European Political Science Review (2025), 1–17
doi:10.1017/S175577392510012X

RESEARCH ARTICLE

Contrasting and combining the historical counterfactual method and the synthetic control method

Federico Podestà[1] and Joan E. Madia[2]

[1]Research Institute for the Evaluation of Public Policies (IRVAPP), Bruno Kessler Foundation, Trento, Italy and [2]Kellogg College, University of Oxford, Oxford, UK
Corresponding author: Federico Podestà; Email: podesta@irvapp.it

(Received 13 January 2025; revised 22 June 2025; accepted 24 September 2025)

Abstract
In recent years, many efforts have been made to bring quantitative and qualitative methods into dialogue. This article also moves in that direction. However, in contrast to most works, the present attempt does not concern the large-N/small-N issue but focuses instead on the sole single case study framework. Within this framework, two counterfactual methods, the historical counterfactual method, the qualitative one, and the synthetic control method, the quantitative one, have gained great importance without however meeting. This paper aims to advance mixed-methods research by bridging the gap between these two approaches. More precisely, it has assessed whether these two methods can be used together to understand what would have happened in a single case Z in the absence of an event X. The case study of the impact of Thatcher's election on the UK pension system is then presented as an example of the joint use of the two methods.

Keywords: Synthetic control method; historical counterfactual method; single-case study; causal inference; mixed methods

Introduction

In recent decades, the counterfactual model of causation has gained increasing importance in political science and more generally in the social sciences. Usually, one wonders whether event X caused event Y. The counterfactual model proposes to answer this question assuming that if the antecedent X had not occurred (or had occurred differently), the outcome Y would not have occurred (or have occurred differently).

A main problem with the counterfactual approach is that the statement "In case Z, if X were not to occur, then Y would not occur either" cannot be evaluated empirically. Indeed, the antecedent X actually occurs in the real world so that the premise is false, and there is no evidence about what would happen if X did not occur. In the potential outcome framework, embraced by several quantitative scholars, this is labelled as the fundamental problem of causal inference (Neyman, 1923; Rubin, 1974; Holland, 1986). In this framework, one can assume that the case Z has only two potential outcomes, Y1 and Y0, which respectively correspond to the treated and untreated potential outcomes. Since Y0 cannot be observable in the (actual) case Z, researchers move to the cross-case setting to estimate the average causal effect. This is done by resorting primarily to randomized experiments or, alternatively, by analyzing observational data (e.g., Morgan and Winship, 2015).

However, important developments have also taken place in the individual case framework, where the attention is precisely focused on what would have happened if the event X had not occurred in the (single) case Z. These developments have essentially concerned the study of macro

phenomena and have involved both quantitative and qualitative research. Specifically, two different counterfactual methods have been raised, spanning both the quantitative and qualitative sides.

In the quantitative scholarship, the synthetic control method (SCM) has been developed in line with the potential outcome framework. It was initially introduced by Abadie and Gardeazabal (2003) and subsequently developed by Abadie et al. (2010; 2015) and by numerous econometricians (for an overview, see Abadie, 2019). The SCM has been adopted to investigate the impact of a large number of political events (for an overview, see e.g., Gilchrist et al. 2023). For instance, the SCM's inventors employed it to answer the following counterfactual questions: 'If the Basque Country had not experienced terrorist conflict, would its economic development be more pronounced?' (Abadie and Gardeazabal, 2003); 'If California had not implemented Proposition 99 (a large-scale tobacco control program), would its per-capita cigarette consumption be higher?' (Abadie et al., 2010); 'If 1990 German reunification had not occurred, would West Germany be more economically prosperous?' (Abadie et al., 2015).

On the other hand, several qualitative scholars use a counterfactual approach deriving from the historical explanation that, for convenience, can be termed the historical counterfactual method (HCM).[1] Its earliest advocate can be identified in Max Weber (Reiss, 2009). More recently, it was developed and adopted by numerous political scientists and sociologists (e.g., Fearon, 1991; Tetlock and Belkin, 1996; Lebow, 2000; Levy and Goertz, 2007; Harvey, 2012; Levy, 2015; 2015; Mahoney and Barrenechea, 2019). The HCM has been adopted primarily to answer counterfactual questions concerning international relations. Some examples are: 'If George W. Bush had not been elected president, would the United States not have invaded Iraq?' (Harvey, 2012); 'If the Archduke Franz Ferdinand had not been assassinated, would World War I (WWI) still have taken place?' (Lebow, 2014); and 'If the United States had not dropped the atom bomb, would the Japanese still have surrendered when they did?' (Newman, 1995).

Although these two methods both allow one to answer counterfactual questions concerning individual cases, they have not interacted so far. To the best of our knowledge, there is no methodological work that seek to bring the HCM and SCM into dialogue. Moreover, the respective proponents of these two methods hardly ever mention the counterpart. Finally, and perhaps most importantly, there are no counterfactual analyses that use the HCM and SCM jointly.

This paper aims to fill these gaps. To this end, the HCM and SCM are contrasted in order to highlight their methodological differences. Highlighting these differences is not only important to understand the extent to which the two methods diverge but also to appreciate when they can be used together and when they cannot be combined. So, once the conditions for combining the two methods have been identified, i.e., a circumscribed variation of an antecedent X affecting a single case Z and an outcome Y operationalizable both through a quantitative indicator and through the occurrence of a certain event (see Section 4), a concrete case of a joint use of them is presented. It regards the impact that Thatcher's election had on the UK pension system (see Section 5).

Before all this, the basic elements of the two respective methods are outlined (see Sections 2 and 3). This is necessary because it is possible that few readers are familiar with *both* methods as a result of the lack of dialogue between them.

[1]Giving a name to this methodological approach also stems from the fact that its developers use different terms to refer to it. For example, Tetlock and Belkin (1996) use the name 'counterfactual thought experiment', Harvey (2015) uses the name 'comparative counterfactual analysis', while Mahoney and Barrenechea (2019) speak generically of 'counterfactual analysis in case-study'.

SCM in a nutshell

To estimate what would have happened in case Z in the absence of event X, the SCM adopts a between-case approach. Specifically, it provides a systematic way to choose comparison units to construct the counterfactual case. In this sense, the SCM operates in line with the logic of comparative case studies and solving some of its drawbacks.

As the inventors of SCM themselves observe, economists engaged in comparative case studies usually estimate the impact of the event of interest by comparing the trajectory of quantitative outcomes, such as mortality rates, average income, crime rates, etc., for the case affected by that event, with the trajectory of the same outcome for one or more comparison units not exposed by that event (see for example Card, 1990; Card and Krueger, 1994).

However, this often implies great uncertainty about the ability of the selected comparison units to reproduce the counterfactual outcome trajectory that the affected units would have experienced in the absence of the event of interest. This is essentially because selecting one real case with the same characteristics of the treated unit with the exception of the occurrence of the event of interest is often a heroic task. Consequently, students who follow traditional comparative strategies are frequently criticized for the subjective measures of affinity between affected and unaffected units (Abadie et al., 2010; 2015).

The SCM precisely seeks to address these shortcomings by formalizing the selection of the comparison units. That selection is de facto based on the assumption that the characteristics of the treated unit can generally be much more accurately approximated by a combination of comparison units than by any single control unit. Accordingly, the so-called donor pool, i.e., a reservoir of potential control units, is used.[2]

Hence, through an appropriate data-driven procedure, the so-called synthetic control unit is constructed. It corresponds to a weighted average of those control units that receive positive weights. The weights assigned to each of the control units are chosen so that the synthetic control unit best reproduces the trajectory of the outcome variable of the affected case and its predictors in a given period prior to the event under examination (i.e., in the pre-treatment period).

