Model Specification and Variable Coding across Tables

i. Model Specification

This section will examine the different model specifications from Table 1 and Table 3 in Scheve and Stasavage (2010), examining how changing Table specifications also alters our interpretations of Scheve and Stasavage (2010)'s hypothesis. Although both Table 1 and Table 3 address the same hypothesis, to test that hypothesis they examine different periods of time, with the former examining the impact of mass mobilization in World War I on tax progressivity between 1900 and 1930, while the latter expands that analysis to all mass mobilizing conflicts between 1850 and 1970. The two specifications employed are as follows, Table 1 uses specification (1), while Table 3 uses specification (2). Both are seen below.

$$toprate_{it} = \alpha + \beta WWI_{it} + \gamma X_{it} + \eta_i + \theta f(T_t) + \epsilon_{it} \quad (1)$$

$$toprate_{it} = \rho toprate_{it-1} + \alpha + \beta WarMobilization_{it} + \gamma X_{it} + \eta_i + \theta f(T_t) + \epsilon_{it} \quad (2)$$

where i is each country and t is each year; top rate is the relative top tax rate; WWI/War mobilization, is measure of war mobilization; X_{it} is a vector of control variables; $f(T_t)$ is a time factor, either year dummy variables or a linear trend. α, β, γ, and θ are given coefficients; η_i represents fixed effects parameters; and finally ε_{it} is the error term (Scheve and Stasavage, 2010).

Table 1 provides six total sub-specifications derived from (1), with new covariates included in each specification, the first being WW1 mobilization, then GDP per Capita, % of left seat share in the national legislature, Male Universal Suffrage, and Revenue to GDP. All the specifications employ Newey-West standard errors to account for the autocorrelation and heteroskedasticity in the model. The first three specifications include a linear trend and country-fixed effects to control for potential confounding factors and improve the model's explanatory power for across-time variation. Across the models, as the number of covariates increases, the number of observations decreases. The latter three specifications include the same covariates as the first three specifications but lack the linear trend, instead employing country and year-fixed effects to account for potential unobserved time-specific factors. However, the generalizability of Table 1 and its specifications are limited due to the significant autocorrelation and heteroskedasticity in the model, forcing the authors to use NW standard errors over Panel Corrected ones. Specifically, NW standard errors do not explicitly account for correlation among observations across different groups, meaning that although the authors get correct estimates, they do not explicitly account for correlation among observations across different groups. PCSE, conversely, corrects for both within-group correlation and between-group correlation and variation, all of which lead to more accurate estimates.

Similar to Table 1, Table 3 also provides six total sub-specifications derived from (2), with new covariates included in each specification, but with one additional covariate, a lagged version of the dependent variable top rate. In including a lagged variable, the authors control for potential confounding factors and improve their models explanatory power. The first three specifications include a linear trend and country-fixed effects. Notably, because the residuals are not autocorrelated in Table 3, the authors now utilize PCSEs, which is more efficient than the NW method, increases the accuracy of their estimated SEs, and allows the authors to reject their null hypothesis better. The latter three specifications include the same covariates but lack the linear trend of the first three specifications, instead using country and decade-fixed effects. The inclusion of decade fixed effects serves to bolster generalizability by controlling for time-invariant factors that might cause endogeneity concerns by controlling for unobserved heterogeneity across countries. In this regard, Table 3 introduces several notable refinements over Table 1, especially through PCSE and the inclusion of a lagged variable, both of which allow the authors to provide more empirically sound evidence against their null hypothesis. By employing a more advanced and rigorous methodology, Table 3 is instrumental in guiding our interpretations of Scheve and Stasavage (2010)s findings, ensuring a more accurate interpretation of the data, and enhancing the overall credibility of the paper's outcomes.

