

Geodesic Synthetic Control Methods for Random Objects and Functional Data

Daisuke Kurisu^{1,*}, Yidong Zhou^{2,*}, Taisuke Otsu³, and Hans-Georg Müller^{2,†}

¹Faculty of Economics, The University of Tokyo, Japan

²Department of Statistics, University of California, Davis, USA

³Department of Economics, London School of Economics, UK

September 1, 2026

Abstract

We introduce a geodesic synthetic control method for causal inference with panel outcomes that are random objects in a geodesic metric space. Examples include distributions, compositions, networks, symmetric positive-definite matrices, trees and functional data, among other data types that require a geometry beyond Euclidean vector spaces. The proposed method replaces Euclidean weighted averages by weighted Fréchet means and defines treatment effects as geodesics connecting untreated and treated potential outcomes. We develop a causal model with stochastic perturbations for object-valued untreated outcomes, establish consistency of the estimated weights for the perturbed population target, derive a bound on post-treatment prediction error, and give sufficient conditions for consistent recovery of the untreated counterfactual. For uncertainty quantification, we propose an intrinsic metric conformal calibration procedure that yields prediction regions for untreated counterfactual objects, confidence sets for geodesic treatment effects, confidence intervals for their magnitudes, and global no-effect tests. The method is illustrated through simulation studies for networks and symmetric positive-definite matrices, and through applications to employment composition changes following the 2011 Great East Japan Earthquake and the impact of abortion liberalization on fertility patterns in East Germany.

Keywords: causal inference, Fréchet mean, geodesic metric space, panel data, program evaluation

*The first two authors contributed equally to this work and are listed alphabetically.

†Corresponding author: hgmueeller@ucdavis.edu.

‡D.K. was supported in part by JSPS KAKENHI Grant Numbers 23K12456, 25K00624, and 26K00323. H.G.M. was partially supported by NSF grant DMS-2310450.

1 Introduction

Since the seminal work of [Abadie and Gardeazabal \(2003\)](#) and [Abadie et al. \(2010\)](#), synthetic control methods (SCMs) have become a widely used framework for evaluating causal effects of policy interventions in observational panel and longitudinal studies. SCMs are designed for settings in which some units are exposed to an intervention while others are not, with outcomes observed both before and after the intervention. The central idea is to construct, from the untreated units, a weighted synthetic control that closely reproduces the pre-treatment trajectory of the treated unit, and then to use this synthetic control as a counterfactual benchmark in the post-treatment period. The method has been applied in a wide range of comparative case studies; see [Abadie \(2021\)](#) for a recent survey.

Substantial methodological developments have extended SCMs in several directions. These include penalized synthetic controls that improve matching and regularization ([Abadie and L'Hour, 2021](#)), SCMs for high-dimensional settings ([Robbins et al., 2017](#)), connections with matrix completion methods ([Amjad et al., 2018](#)), and inferential procedures based on conformal prediction and robust testing ([Chernozhukov et al., 2021a,b](#)). Bias-corrected and augmented synthetic control estimators have been developed to address imperfect pre-treatment fit ([Ben-Michael et al., 2021](#)), with further extensions to staggered adoption settings ([Ben-Michael et al., 2022](#)). In addition, [Arkhangelsky et al. \(2021\)](#) introduced synthetic difference-in-differences, which combines features of SCM and difference-in-differences. These developments, however, have primarily focused on Euclidean outcomes.

Many modern applications involve outcomes that are not naturally represented as scalars or vectors. Examples include distributions ([Petersen and Müller, 2016](#); [Panaretos and Zemel, 2020](#); [Petersen et al., 2022](#)), compositional data ([Scealy and Welsh, 2011](#)), manifold-valued data ([Patrangenaru and Ellingson, 2015](#); [Kurisu and Otsu, 2026](#)), networks ([Zhou and Müller, 2022](#); [Severn et al., 2022](#)), functional data ([Hsing and Eubank, 2015](#); [Wang et al., 2016](#)), and tree-structured data ([Billera et al., 2001](#)). These objects are often

best viewed as elements of a metric space endowed with a geometry appropriate for the type of data one observes. Common choices include the Wasserstein metric for distributions, the arc-length metric for compositional data, Riemannian metrics for manifold-valued data, the Frobenius or power metric when representing networks through their graph Laplacians, the L^2 metric for functional data, and the BHV metric for tree spaces. Many such spaces are geodesic metric spaces, but they generally lack linear operations, such as addition, subtraction, and scalar multiplication, on which classical SCMs rely. This creates a fundamental obstacle for causal inference, since both the construction of a weighted synthetic control and the definition of a treatment effect must be reformulated intrinsically in the geometry of the outcome space. General approaches to address related challenges for data analysis in metric spaces are discussed in [Dubey et al. \(2024\)](#).

Causal inference for non-Euclidean and geodesic outcomes is a recent and rapidly developing area. Existing contributions include methods for average treatment effects ([Lin et al., 2023](#)), difference-in-differences ([Torous et al., 2024](#)), synthetic controls ([Gunsilius, 2023](#); [Gunsilius et al., 2024](#)), and regression discontinuity designs ([Van Dijke, 2025](#)), with many of these focusing on distributional outcomes. For more general random objects in geodesic metric spaces, [Kurusu et al. \(2024\)](#) developed a framework for geodesic average treatment effects, and [Zhou et al. \(2025\)](#) proposed geodesic difference-in-differences. These works demonstrate the promise of causal methods for object-valued outcomes, but a general synthetic control framework for a broad class of geodesic metric spaces along with consistency theory and intrinsic inference has been missing.

In this paper we propose such a framework, the geodesic synthetic control method (GSC), for causal inference with outcomes taking values in a geodesic metric space. The proposed method replaces Euclidean weighted averages by weighted Fréchet means and defines treatment effects through geodesics connecting untreated and treated potential outcomes. This provides an intrinsic extension of the classical SCM to outcomes such

as distributions, compositions, networks, SPD matrices, and functions. For distributional outcomes equipped with the Wasserstein metric the proposed framework includes the distributional synthetic control approach of [Gunsilius \(2023\)](#) as a special case, and encompasses a much wider class of geodesic metrics for distributions such as the Fisher–Rao metric; importantly, it extends beyond distributional data to a much wider class of random objects.

For outcomes in general geodesic metric spaces, we develop statistical theory and inference for GSC under a perturbed geodesic causal model, where untreated object-valued outcomes are generated by random perturbation maps acting on latent unit-specific objects. This formulation accommodates idiosyncratic time- and unit-specific variation and reduces, in the Euclidean case, to a latent-factor-type model with additive noise ([Abadie et al., 2010](#)). We establish consistency of the estimated GSC weights for the minimizer set of the limiting perturbed criterion, derive a post-treatment prediction-error bound, and give sufficient conditions for consistent recovery of the untreated counterfactual. For uncertainty quantification, we propose an intrinsic inference procedure based on metric conformal calibration, using pre-treatment prediction errors measured directly in the outcome metric to construct prediction regions for untreated counterfactual objects, confidence sets for geodesic treatment effects, and a global no-effect test. We assess the method through simulations for networks and symmetric positive-definite matrices and illustrate it in two empirical applications: the effect of the 2011 Great East Japan Earthquake on employment composition in Miyagi Prefecture and the effect of the 1972 abortion legislation on age-specific fertility patterns in East Germany.

The paper is organized as follows. Section [2](#) reviews basic concepts from metric geometry and introduces the geodesic metric spaces used throughout the paper. Section [3](#) presents the GSC estimator, the perturbed causal model and theory. Section [4](#) develops metric conformal calibration for inference. Section [5](#) provides implementation details and network simulations; parallel results for SPD matrices are in Section S.5 of the Supple-

ment. Section 6 presents the empirical applications to employment composition in Japan and fertility patterns in East Germany. Section 7 discusses extensions, including augmented GSC and geodesic synthetic difference-in-differences, and concludes. Proofs and additional technical details are provided in the Supplementary Material.

2 Preliminaries on metric geometry

Let $(\mathcal{M}, d)$ be a separable metric space such that $(\mathcal{M}, \mathcal{B}(\mathcal{M}))$ is a standard Borel space, where $\mathcal{B}(\mathcal{M})$ is the Borel σ -field generated by the metric topology. Let $(\Omega, \mathcal{A}, \Pr)$ be a probability space. An $\mathcal{M}$ -valued random object is a measurable map $Y : (\Omega, \mathcal{A}) \rightarrow (\mathcal{M}, \mathcal{B}(\mathcal{M}))$. The random object Y has a bounded essential range if there exist $\nu_0 \in \mathcal{M}$ and $D < \infty$ such that $\Pr\{d(\nu_0, Y) \leq D\} = 1$. A collection $\mathcal{Y}$ of $\mathcal{M}$ -valued random objects has uniformly bounded essential ranges if the same ν_0 and D satisfy this condition for every $Y \in \mathcal{Y}$. Throughout, the latent signals and random-object outcomes considered below, including those in triangular arrays, are assumed jointly to have uniformly bounded essential ranges over all unit, time, and array indices.

We briefly review the metric-geometric concepts needed to study outcomes taking values in $\mathcal{M}$. A *curve* in $\mathcal{M}$ is a continuous map $\gamma : [0, 1] \rightarrow \mathcal{M}$, whose length is defined as $\ell(\gamma) = \sup \sum_{i=0}^{I-1} d(\gamma(t_i), \gamma(t_{i+1}))$, where the supremum is taken over all possible finite partitions $0 = t_0 \leq \dots \leq t_I = 1$. A curve γ is called a *geodesic* if it satisfies $d(\gamma(s), \gamma(t)) = |t - s|d(\gamma(0), \gamma(1))$ for all $s, t \in [0, 1]$. The metric space $\mathcal{M}$ is called a *geodesic space* if, for any two points $\alpha, \beta \in \mathcal{M}$, there exists a geodesic connecting them, denoted $\gamma_{\alpha, \beta}$. If this geodesic is unique for every pair of points in $\mathcal{M}$, the space is referred to as a *uniquely geodesic space* (Bridson and Haefliger, 1999).

Examples of uniquely geodesic spaces that frequently arise in applications are as follows.

Example 1 (Networks). *Consider the space of simple, undirected, weighted networks with m nodes. Each network can be uniquely represented by its graph Laplacian matrix. The*

space of graph Laplacians $\mathcal{L}$ serves as a natural framework for characterizing network structures (Kolaczyk and Csárdi, 2014; Severn et al., 2022; Zhou and Müller, 2022). When endowed with the Frobenius metric, defined as $d_F(L_1, L_2) = \|L_1 - L_2\|_F$ for $L_1, L_2 \in \mathcal{L}$, the space $\mathcal{L}$ becomes a uniquely geodesic space. In this setting, the geodesic between L_1 and L_2 is given by the linear interpolation $\gamma_{L_1, L_2}(t) = (1 - t)L_1 + tL_2$ for $t \in [0, 1]$.

Example 2 (Compositional Data). *Compositional data, represented by vectors of non-negative proportions summing to one, reside in the simplex $\Delta^{d-1} = \{\mathbf{y} \in \mathbb{R}^d : y_j \geq 0, \sum_{j=1}^d y_j = 1\}$. Adopting the square-root transformation provides a map from Δ^{d-1} to the positive orthant of the unit sphere $\mathcal{S}_+^{d-1}$ (Scealy and Welsh, 2011), where the induced geodesic distance on $\mathcal{S}_+^{d-1}$ is the arc-length distance $d_g(\mathbf{z}_1, \mathbf{z}_2) = \arccos(\mathbf{z}_1^\top \mathbf{z}_2)$. Let $\theta = \arccos(\mathbf{z}_1^\top \mathbf{z}_2)$ be the angle between $\mathbf{z}_1$ and $\mathbf{z}_2$ and $\|\cdot\|$ be the standard Euclidean norm. The geodesic connecting $\mathbf{z}_1$ and $\mathbf{z}_2$ is then*

$$\gamma_{\mathbf{z}_1, \mathbf{z}_2}(t) = \cos(t\theta)\mathbf{z}_1 + \sin(t\theta) \frac{\mathbf{z}_2 - (\mathbf{z}_1^\top \mathbf{z}_2)\mathbf{z}_1}{\|\mathbf{z}_2 - (\mathbf{z}_1^\top \mathbf{z}_2)\mathbf{z}_1\|}, \quad t \in [0, 1].$$

Example 3 (Symmetric Positive-Definite Matrices). *The space of $m \times m$ symmetric positive-definite (SPD) matrices, denoted by Sym_m^+ , plays a crucial role in various fields such as computer vision (Cherian and Sra, 2016), signal processing (Arnaudon et al., 2013) and neuroimaging (Dryden et al., 2009). Various metrics have been introduced to endow Sym_m^+ with a meaningful geometric structure. The Frobenius metric d_F induces unique geodesics that are linear interpolations, $\gamma_{A, B}(t) = (1 - t)A + tB$ for $A, B \in \text{Sym}_m^+$. The space Sym_m^+ also forms a uniquely geodesic space under various other metrics such as the Log-Euclidean metric (Arsigny et al., 2007), the power metric family (Dryden et al., 2009), or the Log-Cholesky metric (Lin, 2019). For instance, under the Log-Euclidean metric, the unique geodesic connecting $A, B \in \text{Sym}_m^+$ is $\gamma_{A, B}(t) = \exp\{(1 - t)\log A + t\log B\}$, where $\exp$ and $\log$ denote the matrix exponential and matrix logarithm, respectively.*

Example 4 (Probability Distributions). *Let $\mathcal{W}$ denote the space of univariate probability distributions on $\mathbb{R}$ with finite second moments, equipped with the 2-Wasserstein distance*

$d_W(\mu, \nu) = \left(\int_0^1 \{F_\mu^{-1}(s) - F_\nu^{-1}(s)\}^2 ds \right)^{1/2}$, where F_μ^{-1} and F_ν^{-1} are the quantile functions of the probability measures μ and ν , respectively. The Wasserstein space $\mathcal{W}$ is a complete and separable uniquely geodesic space (Ambrosio et al., 2008), with geodesics given by McCann’s interpolant (McCann, 1997): $\gamma_{\mu, \nu}(t) = \{\text{id} + t(F_\nu^{-1} \circ F_\mu - \text{id})\}_{\#}\mu$, $t \in [0, 1]$, where id denotes the identity map, F_μ is the cumulative distribution function of μ , and $\tau_{\#}\mu$ denotes the pushforward measure of μ by τ . In addition, several metrics that are computationally tractable for both univariate and multivariate distributions are available. One example is the Fisher–Rao metric $d_{FR}(\mu, \nu) = \arccos \left(\int_{\mathbb{R}} \sqrt{f_\mu(x)f_\nu(x)} dx \right)$, where f_μ and f_ν are the probability density functions of the measures μ and ν , respectively.

Example 5 (Functional Data). In functional data analysis, one considers data consisting of functions that are considered to be the realizations of a square integrable stochastic process, usually assumed to be defined on a finite interval. Such data are common in the social, physical and life sciences (Ramsay and Silverman, 2005; Hsing and Eubank, 2015; Wang et al., 2016). Typically, functional data are assumed to be situated in the Hilbert space $L^2(\mathcal{T})$ of square integrable functions, where $\mathcal{T}$ is a compact interval, most commonly a time domain. In contrast to the other metric spaces we consider in this paper, this space is a vector space and thus permits algebraic operations. It is equipped with the standard L^2 metric $d_{L^2}(g_1, g_2) = \left(\int_{\mathcal{T}} \{g_1(t) - g_2(t)\}^2 dt \right)^{1/2}$, making $L^2(\mathcal{T})$ a uniquely geodesic space, with the unique geodesic between g_1 and g_2 given by $\gamma_{g_1, g_2}(t) = (1 - t)g_1 + tg_2$, $t \in [0, 1]$.

