothesis test performs poorly and has no power when there is a true effect of 0; this reiterates the above explanation of false positive rate and the high Type 1 Error Rate this test has. However, the test performs better with moderate to large effects; power increases to 100% after the effect becomes 0.5.

Question Eleven

For this study, I use the Ordinary Least Squares (OLS) as the estimator to estimate the effect of residing in a post-communist country on perceptions of national sovereignty deficits. I chose OLS because the dependent variable, NSD, is measured as a continuous variable and OLS yields a coefficient that is interpretable as the average difference in NSD between individuals in the treated (post-communist) and non-treated (non-post-communist) groups. OLS in this case is estimating a difference-in-means between the treated and control groups. The target of estimation, the estimand, is the Average Treatment Effect (ATE) of residing in a post-communist country on perceptions of NSD. The ATE in this case is *not* the regression coefficient reported; the regression coefficient reported by the OLS estimator is only an estimate of the ATE. ATE, the estimand, refers to the theoretical, population-level quantity we are trying to learn about using the survey data as a sample. In this case, the estimand refers to the average difference in NSD perceptions between individuals residing in post-communist countries and those in non-post-communist countries across the entire EU population. Using the survey respondents as a sample of this population, we use an estimator (OLS) to find an estimate (a regression coefficient, showing difference-in-means of NSD between post-communist and non-post-communist) of the estimand, the average treatment effect of residing in a post-communist country on NSD in the population.

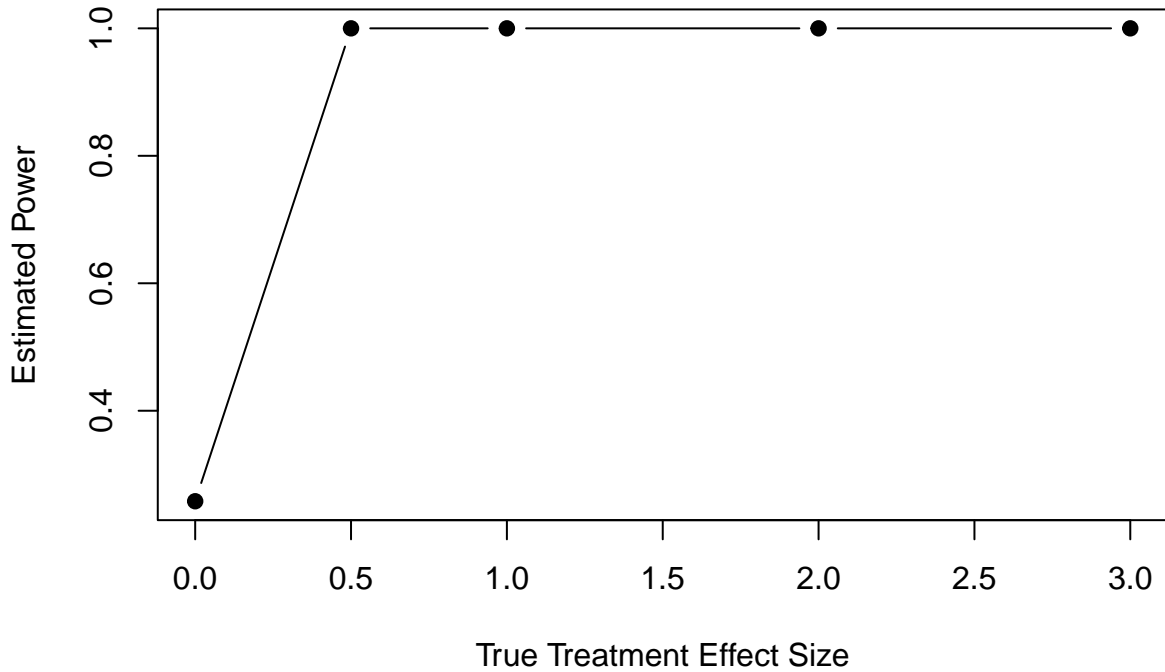## Figure 1: Power Curve for Hypothesis Test