ii. Variable Coding

Beyond the differing specifications employed across the two Tables, both Tables also employ different variable coding. The main difference arises in Table 3, which includes a lagged version of the dependent variable and a recorded mass mobilization variable. Examining the lagged variable further, we see that it is included in the regression specification to capture the long-term trends in the tax rate and examine how the previous year's tax rate may serve to determine the current year's tax rate. In this regard, the inclusion of a lagged variable in Table 3 is significant, enabling the authors to examine the current tax rate as a part of an auto-regressive process whereby the current top tax rate is dependent on the previous year's top tax rates. Including this lagged variable also has a secondary benefit of capturing any autocorrelation between the observations, thereby helping to alleviate the main concern with Table 1, whilst and decreasing unobserved effects in the model. Overall, by employing a lagged variable, Scheve and Stasavage (2010) not only manage to provide a more comprehensive view of the drivers of tax during the research period but do so in a way that also decreases autocorrelation and decreases the number of unobserved effects not captured by the model.

Another difference in the variable coding comes from how the war mobilization variable is utilized. In Table 1, the war mobilization variable is used to identify the years before and after a country's participation in World War I, whereas in Table 3, it encompasses all mass warfare years, irrespective of the specific war. More specifically, in Table 1, the war mobilization variable is a binary variable coded to represent 0 in each year before the country entered the war and 1 thereafter. Conversely, in Table 3, the war mobilization variable is coded to be 1 in any year where there was mass mobilization for any war and 0 otherwise. The impact of this change is that by collapsing the variable into a single value in Table 3, we gain a greater understanding of how larger mobilizations affect income taxation. Through this new variable, Table 3 effectively generalizes the findings beyond World War I, demonstrating that larger mobilizations for war at any point in time, have an impact on taxation. Overall, the changes in variable coding between the two Tables advance our interpretations of Scheve and Stasavage (2010) 's paper by ensuring the robustness and generalizability of their findings to a broader time period and across multiple periods of mobilization.

iii. Distribution of the outcome variable

More broadly, across both Tables, there are a substantial number of zeroes in the outcome tax rate variable. In Table 1, using the filtered data between 1900 and 1930, we observe 79 zero values in our dependent variable 'top-rated tax' out of a potential 248 observations, meaning that 31.85% of the values are zero. This problem increases when we expand the date range to between 1850 and 1970, with approximately 401 zero values in our dependent variable, out of a possible 1248 or 32%. Once observations are dropped for N/A, this % stays the same . All of this indicates that in both Tables, a skewed distribution exists across variables, making estimating effect sizes accurately challenging, especially when the regressions conducted by Scheve and Stasavage (2010) rely on a normal distribution and linearity. More specifically, the large number of zeros negatively affects model fit, decreasing the precision in parameter estimates, and reducing statistical power, leading to potentially inaccurate conclusions being drawn from the regressions. To better account for the high number of zeros, Scheve and Stasavage (2010) could have adopted more flexible modeling assumptions, such as employing a non-monotonic specification and including a quadratic specification to allow for non-linear trends or employing zero-inflated models or hurdle models. However, the authors do not adopt any of these potential solutions, leading to significant concern surrounding the interpretation of the models, specifically the generalizability of the models to other settings and our relative trust in the models to produce a verifiable and accurate effect.