This reproduction makes it such that the trajectory of the outcome variable for the synthetic control unit in the period following the event, i.e., in the treatment period, constitutes a reliable counterfactual scenario. The estimation of the effect of the event under examination can be then computed as the (temporal) difference between the values for the outcome variable of the actual unit and the values of the outcome variable for the synthetic control unit in the treatment period (Abadie et al., 2010; 2015).

Once the effect is estimated, its uncertainty is evaluated through placebo procedure based on the principle of permutation inference. This procedure is proposed by the inventors of the SCM because the use of the standard statistical inference in comparative case studies is difficult because of the small sample nature of the data, the absence of randomization, and because of the fact that probabilistic sampling is not employed to select sample units (for more details, see Abadie et al., 2010; 2015). However, more recently, further procedures to provide a measure of the uncertainty of effects estimated through SCM have been developed (Firpo and Possebom, 2018; Li, 2020; Cattaneo et al., 2021).

The key research steps prescribed by SCM are summarized in an online Appendix. The logic behind this method can be also clarified by reference to one of its inventors' applications: the analysis of the effect of Proposition 99 (Abadie et al., 2010). Proposition 99 was a large-scale tobacco control program implemented in California.

In that study, the case exposed to the event of interest (i.e., the treated unit) was obviously California, while the outcome was cigarette consumption, operationalized as annual per capita cigarette sales in packs at the state level. Accordingly and in line with the potential outcome

[2]Following the terminology adopted by the inventors of the SCM, the treated unit corresponds to the unit exposed to the event of interest, while units not exposed to that event are called comparison units or control units.

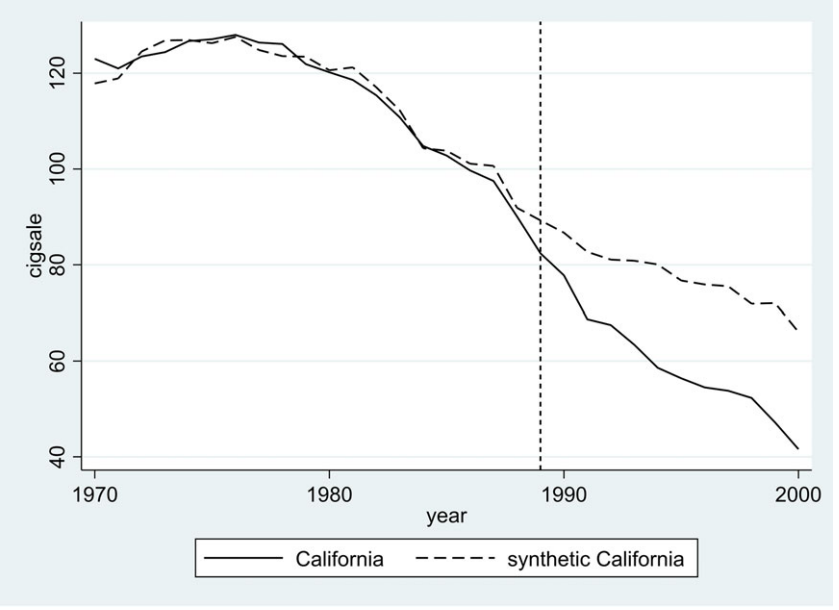

Figure 1. Cigarette consumption for California and the synthetic California (Abadie *et al.*, 2010).

framework, Y1 referred to the values of per capita cigarette consumption in the presence of Proposition 99, and Y0 denoted the values of per capita cigarette consumption in the absence of that treatment. Since the annual trajectory of Y0 was not obviously observable, it had to be reproduced.

To this end, annual state-level panel data for the period 1970-2000 was employed. Proposition 99 was passed in November 1988 and went into effect in January 1989 so that the pre-treatment period lasted from 1970 to 1988, while the treatment period lasted from 1989 to 2000. The dataset included California plus the other US states. However, the set of potential comparison units was constituted by 38 states only because the remaining US states adopted other large-scale tobacco control programs during the sample period.

For the pre-intervention characteristics, a set of cigarette consumption predictors was employed. It included (1) average retail price of cigarettes, (2) per capita state personal income, (3) the percentage of the population aged 15–24, and (4) per capita beer consumption.

Using the data-driven procedure mentioned above, positive weights resulted for Colorado, Connecticut, Montana, Nevada, and Utah, and the value of zero for the other potential controls. Accordingly, a synthetic California was constructed as a weighted average of those five states. Indeed, they were able to accurately reproduce values of cigarette consumption predictors and the trajectory of California cigarette consumption. In particular, the trajectory of California cigarette consumption for the synthetic unit almost overlaps with that of the actual unit in the 1970-1988 period (see Figure 1). Consequently, the synthetic California provided a reasonable approximation of the cigarette consumption that would have occurred in California over the treatment period in the absence of Proposition 99. Hence, the Y1 and Y0 trajectories were compared over the 1989-2000 period to estimate the effect of Proposition 99. In doing this, Abadie et al. (2010) found a large negative effect of Proposition 99. As Figure 1 shows, immediately after the law's passage, the two lines begin to diverge noticeably. While cigarette consumption in the synthetic California continued its moderate downward trend, the real California experienced a sharp decline.[3]

[3]For the sake of simplicity, results of the placebo test are omitted here.

HCM in a nutshell

Like SCM users, HCM proponents are engaged in constructing a well-founded counterfactual scenario. However, the HCM is not a formalized and standardized method like the SCM. For this reason, hereafter it is illustrated how most HCM users pursue this goal. They usually refer to the criterion of plausibility. Specifically, they try to give a plausible answer to the following questions: 'Would the outcome Y still have occurred in the absence of the antecedent X?' and 'How would the sequence of events prior to the outcome Y have been redirected if the antecedent X had not occurred?'. Since there is an endless number of paths that history might have taken in the absence of a certain event X, HCM users try to answer these questions by conjecturing and listing the alternative (virtual) worlds and then selecting one as the most plausible scenario. Therefore, the focus is on the plausibility of these alternative scenarios. In this sense, comparing the relative plausibility of competing – weak versus strong – counterfactual claims concerning the same antecedent is the essence of the HCM (Tetlock and Belkin, 1996; Bunzl, 2004; Levy, 2015; Harvey, 2015).

But how one can determine which counterfactual is most plausible? This can be addressed by taking into account that the path from the counterfactual antecedent X to the counterfactual outcome Y must have observable implications that can be tested somewhere in the real world. This can in turn be generally established through an intensive examination of the individual case under analysis. Specifically, if counterfactuals have left evidence, one can choose between competing counterfactuals and settle which were plausible and which were not on that evidence. Therefore, a detailed within-case analysis is performed to generate informed guesses of what would have happened had the chain of events between X and Y been different (Bennett and Elman, 2006; Mahoney, 2007; Grzymala-Busse, 2011; Bengtsson and Ruonavaara, 2017). For this reason, the HCM is often used in conjunction with process tracing (PT). PT is precisely a within-case method that allows one to trace the causal mechanism, i.e., the causal sequence of events between an initial cause X and a final outcome Y (e.g., Bennett 2010; George and Bennett, 2005; Beach, 2016).