Figure 1: Power Curve for Hypothesis Test

Question Twelve

A good estimator is consistent, precise, and unbiased, meaning it gets the correct value for the estimand on average (Aronow & Miller 2019). Estimators should be unbiased, so as to avoid systematic error in the guessing of the estimand, which in this case is the Average Treatment Effect (ATE) of residing in a post-communist country on perceptions of NSD in the population of individuals in the EU member states (Bowers, Voors and Ichino 2021). We also want our estimator to be consistent, meaning that as the number of units in the study increases, then the probability distribution of the estimator concentrates increasingly around the truth (Bowers and Leavitt 2020).

We can determine the performance of an estimator by observing its estimations of the treatment effect over repeated iterations. Similar to evaluating the performance of our hypothesis test, we again run a simulation study under a known treatment effect after permuting within matched sets to randomly assign treatment and perserve the matching design. In each simulation, I estimate the treatment effect using the OLS regression estimator. Instead of focusing on when the hypothesis test rejects or fails to reject the null hypothesis, we store the estimates of the treatment effect (effect of residing in a post-communist country on perceptions of NSDs). We then take the average of the ATE estimates and compare it to the known treatment effect to assess bias, or how often the estimator gets the treatment effect 'correct' or close to the 'truth.' We complement this by examining the variance of the estimate over the repeated simulations to assess its precision; we ideally want an estimator with low variance. Lastly, we also assess the Mean Squared Error (MSE), which measures both bias and variance, and allows us to see how good our estimator is.

Question Thirteen

To assess the performance of the OLS estimator, I ran a simulation study as described above to determine bias, variance, and the mean squared error (MSE). We ideally want estimators that have low bias, as this

indicates that the estimate is, on average, close to the 'true effect,' which was set to 1 in this case. This simulation study thus assumed that the true treatment effect of living in a post-communist country lead to a NSD 1 point higher, on average, compared to those not living in post-communist countries. The results of this simulation study indicate that the OLS estimator performs quite well under the known treatment effect of 1. The estimator had a bias of 0.06. meaning that the estimator was, on average, about 0.06 off from the true effect of 1. This demonstrates that the selected estimator has low bias and tends to be quite close to the true treatment effect.

In addition to low bias, the estimator also demonstrated low variance (0.002). This low variance indicates that the estimated treatment effect tended to be quite consistent across the repeated samples. Lastly, the estimator also demonstrated a low mean squared error (MSE) of 0.005. This suggests that the estimator performed well overall, with low bias and low variance. From this we can conclude that the OLS estimator performs well and tends to be accurate and consistent in its estimation of the Average Treatment Effect (ATE) of residing in a post-communist country on perceptions of a NSD. The estimator's quality performance in terms of unbiasedness, precision, and consistency makes it a credible and reliable choice for estimating the treatment effect of residing in a post-communist country on perceptions of a NSD. The estimator's good performance can also lead us to cautiously claim that we can draw some inferences about the population-level estimand. Because the estimator performed well in terms of low bias and consistency, we can conclude that the estimator, on average, produces estimates close to the true ATE.