3 Geodesic synthetic control methods

3.1 Construction of the estimator

We begin with the conventional SCM for Euclidean outcomes. Suppose outcomes $Y_{j,t} \in \mathbb{R}^q$ are observed for $J+1$ units over T time periods, indexed by $j = 1, \dots, J+1$ and $t = 1, \dots, T$. Let $T_0 > 1$ denote the last pre-treatment period, and assume that only unit $j = 1$ receives

the intervention for $t > T_0$. For unit j at time t , we denote potential outcomes under treatment and no treatment by $Y_{j,t}(I)$ and $Y_{j,t}(N)$, respectively. We observe $Y_{1,t} = Y_{1,t}(I)$ for $t > T_0$ and $Y_{j,t} = Y_{j,t}(N)$ otherwise. The causal effect of interest is

$$\tau_t = Y_{1,t}(I) - Y_{1,t}(N), \quad t = T_0 + 1, \dots, T.$$

In Euclidean settings, the synthetic control estimator constructs a convex combination of the control units that best reproduces the pre-treatment trajectory of the treated unit. Let $\Delta^{J-1} = \{\mathbf{w} = (w_2, \dots, w_{J+1})^\top \in \mathbb{R}^J : w_j \geq 0, \sum_{j=2}^{J+1} w_j = 1\}$ denote the probability simplex, and $\hat{\mathbf{w}} = (\hat{w}_2, \dots, \hat{w}_{J+1})^\top \in \Delta^{J-1}$ be the optimal weight vector. The synthetic control outcome for the treated unit is defined as

$$\hat{Y}_{1,t}^{(\text{SC})} = \sum_{j=2}^{J+1} \hat{w}_j Y_{j,t}, \quad t = T_0 + 1, \dots, T,$$

where the weights are chosen to minimize the mean squared pre-treatment discrepancy,

$$\hat{\mathbf{w}} \in \arg \min_{\mathbf{w} \in \Delta^{J-1}} \frac{1}{T_0} \sum_{t=1}^{T_0} \|Y_{1,t} - Y_{2:J+1,t}^{(\mathbf{w})}\|^2, \quad Y_{2:J+1,t}^{(\mathbf{w})} = \sum_{j=2}^{J+1} w_j Y_{j,t}.$$

The SCM estimator of the causal effect is then given by

$$\hat{\tau}_t^{(\text{SC})} = Y_{1,t} - \hat{Y}_{1,t}^{(\text{SC})}, \quad t = T_0 + 1, \dots, T.$$

The construction of SCM relies on two key components: the weighted mean of the control outcomes $Y_{2:J+1,t}^{(\mathbf{w})}$ and the difference $Y_{1,t} - \hat{Y}_{1,t}^{(\text{SC})}$. Both are natural in Euclidean space but are not directly available for non-Euclidean outcomes, where algebraic operations like addition, subtraction, or scalar multiplication are not defined.

For general metric spaces, the appropriate analogue of the weighted mean is the weighted Fréchet mean (Fréchet, 1948; Ginestet et al., 2012), defined for objects $Y_{2,t}, \dots, Y_{J+1,t} \in \mathcal{M}$ as

$$Y_{2:J+1,t}^{(\mathbf{w})} = \arg \min_{\nu \in \mathcal{M}} \sum_{j=2}^{J+1} w_j d^2(\nu, Y_{j,t}),$$

which reduces to the weighted Euclidean average when $\mathcal{M} = \mathbb{R}^q$ and d is the Euclidean distance. When the outcome space is geodesic, a suitable generalization for otherwise

unavailable differences exists: in the special case of Euclidean space, the difference $\beta - \alpha$ can be equivalently interpreted as movement along the straight line (geodesic) from α to β , which motivates defining causal effects via geodesics (Kurisu et al., 2024).

Accordingly, in a uniquely geodesic space $(\mathcal{M}, d)$ we define the causal effect of the intervention as the geodesic connecting the untreated and treated potential outcomes,

$$\tau_t = \gamma_{Y_{1,t}(N), Y_{1,t}(I)}, \quad t = T_0 + 1, \dots, T.$$

The uniqueness of geodesics ensures that τ_t is well defined. This definition preserves the interpretation of τ_t as encoding both the size and direction of change.

Remark 3.1. *As pointed out by a referee, one may alternatively represent the causal effect through the ordered pair of potential outcomes $(Y_{1,t}(N), Y_{1,t}(I))$ rather than the geodesic $\tau_t = \gamma_{Y_{1,t}(N), Y_{1,t}(I)}$. This ordered-pair representation contains the same endpoint information and would allow a more general formulation without requiring unique geodesics. We adopt the geodesic formulation because it provides a direct intrinsic analogue of the classical Euclidean treatment effect $Y_{1,t}(I) - Y_{1,t}(N)$: in Euclidean space, subtraction represents a displacement from the untreated outcome to the treated outcome, while in a uniquely geodesic metric space the geodesic represents the corresponding intrinsic movement, with its length giving the effect size and its path describing the direction of change. Moreover, the geodesic structure is also used in the causal model in Section 3.2, where geodesic interpolation between a time-specific reference object and a unit-specific latent object provides an analogue of the latent-factor structure in classical synthetic control (Abadie et al., 2010). The uniquely geodesic assumption is not overly restrictive for our purposes, as many spaces of practical interest are uniquely geodesic, including the spaces considered in Examples 1–5 as well as phylogenetic tree spaces equipped with the BHV metric (Billera et al., 2001).*

To estimate τ_t , we extend the weight construction from SCM using weighted Fréchet

means. The synthetic control outcome for the treated unit is

$$\hat{Y}_{1,t}^{(\text{GSC})} = \arg \min_{\nu \in \mathcal{M}} \sum_{j=2}^{J+1} \hat{w}_j d^2(\nu, Y_{j,t}),$$

with weights chosen as

$$\hat{\mathbf{w}} \in \arg \min_{\mathbf{w} \in \Delta^{J-1}} \frac{1}{T_0} \sum_{t=1}^{T_0} d^2(Y_{1,t}, Y_{2:J+1,t}^{(\mathbf{w})}), \quad Y_{2:J+1,t}^{(\mathbf{w})} = \arg \min_{\nu \in \mathcal{M}} \sum_{j=2}^{J+1} w_j d^2(\nu, Y_{j,t}). \quad (1)$$

The *geodesic synthetic control* (GSC) estimator of the treatment effect is then

$$\hat{\tau}_t^{(\text{GSC})} = \gamma_{\hat{Y}_{1,t}^{(\text{GSC})}, Y_{1,t}}, \quad t = T_0 + 1, \dots, T.$$

3.2 Causal model for the geodesic synthetic control method

To analyze the statistical properties of the GSC estimator, we introduce a causal model for the untreated potential outcomes that allows for stochastic perturbations. Each unit j is associated with a unit-specific latent object $U_j \in \mathcal{M}$, and the untreated outcome at time t is generated as

$$Y_{j,t}(N) = g_t(\mathcal{P}_{j,t}(U_j)), \quad j = 1, \dots, J+1, \quad t = 1, \dots, T, \quad (2)$$

where $g_t : \mathcal{M} \rightarrow \mathcal{M}$ is a deterministic measurable map and $\mathcal{P}_{j,t} : \mathcal{M} \rightarrow \mathcal{M}$ is a random perturbation map. The latent object U_j captures persistent unit-level heterogeneity, while $\mathcal{P}_{j,t}$ represents time- and unit-specific stochastic deviations around this latent structure, allowing the observed untreated outcomes to deviate from the latent geodesic signal. Throughout the asymptotic analysis, the latent objects $U_1, \dots, U_{J+1}$ are treated as fixed, and all probabilities and expectations are taken with respect to the perturbation maps.

For illustration, consider a geodesic interpolation,

$$g_t(\mathcal{P}_{j,t}(U_j)) = \gamma_{\mu_t, \mathcal{P}_{j,t}(U_j)}(\alpha_t), \quad \alpha_t \in (0, 1), \quad (3)$$

where $\mu_t \in \mathcal{M}$ is a time-specific reference object. In this model, the untreated outcome is obtained by moving from the common time-specific object μ_t toward the perturbed unit-specific object $\mathcal{P}_{j,t}(U_j)$ along the geodesic connecting them. For Euclidean outcomes with

the Euclidean metric, taking an additive perturbation $\mathcal{P}_{j,t}(U_j) = U_j + \varepsilon_{j,t}$ yields

$$Y_{j,t}(N) = (1 - \alpha_t)\mu_t + \alpha_t U_j + \alpha_t \varepsilon_{j,t},$$

which corresponds to a latent-factor-type model with additive idiosyncratic noise.

Figure 1 summarizes the causal framework. Its top panel shows the latent unit objects, while the lower panels show the observed pre- and post-treatment outcomes used to construct the GSC counterfactual and geodesic effect. Under model (2), the perturbation maps act on the latent objects before the time-specific map g_t ; they are omitted from the diagram for visual clarity.

The formulation in (2) places the latent unit characteristic directly in the outcome space $\mathcal{M}$. This can be relaxed by allowing the latent characteristic to originate in another metric space $\mathcal{M}'$ and enter the outcome space through a time-specific map. Specifically, one may consider $Y_{j,t}(N) = g_t(\mathcal{P}_{j,t}(U_{j,t}))$, where $U_{j,t} \in \mathcal{M}$ may be represented as $U_{j,t} = \Phi_t(V_j)$ for some $V_j \in \mathcal{M}'$ and measurable map $\Phi_t : \mathcal{M}' \rightarrow \mathcal{M}$. In the Euclidean scalar case, taking $\mathcal{M}' = \mathbb{R}^d$ and $\Phi_t(v) = \nu_t^\top v$ for some $\nu_t \in \mathbb{R}^d$ recovers the usual linear factor representation.

We next distinguish between two weighted Fréchet means. For a weight vector $\mathbf{w} \in \Delta^{J-1}$, define the perturbed synthetic outcome

$$\hat{g}_{2:J+1,t}^{(\mathbf{w})} := \arg \min_{\nu \in \mathcal{M}} \sum_{j=2}^{J+1} w_j d^2(\nu, g_t(\mathcal{P}_{j,t}(U_j))),$$

and the corresponding latent synthetic outcome

$$g_{2:J+1,t}^{(\mathbf{w})} := \arg \min_{\nu \in \mathcal{M}} \sum_{j=2}^{J+1} w_j d^2(\nu, g_t(U_j)).$$

The first quantity is the synthetic control constructed from the observed, perturbed control outcomes, used to estimate the GSC weights; the second is its latent-signal counterpart that is used in the identification condition.

For each time point t , let $\mathbb{W}_t^* = \left\{ \mathbf{w} \in \Delta^{J-1} : d(g_t(U_1), \hat{g}_{2:J+1,t}^{(\mathbf{w})}) = 0 \right\}$ denote the set of weights that exactly reconstruct the latent signal of the treated unit at time t . The

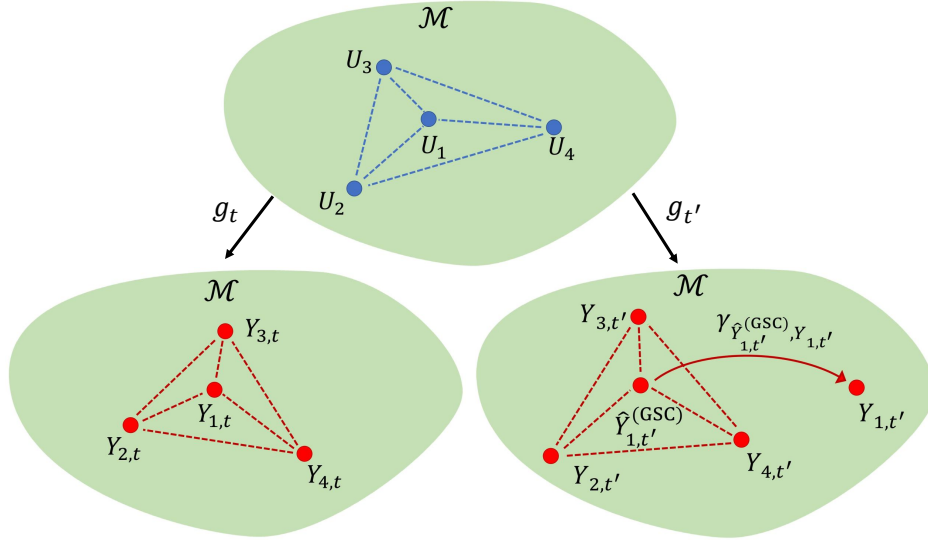


Figure 1: Causal framework for the geodesic synthetic control method. Circles represent random objects in the metric space $\mathcal{M}$. The top panel shows latent unit objects; under model (2), random perturbation maps act on these objects before g_t and are omitted from the diagram for visual clarity. For $t \leq T_0$, the treated outcome $Y_{1,t}$ is matched by a weighted Fréchet mean of controls. For $t' > T_0$, the estimator compares $\hat{Y}_{1,t'}^{(\text{GSC})}$ with the observed $Y_{1,t'}$, and the connecting geodesic represents the estimated causal effect.

following assumption ensures that the synthetic objects are well defined and formalizes exact reconstruction of the latent signal.

Assumption 3.1.

- (i) For each $\mathbf{w} \in \Delta^{J-1}$ and $t = 1, \dots, T$, the weighted Fréchet means $g_{2:J+1,t}^{(\mathbf{w})}$ and $\hat{g}_{2:J+1,t}^{(\mathbf{w})}$ exist uniquely.
- (ii) There exists a nonempty set $\mathcal{W}_0 \subset \Delta^{J-1}$, independent of T_0 , such that $\mathcal{W}_0 \subset \bigcap_{t=1}^{T_0} \mathbb{W}_t^*$ for all sufficiently large T_0 . Moreover, for every fixed $h \geq 1$, any pre-treatment latent balancing weight also reconstructs the treated unit's latent signal at $t = T_0 + h$, that is, $\bigcap_{t=1}^{T_0} \mathbb{W}_t^* \subset \mathbb{W}_{T_0+h}^*$ for all sufficiently large T_0 .

Assumption 3.1(i) ensures existence and uniqueness of the weighted Fréchet means; these properties hold broadly in Hadamard spaces (Sturm, 2003). Assumption 3.1(ii) requires a common set of weights to reconstruct the treated unit's latent trajectory before treatment and at each fixed post-treatment horizon. Section S.2 of the Supplement gives

sufficient conditions for this reconstruction under the signal models and metrics treated there.

We next consider estimation of the weights from perturbed pre-treatment outcomes. Throughout the asymptotic analysis, J is fixed and $T_0 \rightarrow \infty$, and post-treatment conclusions concern $t = T_0 + h$ for fixed $h \geq 1$. The conclusions hold jointly over a fixed finite post-treatment horizon when the stability condition below holds uniformly over that horizon.

The GSC weights are estimated from the perturbed pre-treatment outcomes by

$$\hat{\mathbf{w}} \in \arg \min_{\mathbf{w} \in \Delta^{J-1}} \hat{Q}(\mathbf{w}), \quad \hat{Q}(\mathbf{w}) = \frac{1}{T_0} \sum_{t=1}^{T_0} d^2(Y_{1,t}, Y_{2:J+1,t}^{(\mathbf{w})}),$$

where $Y_{2:J+1,t}^{(\mathbf{w})}$ denotes the weighted Fréchet mean of the observed control outcomes at time t . Under (2), $Y_{2:J+1,t}^{(\mathbf{w})} = \hat{g}_{2:J+1,t}^{(\mathbf{w})}$ for $t \leq T_0$. To describe the population target of this empirical criterion, let

$$\tilde{Q}(\mathbf{w}) = \lim_{T_0 \rightarrow \infty} \mathbb{E}[\hat{Q}(\mathbf{w})], \quad \tilde{\mathcal{W}} = \arg \min_{\mathbf{w} \in \Delta^{J-1}} \tilde{Q}(\mathbf{w}), \quad (4)$$

where $\tilde{\mathcal{W}}$ is the possibly non-singleton minimizer set of the limiting perturbed criterion.