However, if one relies solely on a bottom-up, inductive, process-tracing approach, the construction of a plausible counterfactual remains problematic. In other words, since a counterfactual proposition can never be subjected to any direct empirical test, theoretical knowledge should be used as well. Historical counterfactual analysis is also a theory-driven activity. The problem is that we have few well-established theories in political science and more generally in social sciences. We have, at best, contested theories and probabilistic relationships of modest strength that are conditional and often contextually dependent rather than universal. Accordingly, given the aim of explaining a single case and exploring a counterfactual under a specific set of historical conditions, contingent generalizations can often substitute for more general theoretical laws. We can also use a plurality of different theories to substantiate different links in the causal chain, as long as each theory is well established in its own domain and the propositions derived from those theories are reasonably consistent with each other (Tetlock and Belkin, 1996; Levy, 2015, 395).

Certainly, some degree of uncertainty can remain in the construction of a plausible counterfactual. HCM users do not habitually employ procedures to assess the uncertainty of the conjectured effect as do users of SCM and other quantitative techniques. However, some developments are worth noting in this direction. Perhaps the most relevant is that of Coppock and Kaur (2022) who integrate HCM with an approach from the quantitative causal inference literature called extreme value bounds. Extreme value bounds represent the logical range of average causal effects that are consistent with the world as one can observe it. The bounds start out quite wide, but, as these authors show, a systematic aggregation of counterfactual guesses can shrink fundamental uncertainty about average causal effects in a principled manner.

Furthermore and more habitually, the HCM uses the logic of necessary and sufficient conditions as a basic explanatory framework. Essentially, a necessary condition counterfactual

Table 1. Sequence of choices leading to war – 2002–2003

T 12	Invade and occupy Iraq with allies.
T 11	Issue final ultimatum to Saddam.
T 10	Reject France's demands to take military action off the table.
T 9	Go back to the UN to negotiate a second resolution.
T 8	Interpret Blix report(s) (claiming that no weapons were found in Iraq) as confirming. material breach and non-compliance.
T 7	Interpret Iraq's Weapons of Mass Destruction dossier as unacceptable.
T 6	Negotiate a strong resolution with a very clear mandate and rigid requirements for full, complete, and unfettered compliance.
T 5	Approach the United Nations to obtain a unanimously endorsed resolution to restart coercive inspections.
T 4	Deploy 200,000 troops backed by Congressional authorization to enhance the credibility of the coercive threat needed for a strong UN resolution.
T 3	Obtain congressional authorization to use force (before going to the UN).
T 2	Prioritize a multilateral approach to Iraq impasse.
T 1	Return UN inspectors to Iraq.

Source: adapted from Harvey (2012).

implies that the occurrence of the antecedent X is a necessary condition for the occurrence of the outcome Y. Correspondingly, the absence of the antecedent X is a sufficient condition for the non-occurrence of the outcome Y (Levy and Goertz, 2007; Mahoney et al., 2009; Mahoney and Barrenechea, 2019).

The key research steps prescribed by HCM are summarized in an online Appendix. In addition, to better clarify how the method operates, Harvey's (2012) counterfactual analysis of the Iraq War is hereafter briefly summarized. In this study, the conventional wisdom that the US decision to invade Iraq in 2003 was a product of the ideological agenda, misguided priorities, intentional deceptions and grand strategies of President George W. Bush and prominent 'neoconservatives' and "unilateralists" on his national security team is evaluated. The central tenet of this hypothesis is very clear: a Bush administration dominated by powerful neoconservatives was a necessary condition for the Iraq War. Accordingly, if George W. Bush had not been elected president, the United States would not have invaded Iraq.

Harvey regarded this as an implausible counterfactual: a claim without historical evidence about relevant causal mechanisms and a "theory" without theoretical content. Hence, he developed a plausible counterfactual via a detailed analysis of the historical events that unfolded in the months leading up to the onset of the Iraq War and via a theory-based interpretation of those events.

Specifically, consistent with the PT logic, he started from the final outcome, i.e., the invasion of Iraq, and moved backwards in time to trace a sequence of decisions during the period 2002-2003 (see Table 1). By so doing, the decision to invade Iraq cannot be viewed as a single and independent choice but the result of successive decisions taken by multiple actors and receiving a widespread bipartisan support. Consequently, it would be very difficult for neoconists or other proponents of Gore-peace to challenge the wisdom or rationality of the choices made at each of the previous stages. In other words, the path-dependent decisions that created the momentum toward invasion would almost certainly have been repeated under a Gore presidency.

Therefore, Harvey found scarce theoretical and historical support in favor of the view that if Al Gore had been elected President in place of Bush, the United States would not have invaded Iraq. So, if the negation of the actual antecedent implies no difference for the counterfactual outcome, the Bush election cannot be considered the necessary condition for the Iraq War.

However, Harvey's study might suggest that, by reconstructing and interpreting the causal chain between the antecedent and the outcome, the causal power of the antecedent is inevitably

lost. The traced series of decisions seems to create a path-dependent and unavoidable momentum to the final outcome. Yet, this is not a foregone conclusion. It depends on the time period one considers. In some cases, one may be faced with the so-called critical junctures. During a critical juncture, the structural (i.e., institutional, political, organizational, etc.) influences on policy-making are significantly relaxed and the potential impact of key political actors is expanded. Therefore, different chances of development are possible and prior conditions do not necessarily determine the direction of the process (Capoccia and Kelemen, 2007; see also Section 5).

A difficult relationship?

As observed above, both the SCM and HCM attempt to make causal inference starting from the following general counterfactual question: 'If the antecedent X did not occur in the (single) case Z, would the outcome Y not have occurred or would it have occurred differently?'. Nevertheless, the two methods were adopted to empirically answer different versions of this question without ever being used together.

Accordingly, this section points out which methodological constraints effectively prevent a combination of the HCM and SCM and what scope there is for a joint adoption of them. To this end, the focus will be on the manner in which the counterfactual question is concretely formulated, i.e., which types of antecedent and outcome are included in it and how they are operationalized.

The HCM is normally used to answer counterfactual questions in which both the antecedent and the outcome are constituted by particular historical events – i.e., well-bounded episodes in the history of a case marked by a particular occurrence or specific pattern of activity (García-Montoya and Mahoney (2023) – that may only occur or not occur. A war or the assassination of a political leader are events that can only happen or not happen. If a war does not happen there will be peace, but this is still a categorically different event. In line with the historical explanations, the HCM refers to categories that are mutually exclusive and collectively exhaustive (Mahoney, 1999). Therefore, the HCM adopts a binary logic: The following necessary condition counterfactual is proof of this. If we hypothesize X caused Y, we assume that Y would not have happened in the absence of X (Grynaviski, 2013).[4]

On the other hand, the SCM is routinely used to assess to what extent the outcome evolves differently if a particular event does not occur. The antecedent is also in this case an event that may occur or not occur, while the outcome is operationalized through a continuous (quantitative) variable (see Section 2).

In this respect, the main difference between the SCM and HCM concerns the outcome and not the antecedent. They could consequently be combined to provide different answers to what may have happened in the absence of a same (particular) event. It would be sufficient to operationalize the outcome included in the counterfactual question of interest through both a quantitative variable and a particular (categorical) event.

However, before that, it is necessary that the counterfactual question from which one starts allows both methods to make a valid causal inference, i.e., concluding that any (quantitative or qualitative) change in the outcome is exclusively attributable to the occurrence of the antecedent. The SCM and HCM – like other scientific methods and *in primis* experimental design – have to control for potential confounding factors (e.g., Sekhon, 2004). However, the two methods adopt different strategies to pursue this objective. This is essentially because they use two different approaches to reconstruct the counterfactual scenario. As said above, the HCM adopts a within-case approach while the SCM is used in a between-case framework. This precisely entails that the two methods are generally adopted to answer different counterfactual questions.

