Research Institute for the Evaluation of Public Polices

http://irvapp.fbk.eu

FBK-IRVAPP Working Paper series

Research Institute for the Evaluation of Public Policies

IRVAPP
RESEARCH INSTITUTE FOR
THE EVALUATION OF PUBLIC POLICIES
FONDAZIONE BRUNO KESSLER

Studying the Welfare State by Analysing Time-Series-Cross-Section Data

Federico Podestà

June 2023

FBK-IRVAPP Working Paper No. 2023-03

Studying the Welfare State by Analysing Time-Series-Cross-Section Data

Federico Podestà
FBK-IRVAPP
podesta@irvapp.it

FBK-IRVAPP Working Paper No. 2023-03

June 2023

Research Institute for the Evaluation of Public Policies
Bruno Kessler Foundation
Vicolo dalla Piccola 2, 38122 Trento (Italy)

Phone: (+39) 0461.314209
Fax: (+39) 0461.314240

E-mail: info@irvapp.it
Website: http://irvapp.fbk.eu

Studying the Welfare State by Analysing Time-Series-Cross-Section Data

Federico Podestà

Abstract

For a few decades now, quantitative researchers interested in studying welfare states have been analysing time-series-cross-section (TSCS) data relatively regularly. Given that welfare state researchers operate within an observational data framework, they seek to exploit the characteristics of TSCS data to make causal inferences. However, this objective remains quite difficult. Accordingly, the chapter aims to critically illustrate some of the most relevant TSCS techniques used in recent years. Much of the chapter regards TSCS regression, as it is the most widely used econometric tool for estimating causal effects regarding several welfare state features in a TSCS setting. The concluding part of the chapter regards the synthetic control method. This method requires a dedicated section because, although it has been widely used in numerous strands of research, it has arguably not yet been sufficiently exploited for the study of social policy.

Keywords: time-series-cross-section analysis; welfare state; causal inference; regression; synthetic control method.

Acknowledgments: For helpful comments on previous versions, I thank Enrico Rettore and Alessandro Tomdini. A revised version of this manuscript will be published in the volume entitled ‘Handbook of the Political Economy of Social Policy’ edited by Bent Greve, Amilcar Manuel Reis Moreira and Minna Van Gerven (Edward Elgar).

1. Introduction

For a few decades now, quantitative researchers interested in studying welfare state relatively regularly analyse time-series-cross-section (TSCS) data. TSCS data are characterized by repeated observations (often annual) on the same fixed (non-sampled) spatial units (usually states or countries). This means that arrays of data are those that combine cross-sectional data on N spatial units and T time periods to produce a data set of $N \times T$ observations. Here, the typical range of units of analysis would be about 20, if one focuses on developed countries, with each unit observed over a relatively long time period, around 20-50 years.

The increasing popularity of TSCS analysis is due to several reasons. Some of them have to do with the TSCS data structure, such as alleviating the traditional problem of cross-national comparison, i.e., the well-known imbalance between too many variables and too few (national) cases, and the obvious possibility of simultaneously examining cross-sectional and temporal variation.

Another reason for the growing popularity of TSCS analysis is due to the increasing availability of TSCS datasets. The most prominent comparative databases for social policy are probably the *Comparative Welfare States Data Set* (Brady et al. 2020), the *Social Citizenship Indicator Programme* (Korpi and Palme 2007) and the *Comparative Welfare Entitlements Dataset* (Scruggs et al. 2014).

Furthermore, in the wake of the recent "causal inference revolution", TSCS analysis is adopted to identify causal relationships concerning several welfare state topics. Unable to use controlled experimental designs – political institutions and social programs cannot be randomly assigned to different countries –, analysts remain in the observational data framework and try to exploit the characteristics of TSCS data to make causal inferences (Xu 2020). However, this objective remains quite difficult. Researchers normally encounter several problems and need to adopt specific solutions in order to reach reliable conclusions. Accordingly, the chapter aims to critically illustrate some of the most relevant TSCS techniques used over recent years.

Much of the chapter will be about TSCS regression (Section 2-4). This is essentially because it is the most widely used econometric tool for estimating causal effects in a TSCS setting. In this regard, it will first be highlighted that the conditions required to make causal inferences via TSCS regression, i.e., stationarity and exogeneity, are often violated in the study of the welfare state (see Section 2). That said, the illustration proceeds by introducing econometric techniques that one can adopt to address these problems. Concerning the non-stationarity issue, the methods of cointegration will be considered (see Section 3). On the other hand, the fixed effects model and the instrumental variable approach will be discussed regarding the endogeneity issue (see section 4).