For a nonempty set $A \subset \Delta^{J-1}$, write $d_1(\mathbf{w}, A) = \inf_{\mathbf{v} \in A} \|\mathbf{w} - \mathbf{v}\|_1$.

For $x \in \mathbb{R}^p$, let $\|x\|_1 = \sum_{j=1}^p |x_j|$ denote the ℓ_1 norm. Under the standing bounded-range condition, the following assumptions ensure uniform convergence of the empirical perturbed criterion and consistency of the estimated weights for the perturbed population target.

Assumption 3.2.

- (i) *The limiting criterion $\tilde{Q}$ in (4) exists as a finite continuous function, $\tilde{\mathcal{W}}$ is nonempty, and, for every $\varepsilon > 0$,*

$$\inf_{\{\mathbf{w} \in \Delta^{J-1}: d_1(\mathbf{w}, \tilde{\mathcal{W}}) \geq \varepsilon\}} \tilde{Q}(\mathbf{w}) > \inf_{\mathbf{w} \in \Delta^{J-1}} \tilde{Q}(\mathbf{w}).$$

A measurable minimizer $\hat{\mathbf{w}}$ of $\hat{Q}$ exists and, for every $\mathbf{w} \in \Delta^{J-1}$, $\hat{Q}(\mathbf{w}) \xrightarrow{p} \tilde{Q}(\mathbf{w})$ as $T_0 \rightarrow \infty$.

(ii) For every $\varepsilon > 0$,

$$\lim_{\delta \downarrow 0} \limsup_{T_0 \rightarrow \infty} \Pr \left\{ \sup_{\substack{\mathbf{w}_1, \mathbf{w}_2 \in \Delta^{J-1} \\ \|\mathbf{w}_1 - \mathbf{w}_2\|_1 < \delta}} \frac{1}{T_0} \sum_{t=1}^{T_0} d \left(\widehat{g}_{2:J+1,t}^{(\mathbf{w}_1)}, \widehat{g}_{2:J+1,t}^{(\mathbf{w}_2)} \right) > \varepsilon \right\} = 0.$$

Assumption 3.2(i) provides pointwise convergence and separation of the limiting criterion, while part (ii) imposes uniform local stability of the weighted Fréchet mean map. Together with the standing bounded-range condition in Section 2, these conditions yield the stochastic equicontinuity and uniform convergence of $\widehat{Q}$ used to establish consistency of the estimated weights for $\widetilde{\mathcal{W}}$; see [van der Vaart and Wellner \(2023\)](#) for the underlying empirical-process and M-estimation arguments. Section S.3 of the Supplement gives assumption-by-assumption sufficient conditions for Assumption 3.2 in the outcome spaces considered here. In particular, stationarity and ergodicity, or suitable weak-dependence conditions, provide a law of large numbers for the criterion in part (i), while Hilbert-coordinate or geodesic-convexity arguments establish part (ii).

Theorem 3.1. *Suppose that Assumption 3.1(i) and Assumption 3.2 hold. Then*

(i) As $T_0 \rightarrow \infty$,

$$d_1(\widehat{\mathbf{w}}, \widetilde{\mathcal{W}}) = o_p(1).$$

Since $\widetilde{\mathcal{W}}$ is nonempty and compact, choose a measurable nearest-point selection $\widetilde{\mathbf{w}} \in \widetilde{\mathcal{W}}$ satisfying $\|\widehat{\mathbf{w}} - \widetilde{\mathbf{w}}\|_1 = d_1(\widehat{\mathbf{w}}, \widetilde{\mathcal{W}})$. For a fixed $h \geq 1$, suppose, in addition, that the post-treatment stability condition

$$\sup_{\substack{\mathbf{w}_1, \mathbf{w}_2 \in \Delta^{J-1} \\ \mathbf{w}_1 \neq \mathbf{w}_2}} \frac{d(\widehat{g}_{2:J+1,T_0+h}^{(\mathbf{w}_1)}, \widehat{g}_{2:J+1,T_0+h}^{(\mathbf{w}_2)})}{\|\mathbf{w}_1 - \mathbf{w}_2\|_1^\kappa} = O_p(1)$$

holds for some fixed $\kappa \in (0, 1]$. Then $d(\widehat{g}_{2:J+1,T_0+h}^{(\widehat{\mathbf{w}})}, \widehat{g}_{2:J+1,T_0+h}^{(\widetilde{\mathbf{w}})}) = o_p(1)$ and, with $t = T_0 + h$,

$$\begin{aligned} & d(Y_{1,t}(N), \widehat{Y}_{1,t}^{(\text{GSC})}) \\ & \leq d(g_t(\mathcal{P}_{1,t}(U_1)), g_t(U_1)) + d(g_t(U_1), g_{2:J+1,t}^{(\widetilde{\mathbf{w}})}) + d(g_{2:J+1,t}^{(\widetilde{\mathbf{w}})}, \widehat{g}_{2:J+1,t}^{(\widetilde{\mathbf{w}})}) + o_p(1). \end{aligned} \quad (5)$$

(ii) Suppose, in addition, that Assumption 3.1(ii) holds and $\widetilde{\mathcal{W}} \subset \bigcap_{s=1}^{T_0} \mathbb{W}_s^*$ for all sufficiently large T_0 . Then, for every fixed $h \geq 1$,

$$d(g_{T_0+h}(U_1), g_{2:J+1, T_0+h}^{(\tilde{\mathbf{w}})}) = 0$$

for all sufficiently large T_0 .

Theorem 3.1 establishes consistency of the estimated weights for $\widetilde{\mathcal{W}}$ and decomposes the post-treatment prediction error into the treated-outcome perturbation, latent-signal imbalance, synthetic-control perturbation, and an asymptotically negligible weight-estimation term. Part (ii) eliminates the latent-signal term when every population-target weight also balances the pre-treatment latent signal and this balance persists after treatment. Corollary S1 of the Supplement gives sufficient conditions for consistent recovery of the untreated counterfactual.

Remark 3.2 (*Geodesic synthetic control method with covariates*). Suppose covariates $X_{j,t} \in \mathcal{M}_X$, $j = 1, \dots, J+1$ and $t = 1, \dots, T_0$, are available, where $(\mathcal{M}_X, d_X)$ may differ from the outcome space $(\mathcal{M}, d)$. Covariates can be incorporated by augmenting the pre-treatment matching criterion. For example, for a scale parameter $\eta \geq 0$, one may obtain weights

$$\widehat{\mathbf{w}}(\eta) \in \arg \min_{\mathbf{w} \in \Delta^{J-1}} \left\{ \frac{1}{T_0} \sum_{t=1}^{T_0} d^2(Y_{1,t}, Y_{2:J+1,t}^{(\mathbf{w})}) + \eta \frac{1}{T_0} \sum_{t=1}^{T_0} d_X^2(X_{1,t}, X_{2:J+1,t}^{(\mathbf{w})}) \right\},$$

where

$$Y_{2:J+1,t}^{(\mathbf{w})} = \arg \min_{\nu \in \mathcal{M}} \sum_{j=2}^{J+1} w_j d^2(\nu, Y_{j,t}), \quad X_{2:J+1,t}^{(\mathbf{w})} = \arg \min_{\nu \in \mathcal{M}_X} \sum_{j=2}^{J+1} w_j d_X^2(\nu, X_{j,t}).$$

The corresponding covariate-adjusted GSC estimator is

$$\widehat{Y}_{1,t}^{(\text{COV})} = \arg \min_{\nu \in \mathcal{M}} \sum_{j=2}^{J+1} \widehat{w}_j(\eta) d^2(\nu, Y_{j,t}), \quad \widehat{\tau}_t^{(\text{COV})} = \gamma_{\widehat{Y}_{1,t}^{(\text{COV})}, Y_{1,t}}, \quad t = T_0 + 1, \dots, T.$$

This formulation balances both pre-treatment outcomes and covariates and reduces to the usual Euclidean SCM construction when both spaces are Euclidean. If multiple covariates are available, one may work in the product space $\mathcal{M}_X = \mathcal{M}_1 \times \dots \times \mathcal{M}_K$ with product

metric $d_X^2 = d_1^2 + \dots + d_K^2$. Unlike in the Euclidean setting, object-valued outcomes and covariates need not admit a common vector representation, so the metric-space formulation balances each component through its own intrinsic geometry.

4 Inference by metric conformal calibration

We next develop an inference procedure that calibrates prediction errors over time for the treated unit. The procedure is inspired by counterfactual prediction and conformal inference ideas for Euclidean synthetic control methods (Chernozhukov et al., 2021b), but is formulated intrinsically in the metric space and does not require linear structure.

4.1 Pointwise prediction regions and tests

We first consider pointwise inference for the untreated counterfactual outcome $Y_{1,t}(N)$ at a fixed post-treatment time $t > T_0$. The main idea is to use pre-treatment prediction errors of the treated unit to calibrate the uncertainty of the GSC counterfactual prediction. To avoid using in-sample residuals for calibration, we split the pre-treatment periods into a training set and a calibration set.

Let $\mathcal{T}_{\text{tr}} \subset \{1, \dots, T_0\}$ and $\mathcal{T}_{\text{cal}} \subset \{1, \dots, T_0\} \setminus \mathcal{T}_{\text{tr}}$ denote the training and calibration periods, respectively. Using only the training periods, define the GSC training weights

$$\hat{\mathbf{w}}^{\text{tr}} \in \arg \min_{\mathbf{w} \in \Delta^{J-1}} \frac{1}{|\mathcal{T}_{\text{tr}}|} \sum_{s \in \mathcal{T}_{\text{tr}}} d^2(Y_{1,s}, Y_{2:J+1,s}^{(\mathbf{w})}),$$

where, for any $\mathbf{w} = (w_2, \dots, w_{J+1})^\top \in \Delta^{J-1}$,

$$Y_{2:J+1,s}^{(\mathbf{w})} = \arg \min_{\nu \in \mathcal{M}} \sum_{j=2}^{J+1} w_j d^2(\nu, Y_{j,s}).$$

To distinguish the training-based prediction from the full-pre-treatment GSC prediction in Section 3, write

$$\hat{Y}_{1,s}^{(\text{GSC}, \text{tr})} = \arg \min_{\nu \in \mathcal{M}} \sum_{j=2}^{J+1} \hat{w}_j^{\text{tr}} d^2(\nu, Y_{j,s}), \quad s = 1, \dots, T.$$

For any time s , define the latent no-treatment prediction error $R_s = d(Y_{1,s}(N), \hat{Y}_{1,s}^{(\text{GSC}, \text{tr})})$. This score is intrinsic: it is a real-valued discrepancy measured directly by the metric d on $\mathcal{M}$, and does not require subtraction of random objects or a linear embedding. The residual geodesic from $\hat{Y}_{1,s}^{(\text{GSC}, \text{tr})}$ to $Y_{1,s}(N)$ contains directional information, but residual geodesics with different endpoints have no canonical ordering in a general geodesic metric space and therefore cannot be ranked directly for conformal calibration. We use their lengths for calibration and retain directional uncertainty through the geodesic effect set defined below. For pre-treatment periods $s \leq T_0$, the treated unit is untreated, so $Y_{1,s}(N) = Y_{1,s}$ and the score is observed. For post-treatment periods $s > T_0$, the score involves the unobserved counterfactual $Y_{1,s}(N)$ and is therefore not observed unless a sharp null hypothesis is imposed.

We use the calibration scores $R_s = d(Y_{1,s}, \hat{Y}_{1,s}^{(\text{GSC}, \text{tr})})$, $s \in \mathcal{T}_{\text{cal}}$, to estimate the typical size of the GSC prediction error under no treatment. Let $m = |\mathcal{T}_{\text{cal}}|$, $k_\alpha = \lceil (m+1)(1-\alpha) \rceil$, and $R_{(k)}$ denote the k -th order statistic of $\{R_s : s \in \mathcal{T}_{\text{cal}}\}$. Define $q_{1-\alpha} = R_{(k_\alpha)}$, with the convention that $q_{1-\alpha} = +\infty$ if $k_\alpha > m$. For a post-treatment period $t > T_0$, the $(1-\alpha)$ prediction region for the untreated counterfactual object is

$$\mathcal{C}_t^{(\text{GSC})}(1-\alpha) = \left\{ \omega \in \mathcal{M} : d(\omega, \hat{Y}_{1,t}^{(\text{GSC}, \text{tr})}) \leq q_{1-\alpha} \right\}.$$

Thus $\mathcal{C}_t^{(\text{GSC})}(1-\alpha)$ is the closed metric ball centered at the GSC counterfactual prediction $\hat{Y}_{1,t}^{(\text{GSC}, \text{tr})}$ with radius calibrated by held-out pre-treatment prediction errors. Large values of $d(Y_{1,t}, \hat{Y}_{1,t}^{(\text{GSC}, \text{tr})})$ relative to this calibration distribution provide evidence that the observed treated outcome is incompatible with the no-treatment counterfactual trajectory.

Assumption 4.1 (Exchangeability of calibration and target scores). *Fix a post-treatment time point $t \in \{T_0 + 1, \dots, T\}$. Conditional on the data used to estimate the GSC weights, the calibration scores $\{R_s : s \in \mathcal{T}_{\text{cal}}\}$ together with the target score R_t are exchangeable.*

Assumption 4.1 concerns temporal invariance of the prediction-error scores under the no-treatment law, conditional on the training data. It does not require independence across

units. The following result gives concrete sufficient conditions in terms of the perturbation vectors and the time maps in model (2).

Proposition 4.1 (Sufficient conditions for score exchangeability). *Let $\mathcal{F}_{\text{tr}}$ denote the σ -field generated by the data used to estimate $\widehat{\mathbf{w}}^{\text{tr}}$, and, for each calibration or target period s , define $Z_{j,s} = \mathcal{P}_{j,s}(U_j)$ and*

$$Z_s = (Z_{1,s}, \dots, Z_{J+1,s}).$$

Suppose that, conditional on $\mathcal{F}_{\text{tr}}$, the vectors Z_s are exchangeable over the calibration and target periods. Further assume that either $g_s = g$ throughout these periods or $g_s = h_s \circ g$, where each $h_s : \mathcal{M} \rightarrow \mathcal{M}$ is a deterministic bijective isometry applied to every unit, and that the relevant weighted Fréchet mean map is unique and measurable. Then Assumption 4.1 holds.

This proposition permits arbitrary contemporaneous dependence among the components of each Z_s . For calibration and target blocks of equal length, conditional exchangeability of the corresponding block vectors formed from $\{Z_s\}$ similarly implies Assumption 4.2 below. In particular, conditional independence and identical distribution of Z_s across the relevant periods are sufficient for both assumptions. Exchangeability of the perturbation vectors alone, however, is insufficient if unrestricted temporal variation in g_s changes the distribution of the resulting scores. When exact exchangeability is implausible, one may instead impose stationarity and weak-dependence conditions directly on the metric prediction-error process and use moving, overlapping, or blockwise calibration schemes, following Chernozhukov et al. (2021b). Under the corresponding temporal conditions, these procedures provide approximate or asymptotic validity rather than the exact finite-sample guarantee established in Theorem 4.1.