Bias: 0.06047221 Variance: 0.001632393 MSE: 0.005287649

Question Fourteen

In the table below, I present results based on my real data as a mock analysis to illustrate the kind of results my final analysis may yield. I used Ordinary Least Squares (OLS) regression on my matched data to examine the effect of residing in a post-communist country on perceptions of a National Sovereignty Deficit. The coefficient for post-communist is 0.664, which means that, on average, individuals residing in post-communist countries in the EU report a NSD 0.664 points higher than their counterparts who do not live in post-communist countries. This is the estimated treatment effect of residing in a post communist country that remains after creating groups similar in background covariates such as age, gender, education, and political ideology, effectively 'controlling for' them through a matching design. Because the outcome variable of NSD is continuous with a possible range of values from -10 to 10, the regression coefficient reported here indicates a tendency for respondents from post-communist countries to report NSDs about 0.664 higher than their counterparts. That this coefficient is positive is in line with the theory that those in post-communist countries are likely to have *greater* perceptions of NSDs. While the difference is modest in magnitude, the observation that there is a systematic difference between the groups is substantively informative for the theory, and supports the hypothesis. That the p-value associated with this estimated effect is extremely small (p<0.001) indicates that, under the assumptions of the null hypothesis of no actual effect of residing in a post-communist country on perceptions of a NSD, observing such an estimate is very unlikely. This suggests strong evidence against the assumption of no effect and supports the conclusion that those residing in post-communist countries are more likely to perceive greater NSDs than their counterparts. However, the low r-squared value indicates that post-commmunist status only explains about 1% of the variation in NSDs, suggesting that there are other factors, quite possibly several other factors, that better explain this variation.

Additional Commentary: I also want to note how important it is to interpret these regression results within the context of the performance assessments of both the hypothesis test and the estimator. A simulation study found the estimator to be unbiased, consistent, and precise in its estimation of the Average Treatment Effect (ATE). The high-quality performance of the estimator leads me to conclude that the estimated treatment effect of 0.664 from this regression is likely an accurate and credible estimate. This conclusion is complemented by the assessment of the hypothesis test's high power for non-zero treatment effects, increasing our ability to safely reject the null hypothesis of no effect. However, the high power of the hypothesis test and unbiasedness of the estimator is undermined by the hypothesis test's high false positive rate. While the estimator is likely to correctly estimate and identify non-zero treatment effects, it remains

too prone to falsely rejecting the null hypothesis of no effect. This suggests a need for caution in interpreting results, especially p-values. It also suggests a need for further refinement of the research design, such as identifying and adding other covariates to match on, combined with improved covariate balance.

% Table created by stargazer v.5.2.3 by Marek Hlavac, Social Policy Institute. E-mail: marek.hlavac at gmail.com % Date and time: Thu, May 15, 2025 - 21:09:05

Table 3: OLS Regression Results

	Dependent variable:
	National Sovereignty Deficit
Post-Communist Country	0.664***
	(0.043)
Constant	2.011***
	(0.030)
Observations	23,844
R^2	0.010
Adjusted R^2	0.010
Residual Std. Error	1.891 (df = 23842)
F Statistic	234.094*** (df = 1; 23842)
Note:	*p<0.1; **p<0.05; ***p<0.01

Works Cited

Aronow, P. M., & Miller, B. T. (2019). *Foundations of Agnostic Statistics.* Cambridge: Cambridge University Press.

Bickerton, C., Brack, N., Coman, R., & Crespy, A. (2022). Conflicts of sovereignty in contemporary Europe: a framework of analysis. *Comparative European Politics, 20*(3), 257–274. https://doi.org/10.1057/s41295-022-00269-6

Bowers, J., & Leavitt, T. (2020). Causality and Design-Based Inference. In L. Curini, & R. Franzese (Eds.), *The SAGE Handbook of Research Methods in Political Science and International Relations* (pp. 769-804). SAGE Publishing. https://doi.org/10.4135/9781526486387.n44

Bowers, J., Voors, M., & Ichino, N. (2021). *The Theory and Practice of Field Experiments: An Introduction from the EGAP Learning Days.*

Csergö, Z. (2018). Ethnicity, Nationalism, and the Challenges of Democratic Consolidation. In S. L. Wolchik & J. L. Curry (Eds.), *Central and East European Politics: From Communism to Democracy.* Rowman & Littlefield.

De Vries, C. E., Hobolt, S. B., & Walter, S. (2021). Politicizing international cooperation: The mass public, political entrepreneurs, and political opportunity structures. *International Organization, 75*(2), 306-332.