The concluding part of the chapter (see Section 5) will regard the synthetic control method (SCM) that was initially introduced by Abadie and Gardeazabal (2003) and subsequently developed by Abadie et al. (2010; 2015). This method deserves a specific treatment for important reasons. First, allowing the effect of particular events occurred in single cases to be econometrically analysed, SCM represents a way to bridge the quantitative/qualitative divide in comparative research. Second, although SCM has been widely used in numerous strands of research, it has probably not yet been sufficiently exploited for the study of social policy.

Through the chapter, the discussion will not be purely technical. Statistical notation will be maintained to a minimum to make the argumentation accessible to a large audience. For sake of space, the discussion remains rather general without deepening the individual econometrical issues.

2. TSCS regression

Most quantitative welfare state scholars try to estimate the causal effect of an explanatory variable (x) on an outcome variable (y) by running a TSCS regression. Perhaps the most classic example is the ‘politics matter’ hypothesis according to which left power matters for welfare state development. To test this, social expenditure or some measure of welfare state generosity (y) is regressed against a quantitative indicator of left power (x) plus a set of control variables (e.g., unemployment rate, dependency ratio, trade openness, etc.).

A simple bivariate TSCS linear regression model estimable by the Ordinary Least Squares (OLS) procedure can be formulated as follows:

$$y_{it} = \beta_0 + \beta_1 x_{it} + e_{it} \quad (1)$$

where i = 1,2,….; N; refers to a cross-sectional unit; t = 1,2,….; T; refers to a time period. Thus, y_{it} and x_{it} refer respectively to dependent and independent variables for unit i and time t; and e_{it} is an error term and β_0 and β_1 refer, respectively, to the intercept and the slope parameter.

For TSCS regression to be used to estimate causal effects, two important conditions must be met: stationarity and exogeneity (e.g. Blackwell and Glynn 2018). Stationarity is a requirement first introduced in single time series analysis. Nonetheless, it must also be met in a TSCS setting. This is because the time dimension is particularly relevant in many TSCS datasets. The number of time points (e.g., annual observations) is large enough to use time-series econometrics (Beck and Katz 2004). Broadly speaking, stationary processes have mean and variance that do not change over time. This means that stochastic shocks do not have permanent effects so that the process is mean-reverting. This stationarity requirement implies that the time series are relatively stable over time so that inference about one part of the series can be projected out into the future’. In contrast, if a shock has a cumulative effect and the time series does not denote any inherent tendency to return to its equilibrium value, stationarity is prejudiced, and the process is said to be integrated.

The exogeneity assumption does not specifically regard the structure of TSCS data but concerns more generally the use of regression for reaching causal inferences. Specifically, among the regression assumptions, exogeneity is the most important for estimating a reliable causal relationship. This assumption rules out any correlation between x and the error term. Conversely, If the non-correlation assumption is violated, one or more of the variables put in the right-hand side of the equation are said to be endogenous. This impels inconsistent estimates and inappropriate inferences.

Unfortunately, these two conditions are often fulfilled in TSCS analysis on welfare state. The stationarity requirement is violated because several variables usually analyzed are non-stationary by construction or because they are highly persistent by their own dynamics. Public expenditure on social benefits is generally very smooth so that its current values can be well predicted by the past ones. This is essentially because budgets are not developed from scratch. Instead, policy makers typically begin with the last period’s budget and make incremental changes as deemed necessary (Durr 1993: 215). Left-wing party power is non-stationary by construction if it is measured – as it is done in the *Comparative Welfare States Data Set* – via cumulative score year by year of the share of seats in parliament held by leftist parties. This variable is seen to be continually increasing, denoting a non-constant mean (Kittel and Winner 2005). Furthermore, several econometric tests indicated that unemployment rate, dependency ratio, globalization indices and formal political institutions are nonstationary in a TSCS context (Podestà 2006; Chang and Lee 2010; Sobel and Coyne 2011).

Indeed, Beck and Katz (2011) argued that most of these variables cannot, by definition, be integrated. This is simply because they are bounded variables. For instance, social spending, being expressed as a percentage of GDP, is bounded between zero and one. Therefore, there would be no tendency for it to wander far from its mean and the variance of the observations would grow larger and larger over time. Nevertheless, some econometricians have demonstrated that boundedness is compatible with non-stationarity and developed the notion of bounded unit-root (Cavaliere 2005; Granger 2010).