Theorem 4.1 (Finite-sample pointwise coverage). *Suppose Assumption 4.1 holds. Then, for any $t \in \{T_0 + 1, \dots, T\}$, $\Pr \left\{ Y_{1,t}(N) \in \mathcal{C}_t^{(\text{GSC})}(1 - \alpha) \right\} \geq 1 - \alpha$. Consequently, under*

the pointwise null hypothesis of no treatment effect $H_{0,t} : Y_{1,t}(I) = Y_{1,t}(N)$, the test

$$\phi_t = 1 \left\{ Y_{1,t} \notin \mathcal{C}_t^{(\text{GSC})}(1 - \alpha) \right\} = 1 \left\{ d \left(Y_{1,t}, \hat{Y}_{1,t}^{(\text{GSC}, \text{tr})} \right) > q_{1-\alpha} \right\}$$

has finite-sample size at most α .

The pointwise test can equivalently be reported through the conformal rank p -value

$$\hat{p}_t = \frac{1 + \sum_{s \in \mathcal{T}_{\text{cal}}} 1\{R_s \geq d(Y_{1,t}, \hat{Y}_{1,t}^{(\text{GSC}, \text{tr})})\}}{m + 1},$$

with smaller values providing stronger evidence against $H_{0,t}$. Under $H_{0,t}$, Assumption 4.1 implies $\Pr(\hat{p}_t \leq \alpha) \leq \alpha$; the non-strict comparison is conservative in the presence of ties.

Theorem 4.1 provides a finite-sample valid prediction region for the untreated counterfactual random object $Y_{1,t}(N)$. Since the causal estimand in our framework is the geodesic $\tau_t = \gamma_{Y_{1,t}(N), Y_{1,t}(I)}$, the prediction region for $Y_{1,t}(N)$ naturally induces uncertainty quantification for the geodesic treatment effect itself. Specifically, for $t > T_0$, define

$$\mathcal{G}_t(1 - \alpha) = \left\{ \gamma_{\omega, Y_{1,t}} : \omega \in \mathcal{C}_t^{(\text{GSC})}(1 - \alpha) \right\}.$$

Because the observed post-treatment outcome satisfies $Y_{1,t} = Y_{1,t}(I)$ for $t > T_0$, the set $\mathcal{G}_t(1 - \alpha)$ may be interpreted as a confidence set for the geodesic causal effect.

Corollary 4.1 (Confidence set for the geodesic effect). *Suppose the assumptions of Theorem 4.1 hold. Then, for any $t \in \{T_0 + 1, \dots, T\}$, $\Pr \{ \tau_t \in \mathcal{G}_t(1 - \alpha) \} \geq 1 - \alpha$.*

In addition to the geodesic effect itself, a natural scalar summary is the magnitude of the treatment effect, $\delta_t = d(Y_{1,t}(N), Y_{1,t}(I))$, which corresponds to the geodesic distance between the untreated and treated potential outcomes. The corresponding estimator is $\hat{\delta}_t = d(\hat{Y}_{1,t}^{(\text{GSC}, \text{tr})}, Y_{1,t})$. The conformal prediction region immediately yields a finite-sample confidence interval for δ_t .

Corollary 4.2 (Confidence interval for effect size). *Suppose the assumptions of Theorem 4.1 hold. Then, for any $t \in \{T_0 + 1, \dots, T\}$, $\Pr \left\{ \delta_t \in \left[\{\hat{\delta}_t - q_{1-\alpha}\}_+, \hat{\delta}_t + q_{1-\alpha} \right] \right\} \geq 1 - \alpha$, where $\{a\}_+ = \max(a, 0)$.*

Corollary 4.2 shows that uncertainty quantification for the untreated counterfactual outcome directly translates into uncertainty quantification for the magnitude of the geodesic treatment effect through the metric structure of the outcome space. Unlike for the special case of Euclidean synthetic control methods, the proposed inference procedure is formulated intrinsically in the metric space.

4.2 A global no-effect test

We now develop a global test for detecting whether the treatment induces any departure from the untreated trajectory over the entire post-treatment period. Specifically, consider the null hypothesis

$$H_0^{\text{glob}} : Y_{1,t}(I) = Y_{1,t}(N), \quad t = T_0 + 1, \dots, T.$$

Under H_0^{glob} , the observed post-treatment trajectory coincides with the untreated potential trajectory, so deviations between the observed treated unit and its geodesic synthetic control should resemble the treated unit's own pre-treatment prediction errors.

Let $\mathcal{B}_* = \{T_0 + 1, \dots, T\}$ denote the post-treatment block and let $T_1 = T - T_0$ denote the number of post-treatment periods. To calibrate the post-treatment prediction error, partition the remaining pre-treatment periods into disjoint calibration blocks $\mathcal{B}_1, \dots, \mathcal{B}_K \subset \{1, \dots, T_0\} \setminus \mathcal{T}_{\text{tr}}$, each of length T_1 . For a block $\mathcal{B}$ and $r \in [1, \infty)$, define the block score

$$S_r(\mathcal{B}) = \left\{ \frac{1}{|\mathcal{B}|} \sum_{t \in \mathcal{B}} d^r \left(Y_{1,t}, \hat{Y}_{1,t}^{(\text{GSC}, \text{tr})} \right) \right\}^{1/r},$$

and for $r = \infty$, define $S_\infty(\mathcal{B}) = \max_{t \in \mathcal{B}} d \left(Y_{1,t}, \hat{Y}_{1,t}^{(\text{GSC}, \text{tr})} \right)$.

The choice of r determines the sensitivity of the test to different alternatives. The statistic S_∞ is particularly sensitive to large but localized treatment effects, whereas smaller values such as $r = 1$ or $r = 2$ are more sensitive to persistent deviations from the synthetic counterfactual trajectory.

Under H_0^{glob} , the post-treatment block score is observable and given by

$$S_r(\mathcal{B}_*) = \left\{ \frac{1}{T_1} \sum_{t=T_0+1}^T d^r \left(Y_{1,t}, \hat{Y}_{1,t}^{(\text{GSC}, \text{tr})} \right) \right\}^{1/r},$$

with the obvious modification when $r = \infty$. We then define the conformal p -value

$$\hat{p}_r = \frac{1 + \sum_{k=1}^K 1\{S_r(\mathcal{B}_k) \geq S_r(\mathcal{B}_*)\}}{K + 1}.$$

Assumption 4.2 (Exchangeability of block scores under no treatment). *Conditional on the data used to estimate the GSC weights, and under H_0^{glob} , the block scores $S_r(\mathcal{B}_1), \dots, S_r(\mathcal{B}_K), S_r(\mathcal{B}_*)$ are exchangeable.*

Assumption 4.2 requires that, under the no-treatment law, the treated unit's metric prediction errors remain distributionally invariant at the block level. As in the pointwise setting, this condition concerns temporal stability and does not require exchangeability or independence across units. Proposition 4.1 gives sufficient conditions.

Theorem 4.2 (Finite-sample validity of the global test). *Suppose Assumption 4.2 holds. Then the test $\phi_r^{\text{glob}} = 1\{\hat{p}_r \leq \alpha\}$ satisfies $\Pr_{H_0^{\text{glob}}} \{\phi_r^{\text{glob}} = 1\} \leq \alpha$.*

Theorem 4.2 establishes finite-sample validity of the proposed global test. The procedure may be interpreted as an end-of-sample stability test for the treated unit's intrinsic prediction errors. Rejection of H_0^{glob} indicates that the post-treatment trajectory lies systematically farther from its geodesic synthetic counterfactual than would be expected based on the treated unit's own pre-treatment prediction errors.

Remark 4.1 (In-time placebo diagnostic). *An important diagnostic for synthetic control analyses is an in-time placebo, or backdating, check (Abadie et al., 2015). Choose a pseudo-treatment time $\tilde{T}_0 < T_0$, estimate the GSC weights using only periods $t \leq \tilde{T}_0$, and construct pseudo-post-treatment synthetic controls for $\tilde{T}_0 < t \leq T_0$. Since no intervention occurred during this period, the discrepancies $d(Y_{1,t}, \hat{Y}_{1,t}^{(\text{GSC}, \tilde{T}_0)})$, $\tilde{T}_0 < t \leq T_0$, should be small if*

the counterfactual prediction is stable. Large pseudo-post-treatment discrepancies indicate potential lack of fit, anticipation effects, or instability of the donor pool. This diagnostic is the metric analogue of the in-time placebo check commonly used in Euclidean SCM, with the usual residuals replaced by intrinsic metric discrepancies.

5 Implementation and simulations

5.1 Implementation details

Algorithm 1 outlines the implementation steps for the proposed geodesic synthetic control method. The optimization problem in Step 1 of Algorithm 1 requires customized approaches that depend on the specific metric and geometry of the underlying metric space. For networks (Example 1), SPD matrices (Example 3), distributions with the Wasserstein metric (Example 4), and functions (Example 5), the optimization reduces to a straightforward convex quadratic optimization problem that can be easily implemented (Stellato et al., 2020).

For the case of compositional data (Example 2), the optimization becomes more challenging, as it involves minimizing a bilevel objective where the weighted Fréchet mean $Y_{2:J+1,t}^{(w)}$ is itself an argmin problem. Due to the complexity of analytical differentiation when nonlinear Fréchet means are involved, we adopt a derivative-free constrained optimization strategy. Specifically, we use the Constrained Optimization BY Linear Approximations (COBYLA) algorithm (Powell, 1994), as implemented in the R package `nloptr` (Johnson, 2008). The weighted Fréchet mean involved in the objective function is obtained using the R package `manifold` (Dai et al., 2021). Further computational details, including explicit formulations of the optimization problems and solver strategies across different geodesic metric spaces, are provided in Section S.1 of the Supplement.

Algorithm 1: Geodesic Synthetic Control Method

Input: data $\{Y_{j,t} : j = 1, \dots, J+1, t = 1, \dots, T\}$.

Output: geodesic synthetic control estimator $\hat{\tau}_t^{(\text{GSC})}$ for $t = T_0 + 1, \dots, T$.

1 $\hat{\mathbf{w}} = (\hat{w}_2, \dots, \hat{w}_{J+1}) \leftarrow$ the estimated synthetic control weights:

$$\hat{\mathbf{w}} \in \arg \min_{\mathbf{w} \in \Delta^{J-1}} \frac{1}{T_0} \sum_{t=1}^{T_0} d^2(Y_{1,t}, Y_{2:J+1,t}^{(\mathbf{w})})$$

where $Y_{2:J+1,t}^{(\mathbf{w})} = \arg \min_{\nu \in \mathcal{M}} \sum_{j=2}^{J+1} w_j d^2(\nu, Y_{j,t})$ for $t = 1, \dots, T_0$;

2 $\hat{Y}_{1,t}^{(\text{GSC})} = \arg \min_{\nu \in \mathcal{M}} \sum_{j=2}^{J+1} \hat{w}_j d^2(\nu, Y_{j,t}) \leftarrow$ the synthetic control unit;

3 $\hat{\tau}_t^{(\text{GSC})} = \gamma_{\hat{Y}_{1,t}^{(\text{GSC})}, Y_{1,t}} \leftarrow$ the geodesic synthetic control estimator.

5.2 Simulations for random-object outcomes

We evaluate network outcomes represented by graph Laplacians of weighted 10-node two-block networks under the Frobenius metric; parallel SPD results are reported in Section S.5 of the Supplement. The perturbed geodesic model (3) uses $\alpha_t = .5$, $J = 20$ donors, within- and between-community edge probabilities .75 and .10, and nonzero latent edge weights drawn from $\text{Unif}(.5, 1.5)$. We draw oracle weights $\mathbf{w}^* \in \Delta^{J-1}$, set the treated latent network to the corresponding donor combination, and use a common periodic reference network. Each nonzero edge weight is multiplied independently across edges, units, and time by $\exp(\sigma\xi)$, where $\xi \sim \text{Unif}[-\sqrt{3}, \sqrt{3}]$, and we set $\sigma = .10$.

For estimation, we use $T_0 \in \{20, 100, 500\}$, one post-treatment period, and 500 replications. Table 1 reports the post-treatment prediction gap $d(\hat{g}_{2:J+1,T_0+1}^{(\hat{\mathbf{w}})}, \hat{g}_{2:J+1,T_0+1}^{(\tilde{\mathbf{w}})})$, the realized post-treatment error $d(Y_{1,T_0+1}(N), \hat{g}_{2:J+1,T_0+1}^{(\tilde{\mathbf{w}})})$, the weight-estimation error $\|\hat{\mathbf{w}} - \tilde{\mathbf{w}}\|_1$, and the oracle-weight gap $\|\tilde{\mathbf{w}} - \mathbf{w}^*\|_1$, where $\tilde{\mathbf{w}}$ denotes the selected minimizer of $\tilde{Q}$ under the current simulation design. The post-treatment prediction gap and weight-estimation error decrease with T_0 , whereas the realized post-treatment error and oracle-weight gap remain stable, consistent with the decomposition in Theorem 3.1. These patterns have direct analogues in Euclidean synthetic control: idiosyncratic disturbances can lead the population criterion to select weights that differ from latent-balancing weights (Ferman and

Quantity	$T_0 = 20$	$T_0 = 100$	$T_0 = 500$
Post-treatment prediction gap	.068 (.001)	.031 (<.001)	.014 (<.001)
Realized post-treatment error	.374 (.003)	.373 (.003)	.367 (.003)
$\ \hat{\mathbf{w}} - \tilde{\mathbf{w}}\ _1$	.120 (.001)	.055 (.001)	.025 (<.001)
$\ \tilde{\mathbf{w}} - \mathbf{w}^*\ _1$	.063 (.001)	.062 (.001)	.062 (.001)

Table 1: Network estimation results for $T_0 \in \{20, 100, 500\}$ and $\sigma = .10$, based on 500 replications. The post-treatment prediction gap compares the post-treatment predictions based on $\hat{\mathbf{w}}$ and $\tilde{\mathbf{w}}$, while the realized post-treatment error compares the untreated post-treatment outcome with the prediction based on $\hat{\mathbf{w}}$. The remaining rows report the ℓ_1 weight-estimation error and oracle-weight gap. Entries are Monte Carlo means with Monte Carlo standard errors in parentheses.

Pinto, 2021), while a newly realized post-treatment disturbance remains in the counterfactual prediction error (Abadie et al., 2010). Increasing T_0 therefore improves estimation of the population target but need not eliminate the oracle-weight gap or newly realized post-treatment disturbances.

Section S.5 of the Supplement varies T_0 and σ separately, showing that longer pre-treatment periods reduce the weight-estimation error and post-treatment prediction gap and that larger perturbations increase realized post-treatment error.

For inference, $T_0 = 500$ is split into 200 training and 300 calibration periods, followed by five post-treatment periods. Treatment adds a times a normalized between-community edge Laplacian, giving effect magnitude $a \in \{0, .25, .50, .75, 1\}$. Pointwise inference uses the first post-treatment period; the global test uses 60 five-period calibration blocks with the block score set to $r = 2$. The common time component cancels from each prediction error because the weights sum to one, so the i.i.d. perturbations make the calibration and target scores exchangeable conditional on training. Under the null, pointwise coverage is .938 and pointwise and global rejection probabilities are .062 and .050. At $a = .25$, the corresponding powers are .362 and .858, and both reach one at $a = .50$; see Figure 2.