Kopeček, M. (2019). Sovereignty, "Return to Europe" and Democratic Distrust in the East after 1989 in the Light of Brexit. *Contemporary European History, 28*(1), 73–76. https://doi.org/10.1017/S0960777318000851

Krastev, I. (2017). The Refugee Crisis and the Return of the East-West Divide in Europe. *Slavic Review, 76*(2), 291–296. https://doi.org/10.1017/slr.2017.77

Linden, R. H. (2018). The EU and Its New Members: Forging a New Relationship. In S. L. Wolchik & J. L. Curry (Eds.), *Central and East European Politics: From Communism to Democracy.* Rowman & Littlefield.

Norris, P., & Inglehart, R. (2019). *Cultural Backlash: Trump, Brexit, and Authoritarian Populism.* Cambridge University Press.

Pettai, V., & Pettai, E.-C. (2018). Transitional Justice and Memory. In S. L. Wolchik & J. L. Curry (Eds.), *Central and East European Politics: From Communism to Democracy.* Rowman & Littlefield.

Saifeddine, M. (2025, February 7). Trump leads wholesale withdrawals from international organisations. *Euronews.* https://www.euronews.com/2025/02/07/one-by-one-trump-leads-wholesale-withdrawals-from-international-organisations

Szczerbiak, A., & Taggart, P. (2024). Euroscepticism and anti-establishment parties in Europe. *Journal of European Integration, 46*(8), 1171–1191. https://doi.org/10.1080/07036337.2024.2329634

Tasch, B. (2025, April 3). Hungary withdraws from International Criminal Court during Netanyahu visit. *BBC.* https://www.bbc.com/news/articles/c807lm2003zo

Venizelos, E. (2016). Statehood and Sovereignty: The Difficult Equilibrium between European Union and Member States in Crisis Management - Refugee Crisis and Brexit. *European Politeia, 1*, 17–39. https://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=120822529&site=ehost-live

Venizelos, E. (2017). Democratic Legitimacy at National Level and Solidarity between Financially Unequal Member States. Two Structural Problems of European Integration. *European Politeia, 1/2*, 281–302.

Wolchik, S. L., & Curry, J. L. (2018). Democracy, the Market, and the Return to Europe. In S. L. Wolchik & J. L. Curry (Eds.), *Central and East European Politics: From Communism to Democracy.* Rowman & Littlefield.