Consequently, if one models high-persistent bounded variables as simply stationary, inferences can be threatened.
On the other hand, the reasons that generally make one suspect that explanatory variables are endogenous (i.e., omitted variables, measurement error and/or reverse causality) are common in TSCS studies of welfare state. For instance, to test the 'politics matter' hypothesis, one can run the above mentioned TSCS regression neglecting some unobservable factors, such 'enlightened leadership' (Esping-Andersen and Przeworski 2001). By so doing, omitted-variable bias can arise if the 'enlightened leadership' impacts on welfare state development and correlates with left-wing party power. This results in a correlation between the explanatory variable and the error term.

A similar issue arises if variables are measured with some degree of imprecision. It has been shown that state government partisan balance is often affected by measurement error (Klarner 2003). Left-wing party power can clearly suffer from this problem as well. Consequently, the measurement error in this variable becomes part of the error term, thus creating an endogeneity bias.

Lastly, some correlation between the regressor and the error term can arise in the case of reverse causality. This occurs if x and y dependent in relation to each other. This regards, for instance, the globalization-welfare state nexus. It is very possible that, in addition to an impact of globalization on welfare expenditure, the latter influences the former. Indeed, governments may keep the size of government intervention at a quite moderate level to attract foreign direct investments (Potrafke 2015).

The violation of the stationarity and the exogeneity assumptions prevents the estimation of a simple regression model such as that expressed by Equation 1. Moreover, including a set of control variables is not sufficient to solve these problems. TSCS analysts are then induced to adopt further remedies. Econometrics literature has developed specific techniques for analyzing non-stationary data and specific techniques to deal with endogeneity. In the next two sections, these techniques will be summarily illustrated. In doing so, it will be stressed, however, that the blanket is often too short. Adopting certain techniques to address the problems due to non-stationarity leaves the endogeneity issue. Similarly, by adopting techniques to deal with complications related to endogeneity, the problems associated with non-stationarity remain unaddressed.

3. Dealing with non-stationary data

As the requirement of stationarity, also the remedies developed for its violation derive from the econometrics for single-time series. They have been then extended to TSCS analysis (Birkel 2014). The following illustration paces this logic.

In a seminal article, Granger and Newbold (1974) showed that even if the usual t statistics indicate a significant relation, there is no sense in which dependent and independent variables are associated if the time series are non-stationary (i.e., the so-called spurious regression). In spite of this single time series issue, for a long time, scholars of welfare state opted for simple TSCS OLS regression in levels risking getting completely biased results. However, since the early 2000s, the tendency has changed and the TSCS remedies developed to avoid spurious regression were progressively – even if not systematically – adopted (Xu 2022).

The simplest solution to the spurious regressions problem is modelling the relation of interest, taking the first difference of the variables (i.e., $\Delta_1 Y_t = Y_t - Y_{t-1}$). This is simple because the differenced variables are usually stationary even if their (original) levels are not (for a TSCS application in social policy scholarship, see e.g., Kittel and Winner, 2005). The generic form of first-difference regression model for a TSCS setting is the following:

$$\Delta y_{it} = \beta_0 + \beta_1 \Delta x_{it} + e_{it} \quad (2)$$

Although the first-difference model can constitute a simple econometric solution, it involves a substantive drawback. Modeling differenced variables, it focuses on the short-term effects and discards any long-run relationship. Clearly, this also applies to the study of the welfare state. For instance, partisan effects need considerable time to materialize in welfare expenditure (Garrett and Mitchell, 2001: 168).

Accordingly, the econometric procedures which allow the capture of long-run relationships even in the presence of nonstationary processes have gained greater success. They are the two-step method proposed by Engle and Granger (1987) and the single equation error correction model (ECM) developed by, e.g., Banerjee et al. (1998). These procedures were then extended to the TSCS context by Pedroni (1999), Kao and Chiang (2001), and Westerlund (2007).

Both the two-step method and the single equation ECM are aimed to test if non-stationary variables are cointegrated. i.e., they have a long-term equilibrium relationship and all transitory deviations from this equilibrium state – due to possible shocks – are corrected in the long run.

The two-step method proceeds as follows. In the first step a static cointegrating regression of long-run level (non-stationary) variables is estimated. If the residuals of that regression exhibit stationarity, then variables are said to be cointegrated and one may proceed with the ECM to estimate the equilibrium rate and short-run dynamics.