[4]As Mahoney and Barrenechea (2019) observe, the necessary condition counterfactual is more recurrent in the social sciences than other types of counterfactuals. However, the binary logic also concerns the other types of counterfactuals.

Using a between-case approach, the SCM makes use of a set of potential controls for *developing* the counterfactual scenario. This implies that it generally starts from counterfactual questions whose antecedents, such as large-scale policy interventions or prolonged socio-political processes, affect only one or a few units. In so doing, numerous other units remain untreated and as such can precisely be used as controls to construct the synthetic unit (see the study on California's Tobacco Control Program described in Section 2).

On the other hand, in order to construct a counterfactual lacking the event of interest and neutralizing the confounding factors, HCM users generally opt for a 'minimal rewrite of history' rule (Levy, 2015). This essentially stems from the adoption of the within-case approach. Case studies are highly interconnected systems so that a change in one variable induces changes in other variables, which then ripple through the system, generating further changes. Therefore, the HCM assumes that a minimal rewrite of the antecedent should leave disturbing factors unaltered (Tetlock and Belkin, 1996: 23; Levy, 2015; Levy and Goertz, 2007; Mahoney and Barrenechea, 2019).

Therefore, several HCM users start from counterfactual questions whose antecedents concern events affecting individual political leaders. Specifically, one can ask what would have happened in the absence of assassinations, deaths in office, and a different electoral outcome for a certain leader. As in the case of the non-election of George W. Bush (see Section 3), all these personal events provide easily imaginable ways to substitute one political leader (and their belief systems, personalities, risk orientations, etc.) for another without changing the structure of the political and economic systems, social attitudes, and system-level structures (Levy, 2015: 391).

Certainly, not all election results are circumscribed events like the (non-)election of George W. Bush. Most elections imply a more substantial number of changes to generate a different outcome, violating the minimal-rewrite rule.

However, Runhardt argues that, if one takes into account that HCM is not only interested in distinguishing the effect of X on Y but also in considering the role of potential intervening factors, it is not required – as the minimal-rewrite rule prescribes – to construct a counterfactual lacking only the event X of interest but identical in everything else to the real case. It should suffice that only those variables related to Y remain equal. More precisely, one should only change the putative cause X and not any other intervening factors related to Y but not on the path between X and Y. Accordingly, one can hypothesize small and contingent changes of the whole 'situation', leaving other aspects fixed (Runhardt, 2022: 1601).

This means, for example, that, as a consequence of a given election result, a certain intervening factor, such as the behavior of a given interest group, will change and, hence, the outcome of interest. Accordingly, in addition to the alteration of the election result, one may let such intervening factors vary as well.

Once shown that the HCM and SCM normally start from different counterfactual questions to be able to make a reliable causal inference, one must clarify in which cases the two methods cannot be used together and in which others they can instead be combined.

With reference to what has just been said about the neutralization of confounding factors in the HCM framework, it is easily understandable that pervasive socio-political processes cannot be included in counterfactual questions addressable by this method. Not being circumscribed events but phenomena permeating entire social contexts, their negation hardly leaves other aspects of the case at hand unaltered. This means that certain socio-political processes examined by the SCM, such as the terrorist conflict in the Basque Country (Abadie and Gardeazabal, 2003) or the mafia activity in the regions of southern Italy (Pinotti, 2015), are antecedents that are hardly analyzable via the HCM. For instance, the Basque conflict encompassed a wide range of terrorist activities. They resulted in almost 800 deaths. Furthermore, Basque entrepreneurs and corporations had been specific targets of violence and extortion. All of these occurrences pervaded the entire Basque country and interconnected with other local features. It is therefore difficult to conjecture that by

denying these processes other predictors of the outcome variable, i.e., the Basque GDP, remain unaltered.

Antecedents normally examined by the SCM that are more in line with the criteria adopted by HCM users to control for confounding factors are public policies. In many cases, the implementation of a public policy is no less a circumscribed event than some of the aforementioned affairs concerning individual political leaders. Moreover, in some counterfactual questions addressed through the HCM the antecedent was precisely a policy decision, e.g., 'If the United States had not dropped the atom bomb, would the Japanese still have surrendered when they did?'.

Combining the SCM and HCM to examine the effects of a given policy appears therefore a viable route. However, as said above, this is possible if the outcome is operationalized both through a quantitative indicator and through the occurrence of a particular event. Usually, this is problematic because policy evaluators operationalize the outcome by referring to the policy objective which generally coincides with a mere quantitative level. Policies are normally implemented to achieve a quantifiable objective, i.e., to reach a specific number or, at least, to reduce/increase values of a certain variable. For instance, reducing cigarette consumption was the main and the sole policy objective of Proposition 99 so that Abadie et al. (2010) used per capita cigarette sales in packs as an outcome variable (see above).

The double operationalization of the outcome – and, hence, the combination of the SCM and HCM – is more applicable if one refers to the policy's unintended effect. This is simply because the unintended effects of a policy may more easily include the occurrence of a certain event. Consider the following two examples. First, if a (single) country implements a natural resources policy, this will not only affect, as expected, its economic growth but, as a consequence of rent-seeking, will also produce an increase in corruption (Badeeb et al., 2017). This unintended effect can be operationalized both via a quantitative indicator, such as the corruption perception index, and via the occurrence of a particular event, such as a political corruption scandal. Second, if a country implements a policy aimed at relaxing immigration restrictions, in addition to an increase in immigration flow, reactions by natives will arise. Also, these unintended effects can be doubly operationalized: both through a scale on attitudes towards immigrants and the formation of an anti-immigrant political group.

This second example allows for a further consideration about the double operationalization. The choice to consider the formation of an anti-immigration group clearly denotes the focus on a negative reaction. In contrast, the selected scale may indicate that natives are on average not hostile toward immigrants. Such an inconsistency does not per se prevent the combination of the HCM and SCM. The researcher should simply take these potential incongruences into account before proceeding with any counterfactual analysis.

Let us now examine whether the counterfactual questions normally addressed by the HCM can be investigated by the SCM. The occurrence of a war as a result of certain personal events/decisions cannot be easily analyzed through the SCM.

First, the SCM appears intuitively inadequate to provide an answer to the counterfactual question: 'If Kennedy had followed the advice of his most hawkish advisers and immediately decided on an air strike against Soviet missile sites in Cuba, might World War III have been triggered in October 1962?'. Second, the lack of detailed time series for any outcome variable and its predictors for Austria-Hungary and its potential controls constitutes a primary impediment to performing a synthetic control analysis about the assassination of the Archduke Franz Ferdinand as the putative cause of the beginning of the WWI.

Moreover, the situation does not improve if one considers a more recent case, such as the Iraq War, for whom data availability is not an issue. As Podestà (2023: 61) observed, serious problems arise in operationalizing the US invasion of Iraq through a quantitative indicator. A first solution would be using the number of US troops who arrived in Iraq. This would imply that the corresponding US time series would denote zero values for the pre-treatment period (e.g., from

1992 to the George W. Bush 2000 election) as well as for the 2001-02 period and a value of about 130,000 for 2003. Using a set of Organisation for Economic Co-operation and Development (OECD) countries as potential control units to reproduce that trajectory would yield an insignificant result. Indeed, since none of the advanced countries sent troops to Iraqi territory, the synthetic line would be entirely composed of zero values. This would trivially mean that, in the absence of George W. Bush, none of the US troops would have arrived in Iraq. Serious problems would also arise if military expenditure were employed to measure the outcome. In this case, the US trajectory for the outcome variable would not be synthetically reproducible because, during the pre-treatment period, the US level of military expenditure was, on average, significantly higher than that of any potential control unit. Furthermore, the Bush impact on the Iraq War would be indiscernible because US military expenditure had begun to increase significantly *before* the Iraq War as a consequence of the Afghanistan War.