(b) Predicted Values and Residuals

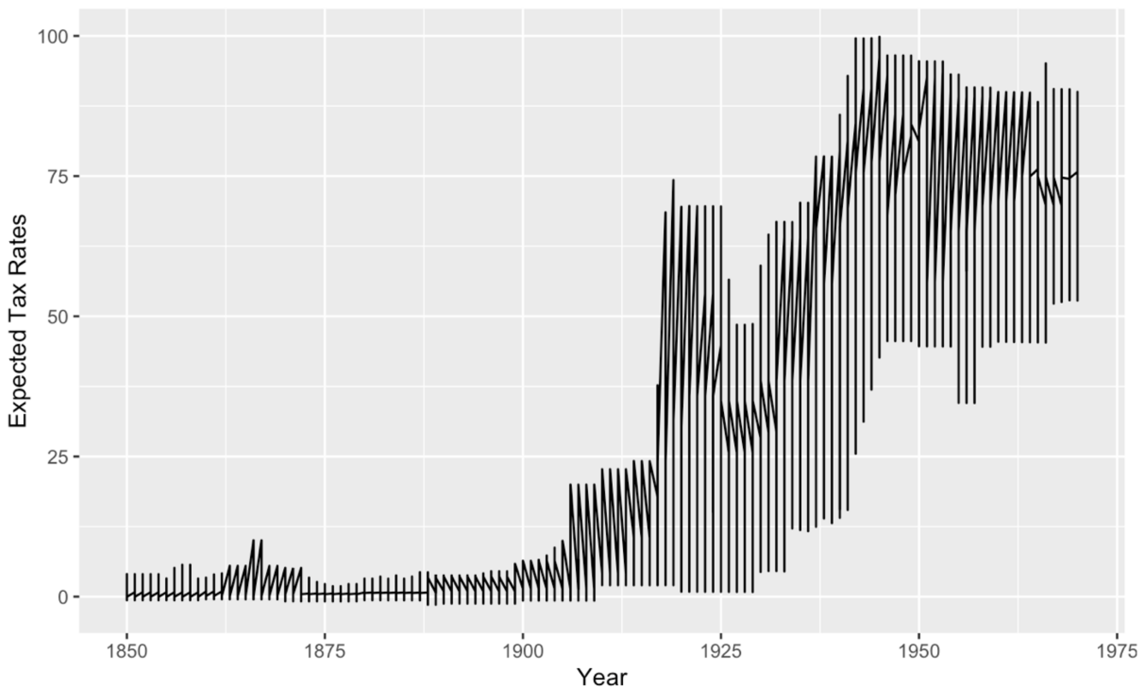
i. Predicted Values

For section b, I chose to examine the predicted values of Table 3 specification 5, which includes covariates, decade and country fixed effects, and estimates SE's using the PCSE method. However, because there are so many predicted values from the regression, specifically, one for each year and each country, I will examine the means of the expected tax rates across countries, which allows for more substantive interpretation of how mobilization impacted tax rates. The means are collated in the Figure (1) and Figure (2) displays the predicted demand for progressive taxation across time.

Figure (1) Predicted Values from Table (3) Specification (5)

ccode <S3: pseries>	expected_tax_rates <dbl>
2	33.13438
20	72.41765
200	37.38541
210	15.45785
220	22.93521
230	12.04959
380	20.10457
740	33.81584

Figure (2) Demand for progressive taxation across time (Predicted Values)