On the other hand, the single-equation ECM is estimated in one step and takes the following form:

$$\Delta y_{it} = \beta_0 + \beta_1 y_{it-1} + \beta_2 x_{it-1} + \beta_3 \Delta x_{it} + e_{it} \quad (3)$$

where the parameter for the lagged dependent variable in levels (β_1) represents equilibrium properties. Specifically, y and x are cointegrated if $-1 < \beta_1 < 0$, while they are not cointegrated if $\beta_1 = 0$. The parameter β_2 for a lagged independent level variable measures the long-term effect of x on y. Lastly, the parameter for a change variable, β_3 represents the short-term impact of x on y (De Boef 2000).

Beyond these differences, the criteria to choose between these two approaches has to do with their different assumptions about the relationship between variables. In the two-step method, the variables are treated symmetrically, i.e., without a clear distinction between left- and right-hand side variables. This means that this approach does not assume causality; rather, it allows that variables are jointly endogenous. This is essentially because cointegrated variables are viewed as the simple manifestations of a same latent variable and an underlying mechanism is viewed as able to restore the equilibrium. Therefore, all disequilibria are restored by adjustments of both y and x (Beck 1992).

This approach could be, for instance, adopted to analyze the relationship between strong social democracy and large welfare states. As Esping-Andersen (2007) argued, it is very possible that they are jointly determined by some unidentified factor that, perhaps, lies deeply buried in history. Take Sweden: left power and welfare state may have moved together in the long run as a result of underlying national peculiarities.

However, this kind of argumentation is not so common in welfare-state scholarship. Students are usually concerned with unidirectional (causal) relationship among variables. The above-mentioned 'politics matter' hypothesis has been generally formulated assuming that left power affects welfare state development and not vice versa. In view of that, it is probable that one expects that social spending adjusts if it is out of equilibrium with government partisanship, but party power does not adjust to move into equilibrium with welfare expenditure (Podestà 2006). This is precisely what the single-equation ECM implies: equilibrium is exclusively restored through an adjustment of y as a result of a change in the independent variable x. For this reason, the single-equation ECM has experienced greater success than the two-step method in welfare state literature.

Nevertheless, this same reason involves a causal inference problem. If a unidirectional relationship between variables is presupposed, the single-equation ECM necessarily requires that the variables put in the right-hand side of the equation be exogenous. However, such an assumption is often even overlooked by both developers and users of this model. Consequently, biased inference can thus arise since different sources of endogeneity affect many macro-level analyses (Freeman 2016).

4. Dealing with endogeneity

4.1. Fixed effects model

Turning to the techniques developed to deal with endogeneity, one must first consider fixed effects (FE) model. It is the standard way for addressing the omitted-variables bias. It is intended to control for unobservable differences among observations based upon observable features. The basic idea is that in specifying the regression model we have failed to include relevant explanatory variables that do not change over time and/or others that do not change across countries. Accordingly, including a set of country and year dummies in the right-hand side of the equation model, we can cover up our ignorance and then solve the omitted-variable bias.

However, this clear advantage is accompanied by several drawbacks which often require the adoption of other empirical strategies or, at least, some caution in the use of FE model itself. Two of these drawbacks are directly related to the erroneous claim of making causal inference merely through the adoption of the FE model. First, the model is evidently incapable of resolving the omitted-variable bias if the bias is not only due to the omission of unobservable variables that do not change over time and/or space, but it is also due to the omission of unobservable variables that vary across country and do not remain constant over time. For instance, the unmeasured 'enlightened leadership' mentioned above could precisely vary across country and change, albeit slightly, over time within each of them.

Second, the FE model requires two additional assumptions to be used for causal inference: (1) past values of x do not directly influence current values of y (lagged effect) and (2) past values of y do not directly affect current values of x (reverse causality). Accordingly, linear FE regressions should be used with respect to the tradeoff between unobserved time-invariant confounders and dynamic causal relationships between x and y. In particular, if, as often happens in TSCS settings, the current values of x depend on the past values of y, one should primarily adjust for this rather than prioritizing the FE model for addressing the omitted-variable bias (Imai and Kim 2019; Hill et al. 2020).

Other drawbacks are less general and more strictly related to substantive issues. One concerns the fact that by including country dummy variables, one can capture the intra-unit variation only. This is because the inclusion of country dummies replaces the dependent and independent variables with their unit-centered deviations. In other words, the FE model removes any of the average unit-to-unit variation from the analysis, focusing on the within-country variation. This contrasts with most research objectives of comparative welfare state researchers, which are often aimed at estimating the effects of explanatory variables that precisely show much more variation across units than over time (Beck and Katz 2004).

The high persistency of numerous explanatory variables used in welfare state scholarship involves a parallel substantive drawback when one opts for the FE model. Country dummies are clearly collinear with any explanatory variables that are slowly changing. This implies that most of the explanatory power of those variables will be absorbed. Consequently, they are unlikely to emerge as either substantively or statistically significant.