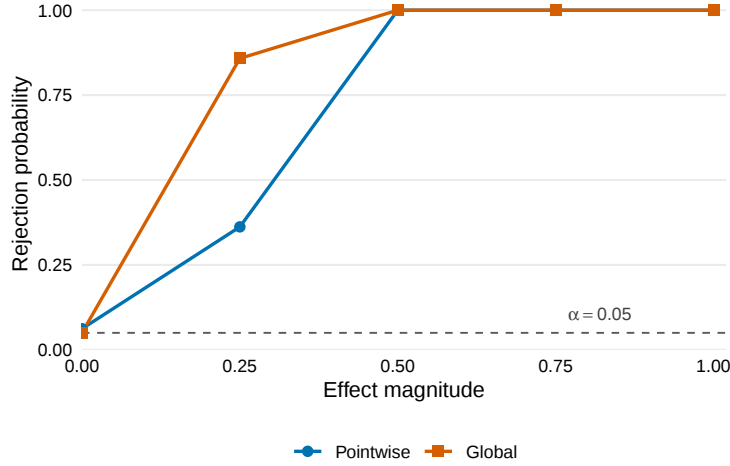


Figure 2: Empirical rejection probabilities for the pointwise and global tests in the network design over 500 replications. The dashed line marks the nominal level $\alpha = .05$; values at zero effect report empirical size.

6 Real data examples

6.1 Employment composition after the 2011 Great East Japan Earthquake

We analyze the impact of the 2011 Great East Japan Earthquake on the employment-population ratio by employment category, using data from the Research Institute of Economy, Trade and Industry (RIETI) database (<https://www.rieti.go.jp/jp/database/R-JIP2021/index.html>). These data provide annual employment figures by industry for each prefecture in Japan from 1994 to 2018. Employment is categorized into three broad groups: primary (agriculture, forestry, and fisheries), secondary (manufacturing, electricity, gas, water supply, waste management and construction) and tertiary (services, including wholesale/retail, finance, and healthcare). For each prefecture and each year, the employment shares (proportions) across these three categories are represented as a three-dimensional compositional vector. The component-wise square root transformation (see Example 2) maps these compositional outcomes onto the positive orthant of the sphere $\mathcal{S}_+^2 \subset \mathbb{R}^3$, which we endow with the geodesic metric. Unlike standard synthetic control

approaches that typically analyze each compositional component separately, the proposed approach considers the entire compositional vector jointly. This holistic perspective reflects the intrinsic geometry of compositional data, thereby capturing interdependencies of employment across the three categories.

The Great East Japan Earthquake occurred on March 11, 2011, causing widespread devastation across northeastern Japan, with Miyagi Prefecture experiencing particularly severe damage. The disaster substantially disrupted regional economic activity and induced major structural changes in employment. To evaluate the effect of the earthquake on employment composition, we treat Miyagi Prefecture as the treated unit. Since the earthquake constituted a broad nationwide shock that may also have affected nearby prefectures, we restrict the donor pool to prefectures located in Western Japan, which were geographically distant from the disaster and therefore less likely to have been directly exposed to its economic and social impacts. Specifically, the donor pool consists of the following 20 prefectures: Shiga, Kyoto, Nara, Wakayama, Tottori, Shimane, Okayama, Hiroshima, Yamaguchi, Tokushima, Kagawa, Ehime, Kochi, Fukuoka, Saga, Nagasaki, Kumamoto, Oita, Miyazaki, and Kagoshima. We analyze annual employment composition data from 1994 to 2013, using 1994–2010 as the pre-treatment period and 2011–2013 as the post-treatment period.

We apply the proposed GSC method to estimate the causal effect of the earthquake on employment composition in Miyagi Prefecture. Figure 3 displays the observed and synthetic post-treatment employment compositions using a ternary plot. Relative to the synthetic counterfactual trajectory, the observed post-treatment compositions exhibit a larger proportion of secondary category employment and a smaller proportion of tertiary category employment, while the primary category remains comparatively stable.

For comparison, we fit conventional SCM separately to the three employment shares. The distinct donor weights do not preserve the unit-sum constraint: the maximum absolute

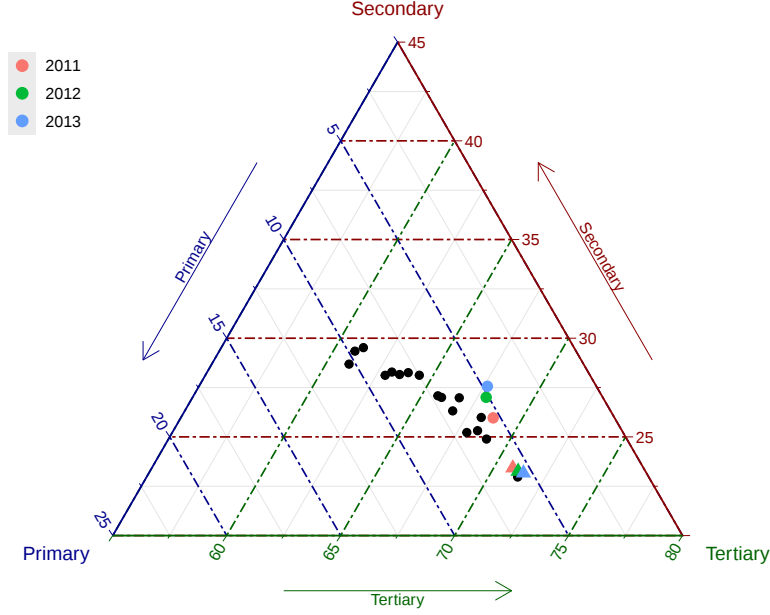


Figure 3: Ternary plot of employment composition in Miyagi Prefecture before and after the 2011 Great East Japan Earthquake. Black points represent the observed pre-treatment trajectory of Miyagi from 1994–2010. Colored circles denote the observed post-treatment compositions for 2011–2013, while colored triangles denote the corresponding geodesic synthetic counterfactuals, with colors indicating year. Relative to the synthetic counterfactual trajectory, the observed post-treatment compositions exhibit a higher proportion in the secondary employment category and a lower proportion in the tertiary category, while the primary category remains comparatively stable.

deviation is .005249 over the held-out pre-treatment period and .002813 after treatment. GSC enforces this constraint through a single joint objective; Section S.5 of the Supplement reports the componentwise predictions.

This shift is consistent with reconstruction-driven demand for secondary-category employment and prolonged disruption of service-sector activity. It suggests that the earthquake temporarily changed Miyagi’s employment composition toward reconstruction work.

Using 1994–1999 for training and 2000–2010 for calibration, the estimated effect magnitudes for 2011–2013 are .0421, .0561, and .0646, with 90% split-conformal intervals [.0066, .0776], [.0206, .0916], and [.0291, .1001]. Each conformal p -value is 1/12; 11 calibration periods do not yield a nontrivial 95% split-conformal interval.

6.2 Impact of the 1972 abortion legislation on fertility patterns in East Germany

The study of fertility rates has attracted substantial attention, with increasing emphasis on analyzing age-specific fertility rates (ASFR), encapsulating more information than summary measures such as the total fertility rate ([Chen et al., 2017](#)). Represented as fertility curves across reproductive ages, ASFR provides rich insights into fertility, allowing researchers to detect age-specific causal effects that aggregate statistics cannot capture.

In March 1972, East Germany introduced legislation permitting women to terminate pregnancies within the first twelve weeks, significantly liberalizing abortion access. Although the primary intention was reproductive autonomy, an unintended consequence was a substantial drop in fertility, with the total fertility rate decreasing notably from 2.1 children per woman in 1971 to 1.5 by 1975 ([Büttner and Lutz, 1990](#)). To gain deeper insights into how this policy affected fertility behavior across different reproductive ages, we applied the proposed GSC to ASFR curves. When viewed as functions over reproductive ages, these are functional data in the space L^2 (see [Example 5](#)).

The data for this analysis were obtained from the [Human Fertility Database](#), which provides ASFR annually for ages 12 to 55 for each calendar year across different countries. We select East Germany as the treated unit, given its distinct policy change, and utilize a control group composed of 20 countries that did not experience comparable abortion liberalization during this period (Austria, Belgium, Bulgaria, Canada, Switzerland, Czechia, Denmark, Spain, Finland, France, Hungary, Ireland, Italy, Japan, Netherlands, Portugal, Slovakia, Sweden, United Kingdom (England and Wales), and United States). We define 1956–1971 as pre-treatment periods and 1972–1975 as post-treatment periods, so that it is possible to capture both immediate and medium-term policy impacts. [Figure 4](#) displays the ASFR curves for East Germany and the control countries from 1956 to 1975. A clear reduction is evident for East Germany during the post-treatment period.

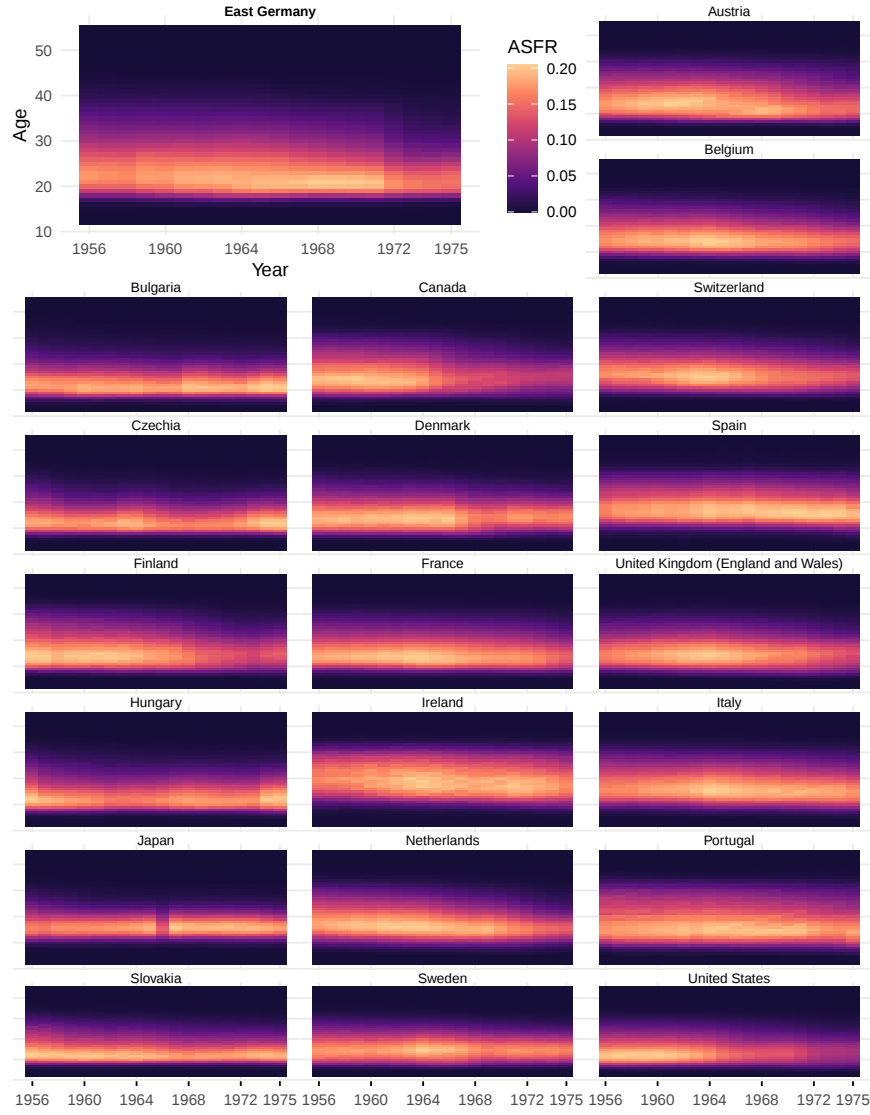


Figure 4: Age-specific fertility rates (ASFR) for East Germany and the control countries, 1956–1975. Each column represents the ASFR across ages 12–55 in the year shown on the horizontal axis.

Figure 5 compares the observed ASFR of East Germany with the synthetic control constructed from the donor countries. The synthetic control closely tracks East Germany’s pre-treatment fertility patterns, while after 1972 the observed trajectory exhibits a clear downward shift relative to the synthetic counterfactual, especially across prime childbearing ages. Using 1956–1961 for training and 1962–1971 for calibration, the estimated effect magnitudes for 1972–1975 are .0813, .1279, .1544, and .1332, with 90% split-conformal intervals [.0207, .1419], [.0673, .1885], [.0938, .2150], and [.0726, .1938]. Each conformal p -

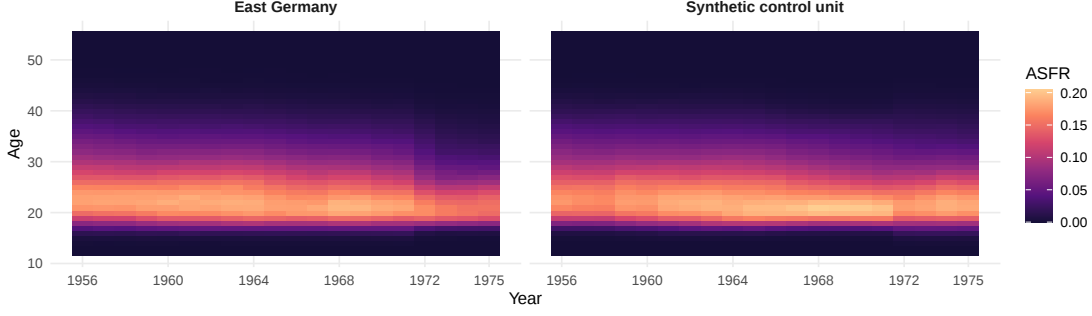


Figure 5: Heat maps of age-specific fertility rates (ASFR) for East Germany (left) and the synthetic control unit (right) from 1956 to 1975. Each column represents the ASFR across ages 12 to 55 (vertical axis) for each calendar year (horizontal axis).

value is $1/11$; 10 calibration periods do not yield a nontrivial 95% split-conformal interval.

As a diagnostic check, we also conduct an in-time placebo analysis by backdating the treatment year to $\tilde{T}_0 = 1967$, as discussed in Remark 4.1. Using only pre-1968 observations to estimate the GSC weights, the pseudo-post-treatment discrepancies for 1968–1971 are 0.032, 0.036, 0.039, and 0.042, respectively. These are substantially smaller than the true post-treatment discrepancies, 0.065, 0.116, 0.141, and 0.122 for 1972–1975. This diagnostic supports the interpretation that the post-1972 divergence reflects the policy effect rather than a comparable pre-existing discrepancy.

7 Discussion

The proposed geodesic synthetic control method extends causal inference to random-object outcomes in geodesic metric spaces by replacing Euclidean weighted averages with weighted Fréchet means and representing treatment effects as geodesics connecting untreated and treated potential outcomes. This provides an intrinsic extension of the classical synthetic control framework to a broad range of object-valued outcomes, including distributions, compositions, networks, SPD matrices, and functional data. Under the perturbed causal model, we established consistency of the estimated weights for the perturbed population target, derived a post-treatment prediction-error bound, and gave sufficient conditions for

consistent recovery of the untreated counterfactual. We also proposed a metric conformal calibration procedure for uncertainty quantification.

Several extensions are natural. First, when the untreated potential outcome of the treated unit is not well approximated by a weighted Fréchet mean of the controls, GSC may suffer from imperfect pre-treatment fit. A possible remedy is an augmented GSC estimator, paralleling the augmented synthetic control method of [Ben-Michael et al. \(2021\)](#). Developing such a method requires regression-based bias correction for metric space-valued outcomes, potentially building on Fréchet regression ([Petersen and Müller, 2019](#)) and geodesic optimal transport regression ([Zhu and Müller, 2026](#)). Second, it would be useful to develop a geodesic analogue of synthetic difference-in-differences ([Arkhangelsky et al., 2021](#)), combining unit reweighting with time reweighting. Recent work on geodesic difference-in-differences ([Zhou et al., 2025](#)) provides a foundation for this direction.