However, this issue is specific to the case at hand and does not undermine the applicability of SCM to examine military expenditure as an outcome. For instance, if one were to examine the effect of Thatcher's electoral victory on the outbreak of the Falklands War, the above problems do not emerge. The UK level of military expenditure was not higher than that of other Western democracies and, at that time, UK military spending was not upward biased for participation in another war. Nevertheless, the increase in UK military spending was so modest and temporary that this indicator appears inadequate to measure the Falklands conflict quantitatively.

The important fact, however, is that SCM can be used to estimate the impact of certain election results. Several recently published studies testify to this. Some synthetic control analyses were performed to estimate the political and economic effects of the election of some political leaders (Grier and Maynard 2016; Silva 2018; Absher et al. 2020). The possibility of using the two methods jointly is therefore not precluded. Grier and Maynard's (2016) study on the impact of Hugo Chavez's election allows one to address two important issues regarding such a potential combination. The first issue concerns the boundaries of the antecedent. Taking into account that no Latin American or Organization of the Petroleum Exporting Countries (OPEC) member countries were affected by the election of Hugo Chavez, all these countries were included in the control group used to construct a synthetic Venezuela.

However, in other case studies, especially those involving non-presidential systems, this way of operating may be too trivial. The non-election of a certain political leader is generally associated with the electoral defeat of their political party. The boundaries of the antecedent can thus be extended to the election victory of a political party or even a given political coalition. In the case of the SCM, this simply means excluding from the control group those countries exposed to the government of those parties. In the case of the HCM, the negation of this (enlarged) antecedent can be seen as a violation of the minimal rewrite rule. However, that negation can remain valid as long as one can still assume that any independent influences on the effect that stand outside the process between X and Y remain the same (see the argument illustrated above).

The second issue concerns the double operationalization of the outcome required to combine the two methods. The synthetic control analysis on Hugo Chavez's election was performed to estimate its impact on some national socio-economic performances, i.e., per capita income, life expectancy, infant mortality, poverty, and inequality. However, the adoption of the HCM is precluded because national socio-economic performances are not measurable through the occurrence of specific events.

Nevertheless, the double operationalization of the outcome is easier if one wants to examine what would have happened in a given domestic policy area in the absence of a certain election result. This is because several domestic policy areas, such as the welfare state, labor market, education, the environment, etc., can be operationalized both through the occurrence of certain events, such as the implementation of certain policy interventions, and quantitative indicators, such as public expenditure.

In view of these argumentations, the next section will present a concrete example of the combined use of the HCM and SCM for analyzing the effect of an electoral result on a particular social policy.

Combining the HCM and SCM in analyzing the Thatcher government's pension reforms

The concrete example presented here concerns the impact that Thatcher's election had on the UK pension system. Since the UK pension system is regarded as one of the most altered social programs by the conservative revolution of the 1980s, one can start from the following counterfactual question: 'If Margaret Thatcher had not been elected, would the UK pension system still have undergone severe cuts?'.

This counterfactual question has never been addressed through the HCM. However, some qualitative studies can be exploited as they have traced the sequence of events characterizing the policymaking of the pension reforms set up by the Thatcher government: Pierson's (1994) seminal work of Reagan and Thatcher's welfare state retrenchments and some more recent historical studies precisely on the Thatcher government's pension policy (Pemberton et al. 2017; Davies et al. 2023).[5] On the SCM side, one can exploit Podestà's (2020) synthetic control analysis which was performed precisely to counterfactually replicate Pierson's (1994) work.

In view of what was said in the previous section, the antecedent can be extended to the electoral victory of the Conservative Party and not of Margaret Thatcher alone. Furthermore, in order to be able to use both methods for counterfactually analyzing the effect of such an electoral result on a social program, this latter must be operationalized through both the enactment of certain reforms and a quantitative indicator. The qualitative studies above mentioned indicate that during Thatcher's administration two pension reforms were effectively set up, the first in 1981 and the second in 1986; while Podestà (2020), in his synthetic control analysis, operationalized the outcome through the pension generosity index using data from the Comparative Welfare Entitlements Dataset (Scruggs et al. 2014). Both the content of these two reforms and the trajectory of this index denote a remarkable pension retrenchment. The 1980 reform established that pensions would be updated only in line with prices, while the 1986 reform gave all citizens the right to leave their occupational pension scheme, or the State Earnings-Related Pension Scheme (SERPS), and put their pension savings into a 'personal pension'. On the other hand, the level of pension generosity fell by about one point from 11.1 in 1979 to 10.2 in 1989, nullifying much of the growth manifested in the previous decade.

In view of that, the counterfactual question mentioned at the beginning of this section can be reformulated into two distinct questions, the first to be answered through the HCM and the second through the SCM. They are: 'If Margaret Thatcher had not won the election, would the 1981 and 1986 pension reforms have been set up anyway?' and 'If Margaret Thatcher had not won the election, would the UK pension generosity still have decreased significantly?'.

Moreover, in this case study, the antecedent denotes a peculiarity that entails a further reformulation of one of the aforementioned two counterfactual questions. Since Margaret Thatcher won the election in 1979 and 1983, one should try to evaluate separately what would have happened if she had lost the two elections. This poses no relevant problem for the HCM. Although this method operates with a temporal sequence of events that may cumulatively affect each other (see Section 3), in this case, the events occurred in such a way that they could be examined separately. Following the two electoral victories, two pension reforms were in fact implemented, one in 1981 and the other in 1986. The HCM can then be used to answer the following two counterfactual questions: (1) 'If Thatcher had lost the election in 1979 would the 1981 pension reform have been implemented anyway?' and (2) 'If Thatcher had lost the election in 1983 would the 1986 pension reform have been implemented anyway?'.[6]

[5]Although Pierson (1994) reconstructed in detail the decision-making of the Thatcher governments from 1979 to 1988 in the area of pensions, he explained the policy outcomes by comparing the British case with the American case. In so doing, he did not provide any counterfactual analysis of what would have happened in the case in which Margaret Thatcher had lost the election. A counterfactual explanation is not even provided by the other qualitative studies above-mentioned.

In contrast, the SCM does not allow the impact of the two election results to be assessed separately. This is essential because it constructs the synthetic control unit by referring to the pre-treatment period. Specifically, Podestà could not construct a synthetic UK that would best reproduce the characteristics of the real UK in the pre-1983 period precisely because from 1979 to 1983 UK was already treated. Conversely, through a combination of OECD countries not exposed to the conservative resurgence, Podestà was able to construct a synthetic UK which resembles relevant characteristics of Britain before Thatcher's 1979 election victory.[7]

In this regard, it should, however, be noted that SCM makes it possible to estimate the impact of the two (consecutive) Thatcher governments. The SCM can in fact accommodate repeatedly administered treatments. It does not merely estimate the effect of a single binary treatment (e.g., X(1) - X(0)); rather, it compares full treatment histories, such as X(1, 1, . . . , 1) - X(0, 0, . . . , 0). When two distinct treatments occur sequentially, the relevant estimand becomes X(1, 1, \ . . . , 1, 2, 2, . . . , 2) - X(0, 0, . . . , 0), where 1 and 2 represent the respective treatments which in the previous case can be referred to the two Thatcher governments.