Code Appendix

```r
knitr::opts_chunk$set(echo = TRUE)
knitr::opts_knit$set(root.dir = "/Users/kelsiquick/Downloads/QuantII/Final/ESS10-subset")
```

```r
library(readr)
library(optmatch)
library(RItools)
library(ggplot2)
library(cobalt)
library(survey)
library(dplyr)
library(knitr)
library(estimatr)
library(modelsummary)
library(stargazer)
```

```r
# Set your working directory

# Read in the dataset
essdata <- read.csv("ESS10-subset.csv")

# Preview the data
head(essdata)
str(essdata)
```

```r
# Save a clean copy of the original data
essdata_raw <- essdata

# Define the list of official EU member state codes (as of 2020)
eu_countries <- c("AT", "BE", "BG", "HR", "CY", "CZ", "DE", "DK", "EE",
                  "ES", "FI", "FR", "GR", "HU", "IE", "IT", "LT", "LU",
                  "LV", "MT", "NL", "PL", "PT", "RO", "SE", "SI", "SK")

# Filter the dataset to only include EU member states
essdata_eu <- essdata[essdata$cntry %in% eu_countries, ]

# Print which EU countries are actually present in the data
cat("EU countries present in ESS10 dataset:\n")
print(sort(unique(essdata_eu$cntry)))


#Creating post-communist binary variable
post_communist_countries <- c("BG", "CZ", "EE", "HR", "HU", "LT", "SI", "SK")

# Create the binary treatment variable: 1 = Post-Communist, 0 = Non-Post-Communist
essdata_eu$post_communist <- ifelse(essdata_eu$cntry %in% post_communist_countries, 1, 0)

# Quick check
table(essdata_eu$post_communist)

table(essdata_eu$cntry, essdata_eu$post_communist)


# Recoding missing data according to ESS codebook
# Recode 'lrscale' (Political Ideology)
essdata_eu$lrscale[essdata_eu$lrscale %in% c(77, 88, 99)] <- NA

# Recode 'keydec' (Ideal Sovereignty)
essdata_eu$keydec[essdata_eu$keydec %in% c(77, 88, 99)] <- NA

# Recode 'keydecc' (Observed Sovereignty)
essdata_eu$keydecc[essdata_eu$keydecc %in% c(77, 88, 99)] <- NA

# Recode 'gndr' (Gender)
essdata_eu$gndr[essdata_eu$gndr == 9] <- NA

# Recode 'agea' (Age)
essdata_eu$agea[essdata_eu$agea == 999] <- NA

# Recode 'eduyrs' (Years of Education)
essdata_eu$eduyrs[essdata_eu$eduyrs %in% c(77, 88, 99)] <- NA


# Check for missing data
colSums(is.na(essdata_eu[, c("lrscale", "keydec", "keydecc", "gndr", "agea", "eduyrs")]))


# Examine what the case would be if all NAs were dropped, and if sample size of PCL
# and non PCL would still be comparable
essdata_complete <- essdata_eu[complete.cases(essdata_eu[,
  c("lrscale", "keydec", "keydecc", "agea", "gndr", "eduyrs")]), ]
table(essdata_complete$post_communist)
```

```r
#Check for extremes/outliers
summary(essdata_complete$agea)
summary(essdata_complete$eduyrs)
summary(essdata_complete$keydec)    # Ideal sovereignty
summary(essdata_complete$keydecc)   # Observed sovereignty
summary(essdata_complete$lrscale)
summary(essdata_complete$gndr)


#Visualizaing main variable distributions
boxplot(essdata_complete$agea, main = "Age Distribution", ylab = "Age")
boxplot(essdata_complete$eduyrs, main = "Years of Education", ylab = "Years of Education")
boxplot(essdata_complete$keydec, main = "Ideal Sovereignty (keydec)", ylab = "Score 0-10")
boxplot(essdata_complete$keydecc, main = "Observed Sovereignty (keydecc)", ylab = "Score 0-10")
boxplot(essdata_complete$lrscale, main = "Left-Right Ideology (lrscale)", ylab = "Score 0-10")


#Looking more into education
table(essdata_complete$eduyrs > 40)
table(essdata_complete$eduyrs > 30)
table(essdata_complete$eduyrs > 25)


# Here I drop extreme education values (over 25 years)
essdata_complete <- essdata_complete[essdata_complete$eduyrs <= 25, ]
# Verifying
summary(essdata_complete$eduyrs)  # Check the new max value is 25


# Create the National Sovereignty Deficit (NSD) variable
essdata_complete$NSD <- essdata_complete$keydec - essdata_complete$keydecc

# Quick check of the NSD variable
summary(essdata_complete$NSD)


# view distribution of NSD (raw)
hist(essdata_complete$NSD,
     main = "Distribution of National Sovereignty Deficit (NSD)",
     xlab = "NSD (Ideal - Observed Sovereignty)",
     col = "lightblue",
     breaks = 20)