Accordingly, Beck and Katz (2004) argue that TSCS analysts should assess (via an F test) whether FEs are needed in the model specification. If not, then there is no problem. If FEs are required, then researchers should make sure they are not losing the explanatory power of slowly changing or stable variables of interest. If variables of interest are being lost because of the inclusion of FEs, the researcher must weigh the gains from including FEs against their costs.

On the other hand, Plümper and Troeger (2007) developed the FE vector decomposition to separately analyze time-invariant and rarely changing variables. This estimator is de facto a three-stage procedure. First, an estimation of the unit FE by the baseline panel FE model excluding the time-invariant but not the rarely changing right-hand side variables. Second, a regression of the FE vector on the time-invariant and/or rarely changing explanatory variables of the original model (by OLS) to decompose the unit-specific effects into a part explained by the time-invariant variables and an unexplained part. And third, an estimation of a pooled OLS model by including all explanatory time-variant variables, the time-invariant variables, the rarely changing variables and the unexplained part of the fixed effects vector.

However, the issue of high-persistent Variables is not only related to the FE model, but also to modelling dynamics. In a more recent paper, Plümper and Troeger (2019) demonstrated that FE estimates can amplify the bias resulting from dynamic misspecification. Accordingly, researchers should choose the correct dynamic specification when they rely on FE estimates. However, a correct dynamic specification must be selected in relation to the properties of the variables included in the regression model even before solving the dilemma 'FE specification vs. non-FE specification'. In this regard, the nonstationary issue is probably paramount. As argued in Section 2, time-invariant and rarely changing variables cannot be usually considered to be stationary. Therefore, to select a correct dynamic specification, one should first test for this and proceed accordingly.

4.2. Instrumental variable approach

When exogeneity assumption is violated due to measurement error or/and reverse causality FE model is unhelpful and estimates are biased and inconsistent. In these circumstances, the regression model may be estimated with instrumental variable (IV) method. It is precisely a generic solution for all endogeneity problem.

As observed in Section 2, if y is regressed on x and x is endogenous, i.e., correlated with the error term, the parameter of interest cannot be consistently estimated. The IV approach overcomes this problem through a third (exogeneous) variable z which is, at the same time, correlated with x and non-correlated with the error term.

Basically, IV estimates can be computed via the so-called "two-stage least squares" (2SLS) procedure. In the first stage, the endogenous variable x is regressed on the instrument z plus all the other exogeneous (control) variables included in the equation of interest. In the second stage, y is regressed on the predicted values from the first stage plus the exogeneous variables. By so doing, the predicted x values can be viewed as the portion of x that is not correlated with the error term, i.e., the exogeneous portion. This is because there is no correlation between z and e and, hence, there will also be no correlation between the predicted x values and e.

Despite the simplicity of this identification strategy, the great challenge for researchers is to find a credible instrument. Indeed, a credible instrument must satisfy two conditions: (1) relevance, i.e., the instrument must be correlated with the endogenous explanatory variable, conditional on the other exogeneous variables (if this correlation is strict, the instrument is said to have a strong first stage; and (2) exclusion restriction: the instrument affects the dependent variable exclusively via its effect on the endogenous variable. Although the first condition can be directly assessed because both x and z are observable, the second is not testable. Consequently, the exclusion restriction assumption is normally supported by providing a convincing narrative based on theoretical and contextual information (Sovey and Green 2011; Becker 2016)

The IV approach is increasingly used in comparative research as well as in the welfare state scholarship. As observed in Section 2, the globalization-welfare state nexus may be affected by reverse causality. For many years, scholars have regressed social expenditure on globalization, ignoring that issue. However, more recently, several analysts have addressed this source of endogeneity via the IV approach (Potrafke 2015). For example, Santos and Simões (2021) use an instrument for globalization based on the assumption that the levels of globalization of a given OECD country are influenced by the levels of globalization of the neighboring countries. To fulfill the exclusion restriction assumption, these authors argue that there is no theoretical or empirical argument linking geographical distance (exogenous) and average globalization levels in the OECD with the decisions about social expenditure of each individual government. For instance, the level of globalization in the OECD as a whole and the distance of Austria from other OECD countries should not influence the behavior of social expenditure in Austria except if they lead to changes in Austria´s own globalization levels, which in turn pressures the Austrian government to change social expenditure.