Although SCM appears more limited than HCM with respect to the above issue, it allows a significantly better assessment of the control of confounding factors. The reliability of the synthetic unit constructed by Podestà (2020) was precisely assessed through its ability to reproduce the pre-treatment trajectory of the outcome variable, i.e., the UK pension generosity, and to approximate the pre-treatment values of a set of covariates (i.e., population aged 65 and older, unemployment rate, trade openness, GDP per capita, neo-corporatism). In particular, the synthetic line and the treated line denote almost coinciding trajectories (see Figure 2).

In contrast, such an explicit and formalized control is not possible using the HCM. One can only provide theoretical arguments about the fact that, by denying the Conservative Party's election victories in 1979 or 1983, the other factors which may have impacted the implementation of the two pension reforms remain fixed. In this regard, one can, for example, conjecture that demographic factors that may have affected the two pension reforms would not have varied in the event of Thatcher's electoral defeats.

That said, let us examine how the two methods reconstruct the counterfactual outcome and how they answer the respective counterfactual questions.

Having created a synthetic UK which adequately mimics the characteristics of the real nation prior to the political treatment of interest, Podestà (2020) was able to count on a satisfactory approximation of the pension-generosity trajectory that would have been experienced by the UK in the absence of Thatcher's election. Hence, comparing the trajectory of the treated line with that of the synthetic line during that period, he could estimate the effect produced by Thatcher's election on UK pension generosity. As Figure 2 shows, the effect is evidently negative. The two lines progressively diverge noticeably due to the fact that the synthetic one denotes a remarkable increase and the treated one exhibits a growing reduction.[8]

Turning to the historical counterfactual analysis, one can first consider that the Labour Party manifestos of 1979 and 1983 present an entirely different policy agenda from those of the Conservatives in all policy areas (Royed, 1996). Accordingly, one may hypothesize that if the Labour Party had won the two elections the 1980 and 1983, pension reforms would not have been

[6]For reasons of simplicity, the two counterfactual questions merely ask whether the two reforms have been (or have not been) implemented. However, in certain cases, this option might be too trivial and the researcher might then wonder whether in the absence of a certain election result the policy would have been set up differently. However, consistent with what was said in Section 4, the logic would remain binary: the non-implementation of the policy Y1 would imply the implementation of the policy Y2.

[7]The UK and USA were considered as the only treated countries because, in the 1980s, the Thatcher and Reagan administrations were the only protagonists of the conservative revolution of the 1980s. However, the results do not change significantly if one discards from the donor pool those countries with governments composed of right-wing parties during the sample period examined (Podestà, 2020).

[8]The estimate of this effect is confirmed by placebo tests. They are not reported here. For more details see Podesta (2020).

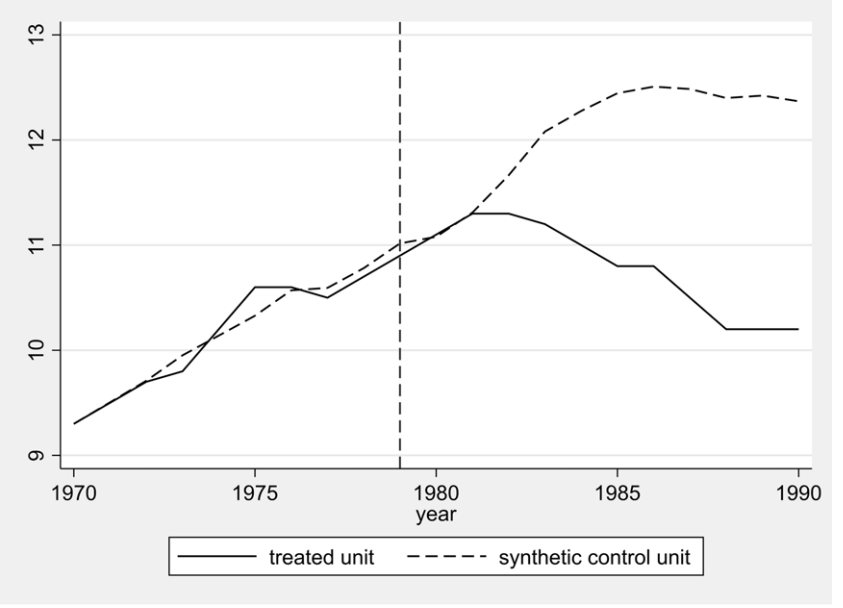

Figure 2. Pension generosity for the UK and the synthetic UK (Podestà, 2020).

implemented. However, consistent with the HCM logic, the plausibility of this simple counterfactual hypothesis cannot only be assessed in relation to the absence of the antecedent, i.e., Thatcher's election, but also in relation to the reconstruction and the scrutiny of the sequence of events leading to the final outcome, i.e., the enactment of the two reforms.

For the sake of space and given its greater relevance, let us consider only the sequence of events prior to the 1986 reform (see Table 2).

This chain of events reveals the importance assumed by a right-wing think tank, namely No. 10 Policy Unit, as a source of neoliberal ideas for reforming the UK pension system (see T3 and T4 in Table 2). One could thus argue that in the absence of these intervening factors Thatcher's second government would not have enacted such a radical pension reform. This is also due to the fact that the Conservative manifesto for the 1983 election was quite cautious on pensions (Davies et al., 2023).

However, someone else might rightly counter-argue that Thatcher should have been operating at a critical juncture (see Section 3). British politics is distinguished by the concentration of power in the executive, and, at that time, interest groups were quite weak and fragmented (Pierson, 1994). Accordingly, the Thatcher government could operate in a very favorable environment. It could thus freely set up a right-wing think thank (see T2 in Table 2) and, despite social opposition (see T5 in Table 2), embrace the neo-liberal spirit of its proposal. Although the 1986 reform did not abolish SERPS, it introduced a new system of voluntary personal pensions (see above).

This allows one to conclude that the Thatcher election was a necessary but not sufficient condition for the enactment of the 1986 pension reform. Equivalently, one can claim that the absence of the Thatcher government was a sufficient condition for the non-enactment of that reform. In other words, if the Labour Party had won the 1983 election, the Policy Unit would never have been set up, making the enactment of the 1986 reform highly unlikely.

However, someone else might rightly counter-argue that the critical juncture in which the Thatcher government operated allowed it to set up a right-wing think thank and, at least partially, embrace its proposals. In other words, if the Labour Party had won the 1983 election it too would have operated under these conditions and the Policy Unit would never have been set up, making

Table 2 Sequence of events leading to the 1986 pension reform

T 6	In 1986, the Social Security Act was promulgated.
T 5	The government was confronted with public hostility that the green paper produced. Opposition from the traditional supporters of social programs was accompanied by employers and the occupational pension funds.
T 4	In June 1985, the Policy Unit published a green paper on social security reform. Its many proposals included the abolition of SERPS. More than this, however, SERPS members were to be compelled to take out a personal pension.
T 3	In 1984, under Norman Fowler's direction, the Policy Unit clearly expressed the advice to abolish SERPS.
T 2	In the wake of the electoral results, Thatcher abolished the Central Policy Review Staff and strengthened the ability of the No. 10 Policy Unit both to develop policy ideas and to promulgate policy change.
T 1	As a result of the June 1983 general election, the Conservatives increased their parliamentary majority from 53.4 percent of seats to 61.1 percent of seats.

the enactment of the 1986 reform highly unlikely. Historical evidence that makes this scenario plausible is the Labour Party leader Neil Kinnock's statement which, even before the green paper was released, announced that a Labour government would reverse any major modifications of SERPS (Pierson, 1994). Accordingly, one may conclude that the Thatcher government was a necessary but not sufficient condition for the enactment of the 1986 pension reform and, equivalently, the absence of the Thatcher government was a sufficient condition for the non-enactment of that reform.