Metric choice is another important modeling issue for GSC. Many object spaces admit multiple meaningful metrics, and the chosen metric determines the distances, weighted Fréchet means, geodesics, synthetic counterfactuals, and treatment-effect magnitudes. Consequently, the metric should be selected based on the scientific question, the geometry of the data type, interpretability, and computational feasibility. For example, SPD matrices may be equipped with the Frobenius, Log-Euclidean, power, or Log-Cholesky metric. When several metrics are plausible, data-adaptive criteria or sensitivity analyses may be useful; see [Petersen and Müller \(2016\)](#) for adaptive metric selection for SPD matrices and [Dubey et al. \(2024\)](#) for a broader discussion. A systematic study of metric selection for GSC, especially in causal inference settings, is an important direction for future work.

Data Availability Statement

The employment data are openly available from the [RIETI R-JIP database](#), and the fertility data are available from the [Human Fertility Database](#).

Acknowledgments

The authors thank the editor, associate editor, and three anonymous referees for their valuable and constructive feedback.

Disclosure Statement

The authors report there are no competing interests to declare.

References

- Abadie, A. (2021) Using synthetic controls: Feasibility, data requirements, and methodological aspects. *Journal of Economic Literature*, **59**, 391–425.
- Abadie, A., Diamond, A. and 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**, 493–505.
- (2015) Comparative politics and the synthetic control method. *American Journal of Political Science*, **59**, 495–510.
- Abadie, A. and Gardeazabal, J. (2003) The economic costs of conflict: A case study of the Basque Country. *American Economic Review*, **93**, 113–132.
- Abadie, A. and L’Hour, J. (2021) A penalized synthetic control estimator for disaggregated data. *Journal of the American Statistical Association*, **116**, 1817–1834.
- Ambrosio, L., Gigli, N. and Savaré, G. (2008) *Gradient Flows in Metric Spaces and in the Space of Probability Measures*. Lectures in Mathematics. ETH Zürich. Birkhäuser Basel, 2nd edn.
- Amjad, M., Shah, D. and Shen, D. (2018) Robust synthetic control. *Journal of Machine Learning Research*, **19**, 1–51.
- Arkhangelsky, D., Athey, S., Hirshberg, D. A., Imbens, G. W. and Wager, S. (2021) Synthetic difference-in-differences. *American Economic Review*, **111**, 4088–4118.
- Arnaudon, M., Barbaresco, F. and Yang, L. (2013) Riemannian medians and means with applications to radar signal processing. *IEEE Journal of Selected Topics in Signal Processing*, **7**, 595–604.
- Arsigny, V., Fillard, P., Pennec, X. and Ayache, N. (2007) Geometric means in a novel vector space structure on symmetric positive-definite matrices. *SIAM Journal on Matrix Analysis and Applications*, **29**, 328–347.

- Ben-Michael, E., Feller, A. and Rothstein, J. (2021) The augmented synthetic control method. *Journal of the American Statistical Association*, **116**, 1789–1803.
- (2022) Synthetic controls with staggered adoption. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, **84**, 351–381.
- Billera, L. J., Holmes, S. P. and Vogtmann, K. (2001) Geometry of the space of phylogenetic trees. *Advances in Applied Mathematics*, **27**, 733–767.
- Bridson, M. R. and Haefliger, A. (1999) *Metric Spaces of Non-Positive Curvature*. Grundlehren der mathematischen Wissenschaften. Springer Berlin, Heidelberg.
- Buss, S. R. and Fillmore, J. P. (2001) Spherical averages and applications to spherical splines and interpolation. *ACM Transactions on Graphics (TOG)*, **20**, 95–126.
- Büttner, T. and Lutz, W. (1990) Estimating fertility responses to policy measures in the German Democratic Republic. *Population and Development Review*, **16**, 539–555.
- Chen, K., Delicado, P. and Müller, H.-G. (2017) Modelling function-valued stochastic processes, with applications to fertility dynamics. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, **79**, 177–196.
- Cherian, A. and Sra, S. (2016) Positive definite matrices: Data representation and applications to computer vision. In *Algorithmic Advances in Riemannian Geometry and Applications: For Machine Learning, Computer Vision, Statistics, and Optimization*, 93–114. Cham: Springer. URL: https://doi.org/10.1007/978-3-319-45026-1_4.
- Chernozhukov, V., Wüthrich, K. and Zhu, Y. (2021a) Distributional conformal prediction. *Proceedings of the National Academy of Sciences*, **118**, e2107794118.
- (2021b) An exact and robust conformal inference method for counterfactual and synthetic controls. *Journal of the American Statistical Association*, **116**, 1849–1864.
- Dai, X., Lin, Z. and Müller, H.-G. (2021) Modeling sparse longitudinal data on Riemannian manifolds. *Biometrics*, **77**, 1328–1341.
- Dryden, I. L., Koloydenko, A. and Zhou, D. (2009) Non-Euclidean statistics for covariance matrices, with applications to diffusion tensor imaging. *Annals of Applied Statistics*, **3**, 1102–1123.
- Dubey, P., Chen, Y. and Müller, H.-G. (2024) Metric statistics: Exploration and inference for random objects with distance profiles. *Annals of Statistics*, **52**, 757–792.
- Eltzner, B. and Huckemann, S. F. (2019) A smeary central limit theorem for manifolds with application to high-dimensional spheres. *Annals of Statistics*, **47**, 3360–3381.
- Ferman, B. and Pinto, C. (2021) Synthetic controls with imperfect pretreatment fit. *Quantitative Economics*, **12**, 1197–1221.
- Fréchet, M. (1948) Les éléments aléatoires de nature quelconque dans un espace distancié. *Annales de l’Institut Henri Poincaré*, **10**, 215–310.

- Ginestet, C. E., Simmons, A. and Kolaczyk, E. D. (2012) Weighted Fréchet means as convex combinations in metric spaces: properties and generalized median inequalities. *Statistics and Probability Letters*, **82**, 1859–1863.
- Gunsilius, F., Hsieh, M. H. and Lee, M. J. (2024) Tangential Wasserstein projections. *Journal of Machine Learning Research*, **25**, 1–41.
- Gunsilius, F. F. (2023) Distributional synthetic controls. *Econometrica*, **91**, 1105–1117.
- Hsing, T. and Eubank, R. (2015) *Theoretical Foundations of Functional Data Analysis, with an Introduction to Linear Operators*, vol. 997. John Wiley & Sons.
- Human Fertility Database (2025) Max Planck Institute for Demographic Research (Germany) and Vienna Institute of Demography (Austria). Available at www.humanfertility.org (data downloaded on March 16, 2025).
- Johnson, S. G. (2008) *The NLOpt nonlinear-optimization package*. URL: <https://github.com/stevengj/nlopt>.
- Kolaczyk, E. D. and Csárdi, G. (2014) *Statistical Analysis of Network Data with R*, vol. 65. Springer New York.
- Kurisu, D. and Otsu, T. (2026) Empirical likelihood for manifolds. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, **88**, 91–119.
- Kurisu, D., Zhou, Y., Otsu, T. and Müller, H.-G. (2024) Geodesic causal inference. *arXiv preprint arXiv:2406.19604*.
- Le, H. and Barden, D. (2014) On the measure of the cut locus of a Fréchet mean. *Bulletin of the London Mathematical Society*, **46**, 698–708.
- Lin, Z. (2019) Riemannian geometry of symmetric positive definite matrices via Cholesky decomposition. *SIAM Journal on Matrix Analysis and Applications*, **40**, 1353–1370.
- Lin, Z., Kong, D. and Wang, L. (2023) Causal inference on distribution functions. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, **85**, 378–398.
- McCann, R. J. (1997) A convexity principle for interacting gases. *Advances in Mathematics*, **128**, 153–179.
- Panaretos, V. M. and Zemel, Y. (2020) *An Invitation to Statistics in Wasserstein Space*. Springer New York.
- Patrangenaru, V. and Ellingson, L. (2015) *Nonparametric Statistics on Manifolds and Their Applications to Object Data Analysis*. CRC Press.
- Petersen, A. and Müller, H.-G. (2016) Fréchet integration and adaptive metric selection for interpretable covariances of multivariate functional data. *Biometrika*, **103**, 103–120.
- Petersen, A. and Müller, H.-G. (2016) Functional data analysis for density functions by transformation to a Hilbert space. *Annals of Statistics*, **44**, 183–218.

- Petersen, A. and Müller, H.-G. (2019) Fréchet regression for random objects with Euclidean predictors. *Annals of Statistics*, **47**, 691–719.
- Petersen, A., Zhang, C. and Kokoszka, P. (2022) Modeling probability density functions as data objects. *Econometrics and Statistics*, **21**, 159–178.
- Powell, M. J. D. (1994) A direct search optimization method that models the objective and constraint functions by linear interpolation. In *Advances in Optimization and Numerical Analysis*, 51–67. Dordrecht: Springer Netherlands. URL: https://doi.org/10.1007/978-94-015-8330-5_4.
- Ramsay, J. O. and Silverman, B. W. (2005) *Functional Data Analysis*. Springer Series in Statistics. Springer New York, 2nd edn.
- Robbins, M. W., Saunders, J. and Kilmer, B. (2017) A framework for synthetic control methods with high-dimensional, micro-level data: Evaluating a neighborhood-specific crime intervention. *Journal of the American Statistical Association*, **112**, 109–126.
- Scealy, J. and Welsh, A. (2011) Regression for compositional data by using distributions defined on the hypersphere. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, **73**, 351–375.
- Severn, K. E., Dryden, I. L. and Preston, S. P. (2022) Manifold valued data analysis of samples of networks, with applications in corpus linguistics. *Annals of Applied Statistics*, **16**, 368–390.
- Stellato, B., Banjac, G., Goulart, P., Bemporad, A. and Boyd, S. (2020) OSQP: An operator splitting solver for quadratic programs. *Mathematical Programming Computation*, **12**, 637–672.
- Sturm, K.-T. (2003) Probability measures on metric spaces of nonpositive curvature. *Heat Kernels and Analysis on Manifolds, Graphs, and Metric Spaces (Paris, 2002)*. *Contemp. Math.*, 338. Amer. Math. Soc., Providence, RI, **338**, 357–390.
- Torous, W., Gunsilius, F. and Rigollet, P. (2024) An optimal transport approach to estimating causal effects via nonlinear difference-in-differences. *Journal of Causal Inference*, **12**, 20230004.
- van der Vaart, A. and Wellner, J. (2023) *Weak Convergence and Empirical Processes: With Applications to Statistics*. Springer New York, 2nd edn.
- Van Dijke, D. (2025) Regression discontinuity design with distribution-valued outcomes. *arXiv preprint arXiv:2504.03992*.
- Wang, J.-L., Chiou, J.-M. and Müller, H.-G. (2016) Functional Data Analysis. *Annual Review of Statistics and its Application*, **3**, 257–295.
- Zhou, Y., Kurisu, D., Otsu, T. and Müller, H.-G. (2025) Geodesic difference-in-differences. *arXiv:2501.17436*.

- Zhou, Y. and Müller, H.-G. (2022) Network regression with graph Laplacians. *Journal of Machine Learning Research*, **23**, 1–41.
- Zhu, C. and Müller, H.-G. (2026) Geodesic optimal transport regression. *Biometrika*, **113**, asaf086.

SUPPLEMENTARY MATERIAL

S.1 Optimization Details

This section provides additional computational details for the implementation of the geodesic synthetic control (GSC) estimator. We present explicit formulations of the optimization problems, describe how weighted Fréchet means are computed in different geodesic metric spaces, and discuss the solver strategies adopted in each case.

S.1.1 Unit weights

The GSC relies on optimizing weights attached to observational units rather than features of the objects themselves. Specifically, the *unit weights* $\hat{\mathbf{w}} \in \Delta^{J-1}$ form convex combinations of control units. Random objects are then averaged via weighted Fréchet means in the respective metric space. Thus, the weights govern the influence of units but do not vary across internal coordinates of the objects (such as quantiles of a distribution). This preserves interpretability in the geodesic extension.

S.1.2 General optimization problems

The estimation of unit weights in GSC is formulated as a minimization problem involving distances between the treated unit and weighted Fréchet means of control units. Specifically, the unit weights $\hat{\mathbf{w}} \in \Delta^{J-1}$ are chosen to minimize the average squared distance between the treated outcome $Y_{1,t}$ and the synthetic control outcome $Y_{2:J+1,t}^{(\mathbf{w})}$ over the pre-treatment period,

$$\hat{\mathbf{w}} \in \arg \min_{\mathbf{w} \in \Delta^{J-1}} \frac{1}{T_0} \sum_{t=1}^{T_0} d^2 \left(Y_{1,t}, Y_{2:J+1,t}^{(\mathbf{w})} \right),$$

where the synthetic control at time t is defined through the weighted Fréchet mean

$$Y_{2:J+1,t}^{(\mathbf{w})} = \arg \min_{\nu \in \mathcal{M}} \sum_{j=2}^{J+1} w_j d^2(\nu, Y_{j,t}).$$

The optimization is carried out over the probability simplex, with the geometry of the underlying metric space determining how the Fréchet means are evaluated.

S.1.3 Spaces with closed-form weighted Fréchet means

For networks (Example 1), symmetric positive-definite (SPD) matrices (Example 3), distributions (Example 4), and functional data (Example 5), the weighted Fréchet mean admits an explicit expression under the metrics considered in this paper. In these cases, the outer optimization problems in Section S.1.2 reduce to convex quadratic programs in $\mathbf{w}$, which can be solved efficiently using standard quadratic programming solvers such as OSQP (Stellato et al., 2020). This yields numerically stable and fast implementations.

S.1.4 Compositional data with the arc-length distance

For compositional data with the arc-length distance (Example 2), the situation is more complex. When mapped to the positive orthant of the sphere, the weighted Fréchet mean has no closed form and must be obtained via iterative algorithms on the manifold. Each evaluation of the objective therefore requires solving an inner Fréchet mean problem, which creates a bilevel optimization problem. Because analytical differentiation is intractable in this nonlinear setting, we employ a derivative-free constrained optimization strategy, specifically the COBYLA algorithm (Powell, 1994) as implemented in the `nloptr` package (Johnson, 2008). The inner Fréchet mean computations are carried out using the `manifold` package (Dai et al., 2021).

S.1.5 Practical remarks

For the tractable cases with closed-form weighted Fréchet means, optimization is straightforward and stable. For compositional data with the arc-length distance, convergence of the iterative mean calculation can be slow when many observations lie close to zero com-

ponents, in which case regularization (e.g., adding a small constant to proportions before normalization) improves stability.

S.2 Verification of Assumption 3.1

This section gives sufficient conditions for Assumption 3.1 under the signal models and metrics treated below: the Frobenius metric for networks, the arc-length metric for compositional data, the Frobenius and Log-Euclidean metrics, the power metric family, and the Log-Cholesky metric for SPD matrices, the 2-Wasserstein metric for univariate distributions, and the L^2 metric for functional data. For $\mathbf{w} = (w_2, \dots, w_{J+1})^\top \in \Delta^{J-1}$, define

$$U_{2:J+1}^{(\mathbf{w})} = \arg \min_{\nu \in \mathcal{M}} \sum_{j=2}^{J+1} w_j d^2(\nu, U_j), \quad g_t^{(\mathbf{w})}(U_{2:J+1}) = g_{2:J+1,t}^{(\mathbf{w})} = \arg \min_{\nu \in \mathcal{M}} \sum_{j=2}^{J+1} w_j d^2(\nu, g_t(U_j)),$$

and $\mathbb{W}^* = \{\mathbf{w} \in \Delta^{J-1} : d(U_1, U_{2:J+1}^{(\mathbf{w})}) = 0\}$. Throughout Sections S.2.1–S.2.5 below, we assume that $\mathbb{W}^*$ is not empty.