options("optmatch_max_problem_size" = Inf)

# Create Mahalanobis distance matrix
distance_matrix <- match_on(
  post_communist ~ agea + eduyrs + lrscale,
  data = essdata_complete,
  method = "mahalanobis"
)

# Exact matching on gender
exact_gender <- exactMatch(post_communist ~ gndr, data = essdata_complete)

# Apply a global caliper on the Mahalanobis distance
caliper_constraint <- caliper(distance_matrix, width = 1.5)  # Adjust as needed
```

```r
# Run Full Matching, combining constraints properly
full_match_result <- fullmatch(
  distance_matrix + exact_gender + caliper_constraint
)

#  Add matched group labels to dataset
essdata_complete$match_group <- full_match_result

# Quick check of matched sets
table(essdata_complete$match_group)
```

```r
length(unique(essdata_complete$match_group))

sum(is.na(essdata_complete$match_group))   # Number of unmatched respondents

table_sizes <- table(essdata_complete$match_group)
summary(table_sizes)   # Get min, median, mean, and max matched set sizes

hist(table_sizes,
     main = "Distribution of Matched Set Sizes",
     xlab = "Set Size",
     col = "lightblue",
     breaks = 20)
```

```r
# Identify Informative Matched Sets (those with both treated and control units)
informative_sets <- with(essdata_complete,
                          tapply(post_communist, match_group, function(x) length(unique(x)) > 1))

informative_ids <- names(informative_sets[informative_sets == TRUE])

# Subset the Data to Informative Sets
ess_informative <- subset(essdata_complete, match_group %in% informative_ids)

# Randomly Sample Informative Matched Sets
set.seed(123)   # For reproducibility
sampled_sets <- sample(informative_ids, size = 500)

# Create Final Sampled Dataset
ess_sample <- subset(ess_informative, match_group %in% sampled_sets)

# Run the Balance Test on the Sampled Data

balanceTest(
  post_communist ~ agea + gndr + eduyrs + lrscale + strata(match_group),
  data = ess_sample
)
```

```r
# balance table for Table X
# balance table with  p-values from balance test output
balance_table <- data.frame(
  Covariate = c("Age", "Gender (Female)", "Education (Years)", "Ideology (LR Scale)"),
  `Treatment Mean` = c(51.1, 1.53, 13.0, 5.39),
  `Control Mean` = c(51.1, 1.51, 13.4, 5.29),
```

```r
  `SMD` = c(0.00, 0.03, -0.11, 0.05),
  `z-Statistic` = c(0.54, 0.96, 2.10, 0.94),
  `p-Value` = c("0.59", "0.34", "0.04 *", "0.35")  # p-Value for education flagged as significant
)

# Display the table
kable(balance_table,
      caption = "Table X. Covariate Balance After Matching (500 Informative Sets)",
      digits = 2,
      align = "lccccc")

# creating the visual
# Create overall balance test table
overall_balance_table <- data.frame(
  Test = c("Unadjusted", "Adjusted"),
  `Chi-Square` = c(5.53, 6.66),
  `df` = c(4, 4),
  `p-Value` = c("0.237", "0.155")
)

kable(overall_balance_table,
      caption = "Table Y. Overall Balance Test (Chi-Square)",
      digits = 3,
      align = "lccc")

set.seed(12345) #for reproducibility

# Create weights from matched groups
match_counts <- table(essdata_complete$match_group, essdata_complete$post_communist)
treated_counts <- match_counts[, "1"]
control_counts <- match_counts[, "0"]

# Assign weights based on match group sizes
essdata_complete$weight <- with(essdata_complete, ifelse(
  post_communist == 1,
  1 / treated_counts[match_group],
  1 / control_counts[match_group]
))

# Function to compute p-value for each simulation (using OLS regression)
compute_p_value <- function(data) {
  model <- lm(NSD ~ post_communist, data = data, weights = weight)
  p_val <- summary(model)$coefficients["post_communist", "Pr(>|t|)"]
  #extracts p value for each estimate
  return(p_val)
}

# Permute treatment assignment within matched groups to preserve matching structure
# simulates null distribution under assumption of no effect
permute_within_groups <- function(data) {
  permuted_data <- data
  permuted_data$post_communist <- unlist(
    tapply(permuted_data$post_communist, permuted_data$match_group, sample)
```