To sum up, the two methods, while following different paths of analysis, both conclude that if Thatcher had not won the election the UK pension system would not have suffered severe cuts. This convergence of results is not a foregone conclusion when one combines the HCM and SCM. In other cases, the results may substantively diverge. Beyond this, the above example shows that using the two methods together allows one to provide a more complete picture of the phenomenon being studied. Addressing different and specific counterfactual questions, using distinct operationalizations of the outcome, and adopting different approaches for reconstructing the counterfactual scenario allow one to address the puzzle under examination from different viewpoints.

Conclusion

In recent years, many efforts have been made to bring quantitative and qualitative methods into dialogue. This article also moves in that direction. However, in contrast to most works, the present attempt does not concern the large-N/small-N issue but focuses instead on the sole single case study framework.

Within this framework, two counterfactual methods, the HCM, the qualitative one, and the SCM, the quantitative one, have gained great importance without however meeting. This paper has thus tried to reduce their distance. More precisely, it has tried to assess whether these two methods can be used together to understand what would have happened in a single case Z in the absence of an event X.

This appears viable under certain conditions only. Both methods must first of all be able to make a valid causal inference. This implies certain constraints because the SCM and HCM use different approaches to adequately control for confounding factors. The joint use of the two methods emerges, however, as feasible if one starts from a counterfactual question that (1) involves an antecedent affecting an individual case making numerous control units available, and, at the same time, (2) hypothesizes a rather circumscribed variation of that antecedent leaving the other factors of Y unaltered.

Yet, this is not sufficient. A further condition must be fulfilled. The outcome must be capable of being operationalized both through a quantitative indicator and through the occurrence of a particular event. This is precisely because the SCM and HCM respectively adopt these two operationalizations of the outcome.

Two classes of counterfactual questions are able to meet these conditions. They are: 'What would have been the unintended effects of a given policy if it had not been implemented?' and 'What would have happened in a given domestic policy area if a certain political result had not occurred?'.

However, these two classes of counterfactual questions may not be the only ones that allow the combination of the SCM and HCM. Their identification is derived from a reworking of the main applications of the two methods. These applications are de facto representative of what the two methods respectively allow for in terms of counterfactual analysis. Nevertheless, exploring further research areas could be promising to identify other counterfactual questions that can be examined by both methods.

Furthermore, starting with a counterfactual question belonging to one of these two classes does not necessarily guarantee the possibility of effectively combining the two methods. The two classes of counterfactual questions are, in general, suitable for fulfilling the above-mentioned conditions for the combination of the SCM and HCM, but it depends a great deal on how the question is empirically treated. The example illustrated above concerning the impact of Margaret Thatcher's election on the UK pension system shows that the two methods can indeed be fruitfully combined and that returns in terms of causal inference may be reached from that combination.

Moreover, it can be noted that the goal of the present work was rather ambitious. As mentioned, it sought to determine under what conditions HCM and SCM can be adopted jointly to answer the same counterfactual question. Certainly, this made it possible to clarify the differences of the two methods and to understand what different answers they can provide to the question, 'What would have happened in a single case Z in the absence of an event X?'. This reduced, however, the possibilities of using the two methods in the same research program. In fact, the use of the two methods can occur under fewer research constraints.

Most empirical studies involve multiple, sometimes sequential questions. The two methods can be thus applied in tandem to examine different stages of the causal chain. For instance, SCM has been adopted by Grier and Maynard (2016) to estimate the impact of Hugo Chavez's election on some Venezuelan socio-economic performances (see above). Since these performances can only be measured through appropriate quantitative variables, HCM cannot be employed for this purpose. It could, however, be employed to answer a question that SCM cannot answer: to determine whether, in the absence of certain events, such as the so-called 'Caracazo', i.e., a wave of popular protests against austerity measures imposed by former President Carlos Andrés Pérez (including rising fuel and transportation prices), Hugo Chavez would still have won the 1998 presidential election. In other words, the event that is used as the antecedent in the synthetic control analysis can be adopted as the outcome when using HCM.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/S175577392510012X.

Acknowledgements. For helpful comments on previous versions, we thank Gary Goertz, Rosa Ruhnard, Luis Schenoni and several FBK-IRVAPP researchers. We would also like to thank three anonymous referees for providing valuable comments and enabling us to improve our article accordingly.

Funding. This study did not receive any funds.

Competing interests. The authors declare that they have no conflict of interest.

References

Abadie, Alberto. "Using Synthetic Controls: Feasibility, Data Requirements, and Methodological Aspects." *Journal of Economic Literature* **59**.2 (2019): 391–425.

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program." *Journal of the American statistical Association* **105**.490 (2010): 493–505.

Abadie, Alberto, Alexis Diamond,, and Jens Hainmueller. "Comparative Politics and the Synthetic Control Method." *American Journal of Political Science* **59**.2 (2015): 495–510.

Abadie, Alberto, and Javier Gardeazabal. "The Economic Costs of Conflict: A Case Study of the Basque Country." *American economic review* **93**.1 (2003): 113–132.

Absher, Samuel, Kevin Grier,, and Robin Grier. "The Economic Consequences of Durable Left-Populist Regimes in Latin America." *Journal of Economic Behavior & Organization* **177** (2020): 787–817.

Badeeb, Ramez Abubakr, Hooi Hooi Lean, and Jeremy Clark. "The Evolution of the Natural Resource Curse Thesis: A Critical Literature Survey." *Resources Policy* **51** (2017): 123–134.

Beach, Derek. "It's all about Mechanisms - What Process-Tracing Case Studies Should be Tracing." *New Political Economy* **21**.5 (2016): 463–472.

Bennett, Andrew. . "Process tracing and causal inference." *Rethinking Social Inquiry*, ed. by Henry Brady and David Collier, Lanham MD: Rowman and Littlefield, 2010, pp. 207–219.

Bennett, Andrew, and Colin Elman. "Qualitative Research: Recent Developments in Case Study Methods." *Annual Review of Political Science* **9** (2006): 455–476.

Bengtsson, Bo, and Hannu Ruonavaara. "Comparative Process Tracing: Making Historical Comparison Structured and Focused." *Philosophy of the Social Sciences* **47**.1 (2017): 44–66.

Bunzl, Martin. "Counterfactual History: A User's Guide." *The American historical Review* **109**.3 (2004): 845–858.

Capoccia, Giovanni, and R. Daniel Kelemen. "The Study of Critical Junctures: Theory, Narrative, and Counterfactuals in Historical Institutionalism." *World politics* **59**.3 (2007): 341–369.

Card, David. "The Impact of the Mariel Boatlift on the Miami Labor Market." *Industrial and Labor Relations Review* **44** (1990): 245–257. [493,494,496].

Card, David, and Alan B. Krueger "Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania." *American Economic Review* **84** (1994): 772–793. [493,496].

Cattaneo, Matias D., Yingjie Feng, and Rocio Titiunik. "Prediction Intervals for Synthetic Control Methods." *Journal of the American Statistical Association* **116**.536 (2021): 1865–1880.