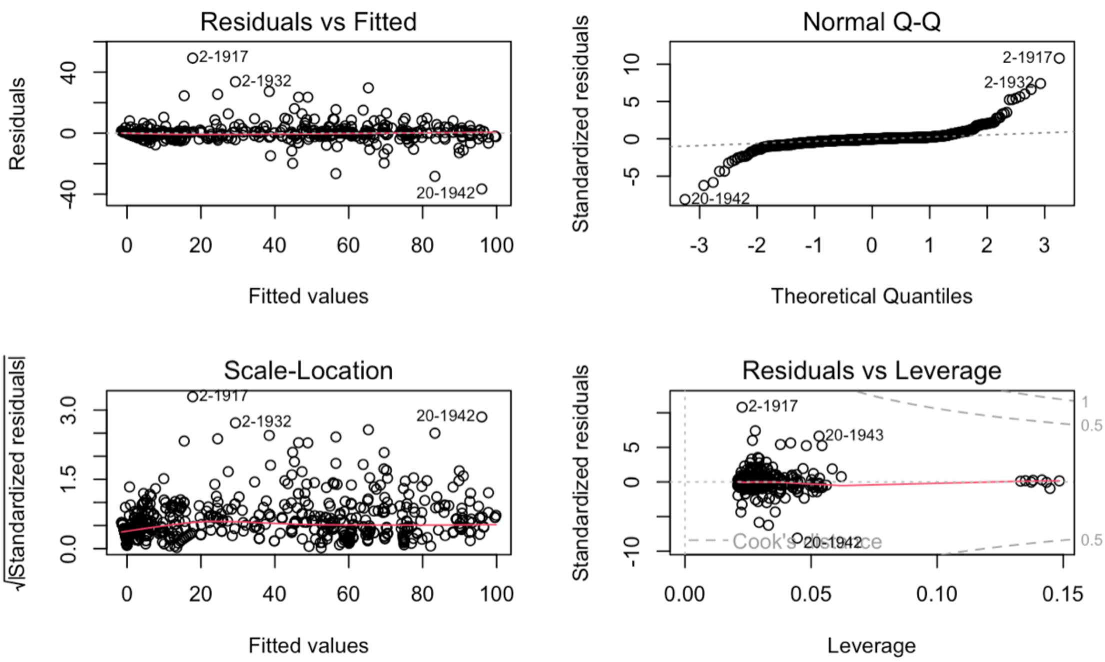
Examining Figure (1) above, we observe that the models' expected mean tax rates for each country are lower in nations that did not mobilize extensively for wars. Specifically, the USA (2), Canada (20), and UK (200), which mobilized extensively for war across 1850-1970, especially WW1, WW2, and Vietnam, have significantly higher mean tax rates than the other nations. Canada, in particular, has an exceptionally high expected top income tax rate of 72.417, almost double the second highest expected tax rate, that of the UK at 37.385. Conversely, in nations that mobilized less extensively across the period, either due to their occupation by other nations, or the war being fought within their territories, such as the Netherlands (210) and France(220), we see significantly decreased predicted tax rates under the model. Similarly, nations that were not invaded or neutral and therefore did not mobilize extensively also see decreased predicted tax rates under the model, with Spain(230) and Sweden(380) seeing some of the smallest predicted values. Finally, Japan (740), a nation that mobilized extensively for wars throughout the period of time being examined, especially WW1 and WW2, but whose mainland was never truly invaded, has a high predicted tax rate in the model. Overall, from this model, we observe a pattern consistent with the conclusions of Scheve and Stasavage (2010), that countries that extensively mobilized for wars exhibit higher top tax rates, while nations who had limited mobilization, either due to their occupation or neutrality, tending to have lower tax rates under the model.

More broadly, in Figure (2), we see that the model's estimated demand for progressive taxation across time, substantially increases during the main periods of wartime for nations across the sample. Beginning at 1860-1865, we see that the model predicted an increase in the top income tax around the period of the US Civil War. In the late 19th century and early 20th century, we see a small spike, likely due to the Boer Wars undertaken by Canada and the UK. During and after WWI, our model spikes significantly, following the trend seen in Scheve and Stasavage (2010)'s Figure (1) on page 540. We see a decrease in the early 1920s, then a large spike following the great depression and an even more pronounced spike that continues to increase as WWII begins in 1939. It's not until after the war, around 1950, that the models begins to predict decreases in the top income tax. We do see one spike in the mid-to-to late 1960s, potentially indicative of the beginning of the Vietnam war.

Notably, although Figure (2) does not break down country-by-country differences, these results do provide additional evidence in support of Scheve and Stasavage (2010)'s argument, that there exists an effect of mass mobilization as a result of warfare on top income tax rates from 1850 to 1970.

ii. Residuals

Figure (3) Residuals Graphs (1 - 4)



Examining the first residual graph above, being the residuals vs fitted graph, we see that there is a strong linear trend of residuals clumping evenly along the red line, between 0 and 100. We see that the residuals very slightly start more positive on the left, and get slightly more negative as we move further to the right, however, the effect is not pronounced. In this graph we also see three key outliers' points 2-1917, 2-1932, and 20-1942, which have large residual values, either extremely positive, around the 20-40 mark, or negative around the -40 mark. These outliers do not massively serve to skew the data however. Overall, across the residuals vs fitted graph, the spread of the residuals is relatively uniform across the x axis, with linearity holding well as evidenced by the red line closely approximating the dotted line.