```r
  )
  return(permuted_data)
}

# Run simulations to estimate false positive rate
n_simulations <- 1000
p_values <- replicate(n_simulations, {
  null_data <- permute_within_groups(essdata_complete)
  compute_p_value(null_data)
})

#  Calculate and report false positive rate at alpha = 0.05
false_positive_rate <- mean(p_values < 0.05)
#calculates how often the p-value was below 0.05,
# thereby counting how often the test registered
# a significant effect when there was none
cat("Estimated False Positive Rate (Type I Error):", false_positive_rate, "\n")

# Optional: Visualize p-value distribution
hist(p_values, breaks = 30, main = "P-Value Distribution Under Null", xlab = "P-Value")
```

```r
set.seed(12345)

# Define Power Simulation Function
power_function <- function(simulations, treatment, outcome, block, effect_size) {
  p_values <- replicate(simulations, {
    # Reshuffle treatment assignment within matched groups
    permuted_treatment <- unlist(tapply(treatment, block, sample))

    # Simulate new outcomes under the assumed true effect
    simulated_outcome <- outcome + permuted_treatment * effect_size

    # Run weighted OLS regression on simulated data
    model <- lm(simulated_outcome ~ permuted_treatment, weights = essdata_complete$weight)

    # Extract p-value for the treatment effect
    p_val <- summary(model)$coefficients["permuted_treatment", "Pr(>|t|)"]
    return(p_val)
  })

  return(p_values)
}

# Simulate Power for a Range of Effect Sizes
effect_sizes <- c(0, 0.5, 1, 2, 3)

power_estimates <- sapply(effect_sizes, function(effect) {
  p_vals <- power_function(
    simulations = 1000,
    treatment = essdata_complete$post_communist,
    outcome = essdata_complete$NSD,
    block = essdata_complete$match_group,
    effect_size = effect
```

```r
  )
  mean(p_vals < 0.05)  # Power is the proportion of p-values below 0.05
})

#  Plot Power Curve
plot(effect_sizes, power_estimates, type = "b", pch = 19,
     main = "Power Curve",
     xlab = "True Treatment Effect Size",
     ylab = "Estimated Power")
```

```r
power_plot <- plot(effect_sizes, power_estimates, type = "b", pch = 19,
     main = "Power Curve",
     xlab = "True Treatment Effect Size",
     ylab = "Estimated Power")
```

```r
set.seed(12345)

# Define True Treatment Effect
true_ATE <- 1.00


# Simulation Function to Estimate Treatment Effects
simulate_estimator <- function(data, true_effect) {
  # Shuffle treatment assignment within matched sets to simulate randomness
  permuted_treatment <- unlist(tapply(data$post_communist, data$match_group, sample))

  # Simulate outcomes under known true effect
  simulated_outcome <- data$NSD + permuted_treatment * true_effect

  # Estimate treatment effect using weighted OLS
  model <- lm(simulated_outcome ~ permuted_treatment, weights = data$weight)
  estimated_effect <- coef(model)["permuted_treatment"]

  return(estimated_effect)
}

# Run Simulations
n_simulations <- 1000
estimated_effects <- replicate(n_simulations, simulate_estimator(essdata_complete, true_ATE))

# Calculate Bias, Variance, and MSE
estimated_mean <- mean(estimated_effects)
bias <- abs(estimated_mean - true_ATE)
variance <- var(estimated_effects)
mse <- mean((estimated_effects - true_ATE)^2)

# Output Results
cat("Bias:", bias, "\n")
cat("Variance:", variance, "\n")
cat("MSE:", mse, "\n")
```

```r
# Calculate weights based on matched set sizes;
# I do this because with so many matched sets,
# R crashes when running a regression with matched groups as a factor

essdata_complete <- essdata_complete %>%
  group_by(match_group) %>%
  mutate(weight = 1 / n())

# Then run the weighted regression
reg_weighted <- lm(NSD ~ post_communist, data = essdata_complete, weights = weight)

# View results
summary(reg_weighted)
```

```r
stargazer(reg_weighted,
          type = "latex",
          title = "OLS Regression Results",
          dep.var.labels = "National Sovereignty Deficit",
          covariate.labels = c("Post-Communist Country"),
          no.space = TRUE,
          digits = 3)
```