Coppock, Alexander, and Dipin Kaur, . "Qualitative Imputation of Missing Potential Outcomes." *American Journal of Political Science* **66**.3 (2022): 681–695.

Davies, Andrew, Jonathan Freeman, and Hugh Pemberton. "Thatcher's Policy Unit and the "Neoliberal Vision"." *Journal of British Studies* **62**.1 (2023): 77–103.

Fearon, James. . "Counterfactuals and Hypothesis Testing in Political Science." *World Politics* **43**.2 (1991): 169–195.

Firpo, Sergio, and Vitor Possebom, . "Synthetic Control Method: Inference, Sensitivity Analysis and Confidence Sets." *Journal of Causal Inference* **6**.2 (2018): 20160026.

García-Montoya, Laura, and James Mahoney, . "Critical Event Analysis in Case Study Research." *Sociological Methods & Research* **52**.1 (2023): 480–524.

George, Alexander L. and Andrew Bennett. *Case Studies and Theory Development in the Social Sciences*. Cambridge, MA: MIT Press, 2005.

Gilchrist, David, Thomas Emery, Nuno Garoupa,, and Rok Spruk, . "Synthetic Control Method: A Tool for Comparative Case Studies in Economic History." *Journal of Economic Surveys* **37**.2 (2023): 409–445.

Grier, Kevin, and Norman Maynard, . "The Economic Consequences of Hugo Chavez: A Synthetic Control Analysis." *Journal of Economic Behavior & Organization* **125** (2016): 1–21.

Grynaviski, Eric. "Contrasts, Counterfactuals, and Causes." *European Journal of International Relations* **19**.4 (2013): 823–846.

Grzymala-Busse, Anna. "Time will Tell? Temporality and the Analysis of Causal Mechanisms and Processes. *Comparative Political Studies* **44**.9 (2011): 1267–1297.

Harvey, Frank. *Explaining the Iraq War: Counterfactual Theory, Logic, and Evidence*, Cambridge: Cambridge University Press, 2012.

Harvey, Frank. "What If" History Matters? Comparative Counterfactual Analysis and Policy Relevance." *Security Studies* **24**.3 (2015): 413–424.

Holland, Paul W. "Statistics and Causal Inference." *Journal of the American Statistical Association* **81**.396 (1986): 945–960.

Lebow, Richard Ned "What's so Different about a Counterfactual?" *World Politics* **52**.4 (2000): 550–585.

Lebow, Richard Ned. . *Archduke Franz Ferdinand Lives!: A World without World War I*. St. Martin's Press, 2014.

Levy, Jack S. and Gary Goertz "Explaining war and peace: case studies and necessary condition counterfactuals." *Explaining War and Peace: Case Studies and Necessary Condition Counterfactuals*, ed. by Jack S. Levy and Gary Goertz. London: Routledge, (2007).

Levy, Jack S. "Counterfactuals, Causal Inference, and Historical Analysis." *Security Studies* **24**.3 (2015): 378–402.
Li, Kathleen T. "Statistical Inference for Average Treatment Effects Estimated by Synthetic Control Methods." *Journal of the American Statistical Association* **115**.532 (2020): 2068–2083.
Mahoney, James. "Nominal, Ordinal, and Narrative Appraisal in Macro-Causal Analysis." *American Journal of Sociology* **104**.4 (1999): 1154–1196.
Mahoney, James. . "The elaboration model and necessary causes." *Explaining War and Peace: Case Studies and Necessary Condition Counterfactuals*, ed. by Gary Goertz., Jack S. Levy. New York, NY: Routledge, 2007, pp. 281–306.
Mahoney, James, Kimball, Erin, and Koivu, Kendra L. . "The Logic of Historical Explanation in the Social Sciences." *Comparative Political Studies* **42**.1 (2009): 114–146.
Mahoney, James, and Rodrigo Barrenechea. "The Logic of Counterfactual Analysis in Case-Study Explanation." *The British Journal of Sociology* **70**.1 (2019): 306–338.
Morgan, Stephen L. and Christopher Winship. *Counterfactuals and Causal Inference: Methods and Principles for Social Research*, 2nd edn, Cambridge: Cambridge University Press, 2015.
Newman, Richard B. "Ending the War with Japan: Paul Nitze's" Early Surrender" Counterfactual." *Pacific Historical Review* **64**.2 (1995): 167–194.
Neyman, Jerzy. "On the Application of Probability Theory to Agricultural Experiments. Essay on Principles. Section 9. Reprint, 1990." *Statistical Science* **5**.4 1923: 465–472. With Dabrowska, Dorota M. and Speed, Terence P.
Pemberton, Hugh, Andrew Davies, Jonathan Freeman, and Trevor Gould . *The Fowler Inquiry into Provision for Retirement and the 1986 Personal Pensions Revolution: Briefing note for panellists*. A witness seminar at the Institute and Faculty of Actuaries, London, 2017.
Pierson, Paul. *Dismantling the Welfare State?: Reagan, Thatcher and the Politics of Retrenchment*. Cambridge University Press, 1994.
Pinotti, Paolo. . The economic costs of organised crime: Evidence from Southern Italy. *Economic Journal* **125** (2015): F203–F232.
Podestà, Federico. "Combining Process Tracing and Synthetic Control Method: Bridging Two Ways of Making Causal Inference in Evaluation Research." *Evaluation* **29**.1 (2023): 50–66.
Podestà, Federico. "Did the Conservative Resurgence Matter? A Counterfactual Replication of Pierson's Analysis on Reagan and Thatcher's Welfare State Retrenchments." *Comparative European Politics* **18** (2020): 819–839.
Reiss, Julian. . "Counterfactuals, Thought Experiments, and Singular Causal Analysis in History." *Philosophy of Science* **76**.5 (2009): 712–723.
Royed, Terry J. "Testing the Mandate Model in Britain and the United States: Evidence from the Reagan and Thatcher eras." *British Journal of Political Science* **26**.1 (1996): 45–80.
Rubin, Donald B. "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies." *Journal of Educational Psychology* **66**.5 (1974): 688.
Runhardt, Rosa W. "Concrete Counterfactual Tests for Process Tracing: Defending an Interventionist Potential Outcomes Framework. *Sociological Methods & Research* (2022): 00491241221134523.
Scruggs, Lyle. (2014). Social Welfare Generosity Scores in CWED 2: A Methodological Genealogy (WP-01; CWED Working Paper Series, p. 19). University of Connecticut. http://cwed2.org/Data/CWED2_WP_01_2014_Scruggs.pdf
Sekhon, Jasjeet S. "Quality Meets Quantity: Case Studies, Conditional Probability, and Counterfactuals. *Perspectives on Politics* **2**.2 (2004): 281–293.
Silva, Bruno Castanho. "Populist Radical Right Parties and Mass Polarization in the Netherlands." *European Political Science Review* **10**.2 (2018): 219–244.
Tetlock, Philip E. and Aaron Belkin. "Counterfactual thought experiments in world politics: logical, methodological, and psychological perspectives," *Counterfactual Thought Experiments in World Politics: Logical, Methodological, and Psychological Perspectives* ed. by Philip E. Tetlock and Aaron Belkin. Princeton, NJ: Princeton University Press, 1996.

Cite this article: Podestà F and Madia JE (2025). Contrasting and combining the historical counterfactual method and the synthetic control method. *European Political Science Review*, 1–17. https://doi.org/10.1017/S175577392510012X