Finding a credible instrument is not the only problem with the IV approach. Also in this case, some problems may arise if one fails to properly address the time series properties underlying the variables included in the regression model. More precisely, both bias and mistaken inference can arise if, as frequently may occur, both the endogenous regressor and the instrument are not stationary. If so, their estimated association in the first stage coincides with a spurious regression. Furthermore, in the special case of reverse causality between the outcome variable and an endogenous regressor, the resulting cointegration of these two variables introduces bias and inconsistency into the IV estimate that can reinforce rather than resolve the identification problem present in OLS estimation (Christian and Barrett 2021).

5. The synthetic control method

Consistently with the quantitative tradition, TSCS regression seeks to estimate the average effect of one or more causes across a population of cases. As said, the TSCS regression is, for example, adopted to estimate the causal effect of partisanship on welfare state development within a set of countries observed for a given period of time. On the other hand, a core goal of qualitative research is the explanation of outcomes in individual cases. Generally, qualitative researchers start with one or more events occurred in certain cases and then move backward toward the causes. For example, they attempt to identify the causes of the creation of especially generous welfare states (Mahoney and Goertz 2006). SCM straddles the two traditions. Like qualitative methods, it focuses on specific events occurred in certain cases. However, it does not attempt to identify their causes. As a quantitative method, it adopts an effects-of-causes approach. It seeks to estimate the impact of those events on certain outcomes.

To better clarify the logic of SCM and how it can be triangulated with qualitative analysis on the welfare state, Podestà's (2020) SCM replication of Pierson's (1994) study of Reagan and Thatcher's welfare state retrenchments is hereafter considered.

Pierson tried to explain why at the end of the so-called conservative resurgence of the 1980s, when Thatcher and Reagan left power, the UK and US welfare states ended up substantially unaltered. Tracing and contrasting the decision making of several social programs, he concluded that the structure of social programs prevented the prospects of dismantling the welfare state. Social programs have become central features of the political landscape, and with them have come dense networks of interest groups and strong popular attachments. However, Pierson's qualitative analysis was not substantiated by any counterfactual indication, i.e., what would have happened in the absence of conservative governments. To fill this gap, Podestà (2020) performed a synthetic control analysis.

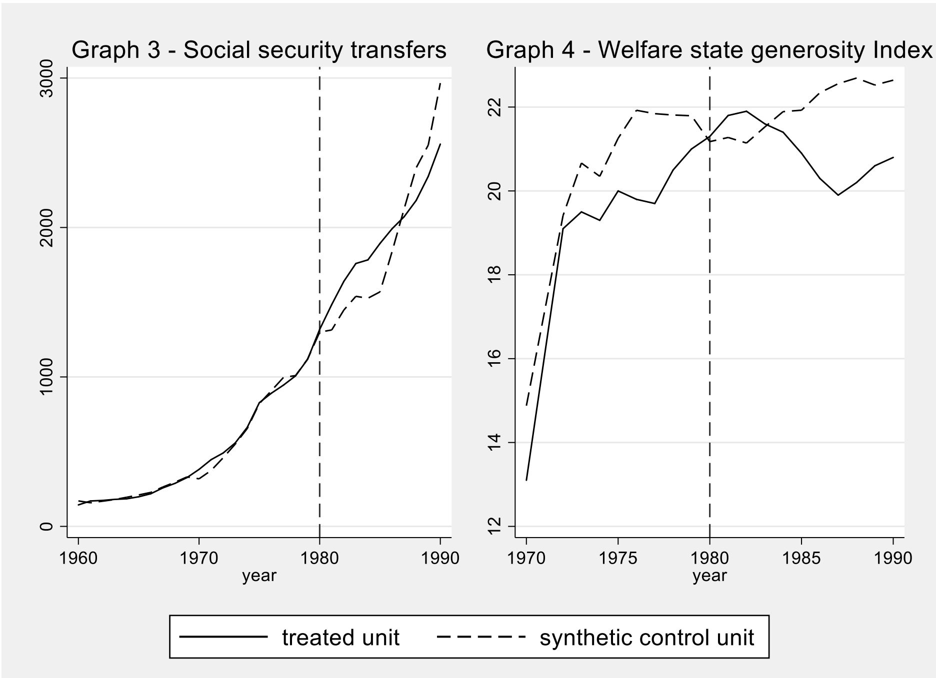
A TSCS dataset composed of 17 OECD countries annually observed for the 1960-1990 period was employed. The USA and UK were the treated units, the countries which experienced the event of interest, i.e., the conservative resurgence); while the other 15 OECD countries included in the dataset formed the donor pool, i.e., the set of potential comparisons. The treatment period, i.e., the period covered by Reagan and Thatcher's administrations, approximately lasted from 1980 to 1990, while the pre-treatment period lasted from 1960 to 1979.