Next, examining the Q-Q plot for normality of errors, we immediately observe residuals deviating at the top and bottom ends of the diagonal line. Specifically, the top of the Q-Q plot is pulled up by the outliers 2-1917, and 2-1932, which skew the data positively beyond what would be expected under standard modelling assumptions. Likewise, the tail of the residuals also has smaller values than would be expected, with 20-1942, serving as the most extreme negative tail value. All of these outlying residuals are further from the mean than what would be expected under a normal distribution. However, outside of these outliers, the residuals follows a relatively normal distribution, with the majority of points falling accurately along the reference line, indicating that residual normality is relatively strong within the data.

Moving onto the scale-location graph, we immediately see that the residuals roughly follow a normal distribution, being spread out across the red line, with some clumping at the 0 value. Specifically, we see a small dip in the average magnitude of the standard residuals around 0, and then another small dip around 60 and 100. This is reflective of the fact that about 1/3 of the data was zero values, as we have explored previously. We again see the outliers' 2-1917, 2-1932, and 20-1942, which have larger residual values than the rest. Beyond these outliers, the residuals are clearly not suffering from heteroskedasticity and autocorrelation, as evidenced by the relatively uniform clumping along the red line, with only minor dips. Overall, there are no identifiable issues in residuals within the scale-location graph.

Finally, examining the residuals vs leverage graph, we immediately see clumping around the 0 to 0.05 region. Notably, the graph has a highly linear trend, as evidenced by the red line closely approximating the dotted line, all of which indicates that the residuals are again linear, and not suffering from heteroskedasticity and autocorrelation. As we can see, the spread of the standardized residuals does change significantly as a function of leverage. Examining cooks distance, although we see our outliers these outliers do not fall outside of the dotted line, indicating that there are no major residuals that have a high level of influence over the model. Overall, there are no identifiable issues with the residuals vs leverage graph.

Notably, the outliers within the data, 2-1917, 2-1932, and 20-1942, represent instances where the model's predicted tax rates significantly deviate from the observed tax rates for these countries during those specific years. These outliers suggest that events during those years impacted the tax rates in the USA and Canada. These include the critical phase of increased WWI mobilization in the USA, and the peak of Canada's war efforts during WWII. However, the presence of these outlying residuals also somewhat rebuffs Scheve and Stasavage (2010) 's final assertion that "in the absence of mass warfare, there may have been nothing inevitable about the development of highly progressive tax systems." Rather, as we see, the year the USA began to recover from the great depression was a year in which the residual was extremely abnormal, indicating that perhaps alternative effects, such as extreme economic recessions, may have an effect on the development of highly progressive tax systems. Overall, the residuals tell a consistent story, with no larger issues surrounding heteroskedasticity or autocorrelation, all of which serves to buttress the findings of Scheve and Stasavage (2010), that mobilization was a key driver of demands for progressive taxation, resulting in notable deviations in the predicted tax rates for those specific years compared to the model trends.

(C) Conclusion

This paper set out with the task of replicating, analysing, and critiquing Scheve and Stasavage's (2010) paper on war mobilization and tax, and has done just that, evaluating their two main models, in the form of Table 1 and Table 3 respectively. After examining the predicted values of one of their regressions specifications, we find that it holds up well under scrutiny, with all the residuals behaving normally, and the key Gauss-Markov assumptions holding. Yet, this is not to say that the analysis does not have issues. Scheve and Stasavage's (2010) paper falls short in a number of respects, particularly surrounding the reliability and accuracy of the effects seen in Table 1, as well as a worrying number of zeros within the distribution of the outcome top tax rate variable that may serve to skew Scheve and Stasavage's (2010) results. Overall, this paper finds that Scheve and Stasavage (2010) correctly estimated the influences of war mobilization on taxation, with their estimated effects holding up well under reproduction, with their findings robust to alternative specifications and consistent with their data and research design.