S.2.1 Networks

Recall that $g_t(U_j) = \gamma_{\mu_t, U_j}(\alpha_t) = \mu_t + \alpha_t(U_j - \mu_t)$, $\alpha_t \in (0, 1)$. For $\mathbf{w} \in \Delta^{J-1}$,

$$\begin{aligned} U_{2:J+1}^{(\mathbf{w})} &= \sum_{j=2}^{J+1} w_j U_j, \\ g_t^{(\mathbf{w})}(U_{2:J+1}) &= \sum_{j=2}^{J+1} w_j \{\mu_t + \alpha_t(U_j - \mu_t)\} = \mu_t + \alpha_t \left\{ \sum_{j=2}^{J+1} w_j U_j - \mu_t \right\} \\ &= \mu_t + \alpha_t(U_{2:J+1}^{(\mathbf{w})} - \mu_t) = g_t(U_{2:J+1}^{(\mathbf{w})}). \end{aligned} \tag{S1}$$

Furthermore,

$$\begin{aligned} d_F^2(g_t(U_1), g_t(U_{2:J+1}^{(\mathbf{w})})) &= \text{tr} \left\{ (g_t(U_1) - g_t(U_{2:J+1}^{(\mathbf{w})}))^\top (g_t(U_1) - g_t(U_{2:J+1}^{(\mathbf{w})})) \right\} \\ &= \alpha_t^2 \text{tr} \left\{ (U_1 - U_{2:J+1}^{(\mathbf{w})})^\top (U_1 - U_{2:J+1}^{(\mathbf{w})}) \right\} = \alpha_t^2 d_F^2(U_1, U_{2:J+1}^{(\mathbf{w})}), \end{aligned} \tag{S2}$$

where $\text{tr}(A)$ is the trace of the matrix A . This yields $\bigcap_{t=1}^{T_0} \mathbb{W}_t^* = \bigcap_{t=T_0+1}^T \mathbb{W}_t^* = \mathbb{W}^*$ and hence Assumption 3.1 holds.

S.2.2 Compositional data

For compositional data, we verify Assumption 3.1 under the following spherical signal model. Let

$$g_t(U_j) = \text{Exp}_{\mu_t}(A_t \text{Log}_{U_1}(U_j)), \quad \mu_t \in \mathcal{S}_+^{d-1},$$

where Exp_μ and Log_μ denote the Riemannian exponential and logarithmic maps on the unit sphere. Let $T_\mu \mathcal{S}^{d-1} = \{v \in \mathbb{R}^d : \mu^\top v = 0\}$ denote the tangent space of the unit sphere $\mathcal{S}^{d-1}$ at $\mu \in \mathcal{S}_+^{d-1}$. Choose a rotation matrix $R_t \in \text{SO}(d)$ such that $R_t U_1 = \mu_t$, and define $A_t = \kappa_t R_t$, where $\kappa_t > 0$ is a constant. Assume that κ_t is chosen so that $\|A_t \text{Log}_{U_1}(U_j)\| < \pi$ and $\text{Exp}_{\mu_t}(A_t \text{Log}_{U_1}(U_j)) \in \mathcal{S}_+^{d-1}$ for every j . These conditions place the tangent vectors in the injectivity domain of Exp_{μ_t} and keep the signal in the outcome space. Then A_t maps $T_{U_1} \mathcal{S}^{d-1}$ into $T_{\mu_t} \mathcal{S}^{d-1}$. Indeed, for any $v \in T_{U_1} \mathcal{S}^{d-1}$, we have $U_1^\top v = 0$, and therefore $\mu_t^\top (A_t v) = \kappa_t U_1^\top R_t^\top R_t v = \kappa_t U_1^\top v = 0$. Thus $A_t v \in T_{\mu_t} \mathcal{S}^{d-1}$, and the expression defining $g_t(U_j)$ is well defined. Note that $g_t(U_1) = \text{Exp}_{\mu_t}(0) = \mu_t$. For every $\mathbf{w} \in \Delta^{J-1}$ and time point t , assume that, almost surely, the discrete distributions assigning masses w_j to the raw latent objects U_j , the latent-signal objects $g_t(U_j)$, and the perturbed objects $g_t(\mathcal{P}_{j,t}(U_j))$ satisfy the conditions of Theorem 1 in [Buss and Fillmore \(2001\)](#). Then all corresponding weighted Fréchet means are unique, including those required by Assumption 3.1(i). Additional uniqueness conditions for Fréchet means on Riemannian manifolds are given in Remark 2.8 of [Eltzner and Huckemann \(2019\)](#). The same necessity-and-sufficiency argument used below, applied at U_1 , gives

$$\mathbb{W}^* = \left\{ \mathbf{w} \in \Delta^{J-1} : \sum_{j=2}^{J+1} w_j \text{Log}_{U_1}(U_j) = 0 \right\}.$$

For $\overline{\mathbb{W}}_t^* = \{\mathbf{w} \in \Delta^{J-1} : \sum_{j=2}^{J+1} w_j \text{Log}_{g_t(U_1)}(g_t(U_j)) = 0\}$, one can verify Assumption 3.1(ii) if

$$\mathbb{W}_t^* = \overline{\mathbb{W}}_t^*. \tag{S3}$$

Indeed, we have

$$\begin{aligned}\mathbb{W}_t^* &= \overline{\mathbb{W}}_t^* = \left\{ \mathbf{w} \in \Delta^{J-1} : \sum_{j=2}^{J+1} w_j \text{Log}_{\mu_t}(\text{Exp}_{\mu_t}(A_t \text{Log}_{U_1}(U_j))) = 0 \right\} \\ &= \left\{ \mathbf{w} \in \Delta^{J-1} : A_t \sum_{j=2}^{J+1} w_j \text{Log}_{U_1}(U_j) = 0 \right\} = \left\{ \mathbf{w} \in \Delta^{J-1} : \sum_{j=2}^{J+1} w_j \text{Log}_{U_1}(U_j) = 0 \right\} = \mathbb{W}^*.\end{aligned}$$

This yields $\bigcap_{t=1}^{T_0} \mathbb{W}_t^* = \bigcap_{t=T_0+1}^T \mathbb{W}_t^* = \mathbb{W}^*$ and hence Assumption 3.1(ii) holds.

Next we show (S3). Specifically, we establish (i) $\mathbb{W}_t^* \subset \overline{\mathbb{W}}_t^*$ and (ii) $\mathbb{W}_t^* \supset \overline{\mathbb{W}}_t^*$.

Proof of (i): Fix an arbitrary $\mathbf{w} = (w_2, \dots, w_{J+1})^\top \in \mathbb{W}_t^*$. By definition, $g_t^{(\mathbf{w})}(U_{2:J+1})$ is a Fréchet mean of a random element X such that $\Pr\{X = g_t(U_j)\} = w_j$, $j = 2, \dots, J+1$. Further, the condition $d(g_t(U_1), g_t^{(\mathbf{w})}(U_{2:J+1})) = 0$ yields $g_t(U_1) = g_t^{(\mathbf{w})}(U_{2:J+1})$. Applying Corollary 3 in [Le and Barden \(2014\)](#), we then have $\sum_{j=2}^{J+1} w_j \text{Log}_{g_t(U_1)}(g_t(U_j)) = 0$. This implies $\mathbf{w} \in \overline{\mathbb{W}}_t^*$ and hence $\mathbb{W}_t^* \subset \overline{\mathbb{W}}_t^*$ holds.

Proof of (ii): Fix an arbitrary $\mathbf{w} = (w_2, \dots, w_{J+1})^\top \in \overline{\mathbb{W}}_t^*$. From the definition of $\overline{\mathbb{W}}_t^*$ and the uniqueness of the Fréchet mean, Theorem 1 in [Buss and Fillmore \(2001\)](#) yields $g_t(U_1) = g_t^{(\mathbf{w})}(U_{2:J+1})$. This implies $\mathbf{w} \in \mathbb{W}_t^*$ and hence $\mathbb{W}_t^* \supset \overline{\mathbb{W}}_t^*$.

S.2.3 SPD matrices

First, we consider the space of SPD matrices with the Frobenius metric. Standard arguments lead to (S1) and (S2). Hence Assumption 3.1 holds.

Second, we consider the space of SPD matrices with the power metric family defined as $d_{F,p}(A, B) = \|F_p(A) - F_p(B)\|_F$, where $p > 0$ is a constant, $\|\cdot\|_F$ is the Frobenius metric, and $F_p : \text{Sym}_m^+ \rightarrow \text{Sym}_m^+$ is the matrix power map defined by $F_p(A) = A^p = U\Lambda^p U^\top$, where $U\Lambda U^\top$ is the usual spectral decomposition of $A \in \text{Sym}_m^+$. The geodesic with respect to $d_{F,p}$ between A and B is given as $\gamma_{A,B}(t) = [A^p + t(B^p - A^p)]^{1/p} = F_{1/p}(A^p + t(B^p - A^p))$.

Recall that $g_t(U_j) = \gamma_{\mu_t, U_j}(\alpha_t)$. For $\mathbf{w} \in \Delta^{J-1}$,

$$\begin{aligned} (U_{2:J+1}^{(\mathbf{w})})^p &= \sum_{j=2}^{J+1} w_j U_j^p, \\ (g_t^{(\mathbf{w})}(U_{2:J+1}))^p &= \mu_t^p + \alpha_t \left\{ \sum_{j=2}^{J+1} w_j U_j^p - \mu_t^p \right\} = (g_t(U_{2:J+1}^{(\mathbf{w})}))^p. \end{aligned}$$

Furthermore,

$$\begin{aligned} d_{F,p}^2(g_t(U_1), g_t(U_{2:J+1}^{(\mathbf{w})})) &= \|(g_t(U_1))^p - (g_t(U_{2:J+1}^{(\mathbf{w})}))^p\|_F^2 \\ &= \alpha_t^2 \left\| U_1^p - \sum_{j=2}^{J+1} w_j U_j^p \right\|_F^2 = \alpha_t^2 d_{F,p}^2(U_1, U_{2:J+1}^{(\mathbf{w})}). \end{aligned}$$

This yields $\bigcap_{t=1}^{T_0} \mathbb{W}_t^* = \bigcap_{t=T_0+1}^T \mathbb{W}_t^* = \mathbb{W}^*$ and hence Assumption 3.1 holds.

Third, we consider the space of SPD matrices with the Log-Euclidean metric, $d_{\text{LE}}(A, B) = \|\log(A) - \log(B)\|_F$. The geodesic with respect to d_{LE} between A and B is $\gamma_{A,B}(t) = \exp(\log A + t(\log B - \log A))$. Recall $g_t(U_j) = \gamma_{\mu_t, U_j}(\alpha_t)$. For $\mathbf{w} \in \Delta^{J-1}$,

$$\begin{aligned} \log(U_{2:J+1}^{(\mathbf{w})}) &= \sum_{j=2}^{J+1} w_j \log(U_j), \\ \log(g_t^{(\mathbf{w})}(U_{2:J+1})) &= \log(\mu_t) + \alpha_t \left(\sum_{j=2}^{J+1} w_j \log(U_j) - \log(\mu_t) \right) = \log(g_t(U_{2:J+1}^{(\mathbf{w})})). \end{aligned}$$

Further,

$$\begin{aligned} d_{\text{LE}}^2(g_t(U_1), g_t(U_{2:J+1}^{(\mathbf{w})})) &= \|\log(g_t(U_1)) - \log(g_t(U_{2:J+1}^{(\mathbf{w})}))\|_F^2 \\ &= \alpha_t^2 \left\| \log(U_1) - \sum_{j=2}^{J+1} w_j \log(U_j) \right\|_F^2 = \alpha_t^2 d_{\text{LE}}^2(U_1, U_{2:J+1}^{(\mathbf{w})}). \end{aligned}$$

This yields $\bigcap_{t=1}^{T_0} \mathbb{W}_t^* = \bigcap_{t=T_0+1}^T \mathbb{W}_t^* = \mathbb{W}^*$ and hence Assumption 3.1 holds.

Fourth, we consider the space of SPD matrices with the Log-Cholesky metric. We identify each SPD matrix with its unique Cholesky factor $L \in \mathcal{L}_m^+$, where $\mathcal{L}_m^+$ is the space of lower-triangular matrices with positive diagonal entries. For $L \in \mathcal{L}_m^+$, let $\lfloor L \rfloor$ and $\mathbb{D}(L)$ denote its strictly lower-triangular and diagonal parts, respectively. The Log-Cholesky metric is defined as

$$d_{\text{LC}}(L_1, L_2) = \left\{ \|\lfloor L_1 \rfloor - \lfloor L_2 \rfloor\|_F^2 + \|\log(\mathbb{D}(L_1)) - \log(\mathbb{D}(L_2))\|_F^2 \right\}^{1/2}.$$

Note that the space $(\mathcal{L}_m^+, d_{\text{LC}})$ is a geodesic metric space (Lin, 2019) and the geodesic with respect to d_{LC} between L_1 and L_2 is

$$\gamma_{L_1, L_2}(t) = \lfloor L_1 \rfloor + t(\lfloor L_2 \rfloor - \lfloor L_1 \rfloor) + \exp \{ \log(\mathbb{D}(L_1)) + t(\log(\mathbb{D}(L_2)) - \log(\mathbb{D}(L_1))) \}.$$

Let $U_j \in \mathcal{L}_m^+$, $j = 1, \dots, J+1$ and $\mu_t \in \mathcal{L}_m^+$, $t = 1, \dots, T$. Recall that $g_t(U_j) = \gamma_{\mu_t, U_j}(\alpha_t)$.

For $\mathbf{w} \in \Delta^{J-1}$,

$$\begin{aligned} \lfloor (U_{2:J+1}^{(\mathbf{w})}) \rfloor &= \sum_{j=2}^{J+1} w_j \lfloor U_j \rfloor, \quad \log(\mathbb{D}(U_{2:J+1}^{(\mathbf{w})})) = \sum_{j=2}^{J+1} w_j \log(\mathbb{D}(U_j)), \\ \lfloor g_t^{(\mathbf{w})}(U_{2:J+1}) \rfloor &= \lfloor \mu_t \rfloor + \alpha_t \left\{ \sum_{j=2}^{J+1} w_j \lfloor U_j \rfloor - \lfloor \mu_t \rfloor \right\} = \lfloor g_t(U_{2:J+1}^{(\mathbf{w})}) \rfloor, \\ \log(\mathbb{D}(g_t^{(\mathbf{w})}(U_{2:J+1}))) &= \log(\mathbb{D}(\mu_t)) + \alpha_t \left\{ \sum_{j=2}^{J+1} w_j \log(\mathbb{D}(U_j)) - \log(\mathbb{D}(\mu_t)) \right\} \\ &= \log(\mathbb{D}(g_t(U_{2:J+1}^{(\mathbf{w})}))). \end{aligned}$$

Then

$$\begin{aligned} d_{\text{LC}}^2(g_t(U_1), g_t(U_{2:J+1}^{(\mathbf{w})})) &= \|\lfloor g_t(U_1) \rfloor - \lfloor g_t(U_{2:J+1}^{(\mathbf{w})}) \rfloor\|_F^2 + \|\log(\mathbb{D}(g_t(U_1))) - \log(\mathbb{D}(g_t(U_{2:J+1}^{(\mathbf{w})})))\|_F^2 \\ &= \alpha_t^2 \|\lfloor U_1 \rfloor - \sum_{j=2}^{J+1} w_j \lfloor U_j \rfloor\|_F^2 + \alpha_t^2 \|\log(\mathbb{D}(U_1)) - \sum_{j=2}^{J+1} w_j \log(\mathbb{D}(U_j))\|_F^2 \\ &= \alpha_t^2 d_{\text{LC}}^2(U_1, U_{2:J+1}^{(\mathbf{w})}). \end{aligned}$$

This yields $\bigcap_{t=1}^{T_0} \mathbb{W}_t^* = \bigcap_{t=T_0+1}^T \mathbb{W}_t^* = \mathbb{W}^*$ and hence Assumption 3.1 holds.