To estimate the impact of Reagan and Thatcher's administrations, the trajectory of the UK and US welfare states in the presence and in the absence of conservative governments were to be compared. However, since the welfare state trajectory in the absence of Reagan's and Thatcher's administrations was not observable, it has been synthetically reproduced. Specifically, two synthetic units, one for the

UK and one for the US were constructed. The reliability of these two synthetic units depended on their ability to best reproduce the pre-treatment trajectories of the two welfare states and the values of their predictors. To this end, the SCM provides a data-driven procedure to choose a weighted combination of comparison units that fits the actual unit better than it does any single comparison unit itself (Abadie et al., 2010; 2015). Hence, the synthetic UK/USA were constructed as a weighted average of those 15 OECD countries that best reproduced the UK/USA outcomes and their most relevant predictors prior to 1980.

Clearly, the use of SCM involves some data requirements. As for TSCS regression, one of these requirements regards the temporal dynamics of the variables of interest. Specifically, the trajectory of the outcome variable must be quite smooth. This in order to be best reproduced synthetically in the pre-treatment period. Some of the synthetic control analyses performed by Podestà (2020) constitute good examples of this issue. For sake of space, only two outcome variables for the United States are considered here. Since social security transfers per capita is remarkably smoothed, the synthetic line almost perfectly overlaps the trajectory of the treated line in the pre-1980 period. Conversely, this is not the case for welfare state generosity Index. This is precisely because this outcome variable denotes a noticeably irregular trend (see Figure 1). Therefore, when the difference between the values of the treated unit and those of synthetic one is large, the use of SCM is not recommended because of the potential for substantial biases (Abadie et al. 2010; Abadie 2021).

Figure 1 – Social security transfers per capita and welfare state generosity Index for the USA: the treated unit vs. the synthetic unit



References

Abadie, A. (2021). Using synthetic controls: Feasibility, data requirements, and methodological aspects. *Journal of Economic Literature*, 59(2), 391-425.

Abadie, A., & Gardeazabal, J. (2003). The economic costs of conflict: A case study of the Basque Country. *American economic review*, 93(1), 113-132.

Abadie, A., Diamond, A., & Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. *Journal of the American statistical Association*, 105(490), 493-505.

Abadie, A., Diamond, A., & Hainmueller, J. (2015). Comparative politics and the synthetic control method. *American Journal of Political Science*, 59(2), 495-510.

Banerjee, A., Dolado, J. J., & Mestre, R. (1998). Error‐correction mechanism tests for cointegration in a single‐equation framework. *Journal of Time Series Analysis*, 19(3), 267-283.

Beck, N. (1992). The methodology of cointegration. *Political analysis*, 4, 237-247.

Beck, N. (1993). The Methodology of Cointegration. *Political Analysis* 4: 237-47.

Beck, N. (2011). Time-series-cross-section data: what have we learned in the past few years?. *Annual Review of Political Science*, *4*(1), 271-93.

Beck, N., & Katz, J. N. (2011). Modeling dynamics in time-series–cross-section political economy data. *Annual Review of Political Science*, *14*, 331-352.

Becker, S. O. (2016). Using instrumental variables to establish causality. *IZA World of Labor*, (250).

Birkel, C. (2014). The analysis of non-stationary pooled time-series cross-section-data. *International Journal of Conflict and Violence (IJCV)*, *8*(2), 222-242.

Blackwell, M., & Glynn, A. N. (2018). How to make causal inferences with time-series cross-sectional data under selection on observables. *American Political Science Review*, *112*(4), 1067-1082.

Brady, B., Huber, E., & Stephens, J. D. (2020). *Comparative Welfare States Data Set*, University of North Carolina and WZB Berlin Social Science Center, 2020.

Cavaliere, G. (2005). Limited time series with a unit root. *Econometric Theory*, 21(5), 907-945.

Chang, C.-P., & Lee, C.-C. (2010). Globalization and growth: A political economy analysis for OECD Countries. *Global Economic Review* 39(2), 151-173.

Christian, P., & Barrett, C. B. (2021). Spurious regressions and panel IV estimation: revisiting the Causes of Conflict. Available at SSRN 3950679.

De Boef, S. (2000). Modeling equilibrium relationships: error correction models with strongly autoregressive data. *Political Analysis,* 9, 14-48.

Durr, R. (1993). An essay on cointegration and error correction models. *Political Analysis,* 4, 185-228.

Engle, R.F. & Granger C.W.J. (1987). Cointegration and error correction: representation, estimation and testing. *Econometrica,* 55, 251-276.

Esping-Andersen, G. (2007). Multiple regression in small-N comparisons. *Comparative Social Research*, 24, 335–342.

Esping-Andersen, G., & Przeworski, A. (2001). Quantitative cross-national research methods. *International encyclopedia of the social and behavioral sciences*, *18*, 12649-12655.

Freeman, J. R. (2016). Progress in the study of nonstationary political time Series: A comment. *Political Analysis*, *24*(1), 50-58.

Garrett, G. & Mitchell, D. (2001). Globalization, government spending and taxation in the OECD. *European Journal of Political Research*, 39, 145-77.

Granger, C.W.J. & Newbold, P (1974) Spurious regressions in econometrics. *Journal of Econometrics* 2: 111-120.

Granger, C.W.J. (2010). Some thoughts on the development of cointegration. *Journal of Econometrics* 158: 3–6.

Hill, T. D., Davis, A. P., Roos, J. M., & French, M. T. (2020). Limitations of fixed-effects models for panel data. *Sociological Perspectives*, 63(3), 357-369.

Imai, K., & Kim, I. S. (2019). When should we use unit fixed effects regression models for causal inference with longitudinal data?. *American Journal of Political Science*, *63*(2), 467-490.

Kao, C., & Chiang, M. H. (2001). On the estimation and inference of a cointegrated regression in panel data. In Nonstationary panels, panel cointegration, and dynamic panels (Vol. 15, pp. 179-222). Emerald Group Publishing Limited.

Klarner, C. (2003). The measurement of the partisan balance of state government. *State Politics & Policy Quarterly*, *3*(3), 309-319.

Kittel, B., & Winner, H. (2005). How reliable is pooled analysis in political economy? The globalization-welfare state nexus revisited. *European Journal of Political Research*, *44*(2), 269-293.

Korpi, W., & Palme, J. (2007). The Social Citizenship Indicator Program (SCIP). Retrieved from http://www.sofi.su.se/spin/about-the-project/social-citizenship-indicator-program-scip-1930-2005-1. 202043

Mahoney, J., & Goertz, G. (2006). A tale of two cultures: Contrasting quantitative and qualitative research. *Political analysis*, 227-249.

Pedroni, P. (1999). Critical values for cointegration tests in heterogeneous panels with multiple regressors. *Oxford Bulletin of Economics and statistics*, *61*(S1), 653-670.

Plümper, T., & Troeger, V. E. (2007). Efficient estimation of time-invariant and rarely changing variables in finite sample panel analyses with unit fixed effects. *Political analysis*, 15(2), 124-139.

Plümper, T., & Troeger, V. E. (2019). Not so harmless after all: The fixed-effects model. *Political Analysis*, 27(1), 21-45.

Pierson, P. (1994). *Dismantling the welfare state?: Reagan, Thatcher and the politics of retrenchment*. Cambridge University Press.

Podestà, F. (2006). Comparing time series cross-section model specifications: the case of welfare state development. *Quality and Quantity*, *40*(4), 539-559.

Podestà, F. (2020). Did the conservative resurgence matter? A counterfactual replication of Pierson's analysis on Reagan and Thatcher's welfare state retrenchments. *Comparative European Politics*, 18, 819-839.

Potrafke, N. (2015). The evidence on globalization. *The World Economy*, 38(3), 509-552.

Santos, M., & Simões, M. (2021). Globalisation, welfare models and social expenditure in OECD countries. *Open Economies Review*, *32*(5), 1063-1088.

Scruggs, L., Jahn, D., & Kuitto, K. (2014). *Comparative Welfare Entitlements Dataset 2* (Version 2014-03).

Sobel, R. S., & Coyne, C. J. (2011). Cointegrating institutions: The time-series properties of country institutional measures. *The Journal of Law and Economics*, 54(1), 111-134.

Sovey, A. J., & Green, D. P. (2011). Instrumental variables estimation in political science: A readers' guide. *American Journal of Political Science*, *55*(1), 188-200.

Westerlund, J. (2007). Testing for error correction in panel data. *Oxford Bulletin of Economics and Statistics*, 69, 709-748.

Xu, Y. (2020) Introduction to the virtual issue: Panel data analysis and regression discontinuity. *Political Analysis*, published online. Available at: https://www.cambridge.org/core/journals/political-analysis/special-collections/panel-data-analysis-and-regression-discontinuity

Xu, Y. (2023). Causal inference with time-series cross-sectional data: a reflection. *Available at SSRN 3979613*.