S.2.4 Univariate probability distributions

Recall that $g_t(U_j) := \gamma_{\mu_t, U_j}(\alpha_t)$, $\alpha_t \in (0, 1)$. Then we have $F_{g_t(U_j)}^{-1}(\cdot) = F_{\mu_t}^{-1}(\cdot) + \alpha_t(F_{U_j}^{-1}(\cdot) - F_{\mu_t}^{-1}(\cdot))$. For $\mathbf{w} \in \Delta^{J-1}$,

$$\begin{aligned} F_{U_{2:J+1}}^{-1}(\cdot) &= \sum_{j=2}^{J+1} w_j F_{U_j}^{-1}(\cdot), \\ F_{g_t(U_{2:J+1})}^{-1}(\cdot) &= \sum_{j=2}^{J+1} w_j F_{g_t(U_j)}^{-1}(\cdot) = \sum_{j=2}^{J+1} w_j \left\{ F_{\mu_t}^{-1}(\cdot) + \alpha_t(F_{U_j}^{-1}(\cdot) - F_{\mu_t}^{-1}(\cdot)) \right\} \\ &= F_{\mu_t}^{-1}(\cdot) + \alpha_t \left\{ \sum_{j=2}^{J+1} w_j F_{U_j}^{-1}(\cdot) - F_{\mu_t}^{-1}(\cdot) \right\} = \left(F_{\gamma_{\mu_t, U_{2:J+1}}(\mathbf{w})}(\alpha_t) \right)^{-1}(\cdot) \\ &= F_{g_t(U_{2:J+1})}^{-1}(\cdot). \end{aligned}$$

Furthermore,

$$d_{\mathcal{W}}^2(g_t(U_1), g_t(U_{2:J+1})) = \alpha_t^2 \int_0^1 \left\{ F_{U_1}^{-1}(p) - \sum_{j=2}^{J+1} w_j F_{U_j}^{-1}(p) \right\}^2 dp = \alpha_t^2 d_{\mathcal{W}}^2(U_1, U_{2:J+1}),$$

which yields $\bigcap_{t=1}^{T_0} \mathbb{W}_t^* = \bigcap_{t=T_0+1}^T \mathbb{W}_t^* = \mathbb{W}^*$ and hence Assumption 3.1 holds.

S.2.5 Functional data

Note that $(g_t(U_j))(s) = (\gamma_{\mu_t, U_j}(\alpha_t))(s) = \mu_t(s) + \alpha_t(U_j(s) - \mu_t(s))$, $s \in \mathcal{T}$. For $\mathbf{w} \in \Delta^{J-1}$,

$$\begin{aligned} U_{2:J+1}^{(\mathbf{w})}(s) &= \sum_{j=2}^{J+1} w_j U_j(s), \\ (g_t^{(\mathbf{w})}(U_{2:J+1}))(s) &= \sum_{j=2}^{J+1} w_j (g_t(U_j))(s) \\ &= \mu_t(s) + \alpha_t(U_{2:J+1}^{(\mathbf{w})}(s) - \mu_t(s)) = (g_t(U_{2:J+1}^{(\mathbf{w})}))(s), \quad s \in \mathcal{T}. \end{aligned}$$

Next,

$$\begin{aligned} d_{L^2}^2(g_t(U_1), g_t(U_{2:J+1}^{(\mathbf{w})})) &= \int_{\mathcal{T}} \left\{ (g_t(U_1))(s) - (g_t(U_{2:J+1}^{(\mathbf{w})}))(s) \right\}^2 ds \\ &= \alpha_t^2 \int_{\mathcal{T}} \left\{ U_1(s) - U_{2:J+1}^{(\mathbf{w})}(s) \right\}^2 ds = \alpha_t^2 d_{L^2}^2(U_1, U_{2:J+1}^{(\mathbf{w})}), \end{aligned}$$

which gives $\bigcap_{t=1}^{T_0} \mathbb{W}_t^* = \bigcap_{t=T_0+1}^T \mathbb{W}_t^* = \mathbb{W}^*$ and hence Assumption 3.1 holds.

S.3 Verification of Assumption 3.2 and perturbation maps

This section provides sufficient conditions for Assumption 3.2 and the standing bounded-range condition in the outcome spaces considered in the paper. We first give sufficient conditions for the criterion convergence, minimizer, and stability requirements in Assumption 3.2. We then construct outcome-specific perturbation maps and record conditions under which the general results apply; the range implications use the signal models in Section S.2. We write $a \geq 0$ for the perturbation scale. Theorem 3.1 permits a fixed perturbation scale, while Corollary S1 uses a triangular array with $a = a_{T_0} \rightarrow 0$.

S.3.1 General sufficient conditions for Assumption 3.2

Assumption 3.2 is verified through two alternative results. Proposition S1 is the Hilbert-coordinate route: cross-moment convergence establishes criterion convergence, and coordinate linearity gives weight stability with exponent $\kappa = 1$. Proposition S2 is the spherical-composition route: a loss-process law of large numbers establishes criterion convergence, and quadratic growth gives weight stability with exponent $\kappa = 1/2$. Neither proposition depends on the other. Section S.3.2 provides temporal conditions for the convergence requirements, and Sections S.3.3–S.3.7 assign each outcome space to the appropriate route.

For the Hilbert-coordinate route, suppose that $\Phi : \mathcal{M} \rightarrow \mathcal{H}$ is an isometry onto a convex subset of a separable Hilbert space $\mathcal{H}$. For $t \leq T_0$, write

$$X_{j,t} = \Phi(Y_{j,t}(N)), \quad j = 1, \dots, J+1.$$

Convexity of $\Phi(\mathcal{M})$ implies that the weighted Fréchet mean $m_t(\mathbf{w}) = \hat{g}_{2:J+1,t}^{(\mathbf{w})}$ is unique and satisfies

$$\Phi(m_t(\mathbf{w})) = \sum_{j=2}^{J+1} w_j X_{j,t}. \tag{S4}$$

Fix $\nu_0 \in \mathcal{M}$, write $z_0 = \Phi(\nu_0)$, and define the empirical cross-moments

$$G_{jk,T_0} = \frac{1}{T_0} \sum_{t=1}^{T_0} \langle X_{j,t} - z_0, X_{k,t} - z_0 \rangle_{\mathcal{H}}, \quad 1 \leq j, k \leq J+1.$$

Proposition S1 (Hilbert-coordinate verification). *Suppose that $\|X_{j,t} - z_0\|_{\mathcal{H}} \leq B$ almost surely, uniformly over the required unit, time, and triangular-array indices, for some $B < \infty$. If*

$$G_{jk,T_0} \xrightarrow{p} G_{jk} \quad \text{for every } 1 \leq j, k \leq J+1,$$

where the limits are deterministic, then Assumption 3.2 holds. Moreover, the pre- and post-treatment stability conditions hold with exponent $\kappa = 1$ whenever the same coordinate bound holds over the corresponding time indices.

Proof. Since $\sum_{j=2}^{J+1} w_j = 1$, the coordinate representation gives

$$\widehat{Q}(\mathbf{w}) = G_{11,T_0} - 2 \sum_{j=2}^{J+1} w_j G_{1j,T_0} + \sum_{j,k=2}^{J+1} w_j w_k G_{jk,T_0}.$$

There are finitely many cross-moments because J is fixed. Their convergence therefore yields uniform convergence of $\widehat{Q}$ over Δ^{J-1} to

$$\widetilde{Q}(\mathbf{w}) = G_{11} - 2 \sum_{j=2}^{J+1} w_j G_{1j} + \sum_{j,k=2}^{J+1} w_j w_k G_{jk}.$$

The cross-moments are uniformly bounded, so convergence in probability also implies convergence of their expectations. Thus this polynomial agrees with the definition $\widetilde{Q}(\mathbf{w}) = \lim_{T_0 \rightarrow \infty} \mathbb{E}[\widehat{Q}(\mathbf{w})]$. It is finite and continuous. Hence its full minimizer set $\widetilde{\mathcal{W}}$ is nonempty and compact, and continuity on the compact simplex gives the separation condition in Assumption 3.2(i). The empirical criterion is jointly measurable and continuous in $\mathbf{w}$; compactness of the simplex and the measurable selection theorem give a measurable empirical minimizer.

For any $\mathbf{w}_1, \mathbf{w}_2 \in \Delta^{J-1}$, (S4) gives

$$\begin{aligned} d(m_t(\mathbf{w}_1), m_t(\mathbf{w}_2)) &= \left\| \sum_{j=2}^{J+1} (w_{1j} - w_{2j})(X_{j,t} - z_0) \right\|_{\mathcal{H}} \\ &\leq B \|\mathbf{w}_1 - \mathbf{w}_2\|_1. \end{aligned}$$

This bound verifies Assumption 3.2(ii) and, when imposed over a fixed post-treatment horizon, the stability condition in Theorem 3.1(i). $\square$

For the spherical-composition route, let K be a common compact geodesically convex subset of the interior of $\mathcal{S}_+^{d-1}$ containing the relevant untreated outcomes and all of their weighted Fréchet means. For

$$F_{t,\mathbf{w}}(\nu) = \sum_{j=2}^{J+1} w_j d^2(\nu, Y_{j,t}(N)),$$

suppose that the weighted Fréchet mean $m_t(\mathbf{w})$ is unique and jointly measurable and that, for some $c > 0$,

$$F_{t,\mathbf{w}}(\nu) - F_{t,\mathbf{w}}(m_t(\mathbf{w})) \geq c d^2(\nu, m_t(\mathbf{w})) \quad (\text{S5})$$

uniformly in t , $\mathbf{w}$, and $\nu \in K$. This quadratic-growth condition follows, for example, from uniform strong geodesic convexity of the squared-distance objectives on a sufficiently small spherical region.

Proposition S2 (Spherical-composition verification). *Suppose the preceding conditions hold and, for each $\mathbf{w} \in \Delta^{J-1}$,*

$$\frac{1}{T_0} \sum_{t=1}^{T_0} \mathbb{E}[\ell_t(\mathbf{w})] \longrightarrow \tilde{Q}(\mathbf{w}), \quad \ell_t(\mathbf{w}) = d^2(Y_{1,t}(N), m_t(\mathbf{w})), \quad (\text{S6})$$

$$\frac{1}{T_0} \sum_{t=1}^{T_0} \{\ell_t(\mathbf{w}) - \mathbb{E}[\ell_t(\mathbf{w})]\} \xrightarrow{p} 0. \quad (\text{S7})$$

Then Assumption 3.2 holds. The corresponding stability condition holds with exponent $\kappa = 1/2$.

Proof. Let $D_K = \text{diam}(K)$. Applying (S5) with $\nu = m_t(\mathbf{w}_2)$ and comparing the two weighted objectives yields

$$c d^2(m_t(\mathbf{w}_1), m_t(\mathbf{w}_2)) \leq 2D_K^2 \|\mathbf{w}_1 - \mathbf{w}_2\|_1.$$

Thus $m_t(\mathbf{w})$ is uniformly Hölder continuous in the weights with exponent $1/2$, which verifies Assumption 3.2(ii). Equations (S6)–(S7) give pointwise convergence of $\hat{Q}$. The same Hölder

bound and the boundedness of K imply that $\tilde{Q}$ is continuous. Compactness then gives a nonempty minimizer set and separation, while joint measurability and continuity give a measurable empirical minimizer, as in the proof of Proposition S1. $\square$

S.3.2 Temporal conditions for criterion convergence

The cross-moment condition in Proposition S1 and the loss law of large numbers in Proposition S2 can be verified under standard temporal conditions. If the joint coordinate-outcome process is strictly stationary and ergodic, the ergodic theorem gives convergence of the bounded cross-moments. Likewise, if $\{\ell_t(\mathbf{w})\}_{t \in \mathbb{Z}}$ is strictly stationary and ergodic for each fixed $\mathbf{w}$, then (S6)–(S7) hold with $\tilde{Q}(\mathbf{w}) = \mathbb{E}[\ell_0(\mathbf{w})]$.

Strict stationarity alone is not sufficient for either conclusion. A second sufficient regime, which also accommodates deterministic time variation in g_t , is that the joint innovation process is strongly mixing with coefficients $\alpha(k) \rightarrow 0$, the relevant cross-products or losses are measurable functions of the current innovation vector, and their Cesàro expectations converge. Because these variables are uniformly bounded, the standard mixing covariance inequality gives

$$\text{Var} \left[\frac{1}{T_0} \sum_{t=1}^{T_0} \{Z_t - \mathbb{E}(Z_t)\} \right] \leq \frac{C_1}{T_0} + \frac{C_2}{T_0} \sum_{k=1}^{T_0-1} \alpha(k) \rightarrow 0,$$

where Z_t denotes any required cross-product or loss. This proves the corresponding law of large numbers by Chebyshev's inequality. When g_t varies deterministically, stationarity or weak dependence is imposed on the induced coordinate or loss process in this manner.

For the perturbation constructions below, let $\mathcal{P}_t = (\mathcal{P}_{1,t}, \dots, \mathcal{P}_{J+1,t})$. The sequence $\{\mathcal{P}_t\}_{t \in \mathbb{Z}}$ is strictly stationary if, for every positive integer k and all integers $t_1, \dots, t_k, \tau$, the joint distributions of $(\mathcal{P}_{t_1}, \dots, \mathcal{P}_{t_k})$ and $(\mathcal{P}_{t_1+\tau}, \dots, \mathcal{P}_{t_k+\tau})$ are identical. The stationary-ergodic regime is imposed on the induced coordinate-outcome or loss process, not only on the innovations. In the mixing regime, the full innovation vector across units is jointly strongly mixing, and current-time measurable transformations inherit its mixing coeffi-

cients. For a triangular array with row-specific coefficients $\alpha_{T_0}(k)$, the sufficient uniform condition is

$$\frac{1}{T_0} \sum_{k=1}^{T_0-1} \alpha_{T_0}(k) \longrightarrow 0,$$

together with the uniform boundedness and Cesàro expectation convergence stated above.

S.3.3 Networks

Consider weighted undirected networks on m nodes, represented by graph Laplacians. Let $\mathcal{E}_m = \{(r, s) : 1 \leq r < s \leq m\}$ be the set of unordered edges and let $B_{rs} = (e_r - e_s)(e_r - e_s)^\top$, where e_r is the r th Euclidean basis vector in $\mathbb{R}^m$. Any weighted graph Laplacian with nonnegative edge weights $b = (b_{rs})_{(r,s) \in \mathcal{E}_m}$ can be written as

$$L(b) = \sum_{(r,s) \in \mathcal{E}_m} b_{rs} B_{rs}.$$

Let $\xi_t = (\xi_{j,t,rs} : j = 1, \dots, J+1, (r, s) \in \mathcal{E}_m)$ be a joint innovation vector. Assume that the induced coordinate-outcome process is stationary and ergodic, or that ξ_t and the cross-moment expectations satisfy the mixing regime above. Define

$$\mathcal{P}_{j,t}(L(b)) = L(b^{(j,t)}), \quad b_{rs}^{(j,t)} = b_{rs} \exp(a \xi_{j,t,rs}), \quad (r, s) \in \mathcal{E}_m.$$

This map preserves nonnegative edge weights and the support of the network. If the latent and reference edge weights and $|\xi_{j,t,rs}|$ are uniformly bounded and a ranges over a bounded interval, then $\|\mathcal{P}_{j,t}(L(b)) - L(b)\|_F \leq Ca$ uniformly in j and t , for a finite constant C . With $\Phi(L) = L$, graph Laplacians form a convex subset of the Hilbert space of symmetric matrices. Thus the bounded-range and stability conditions hold, and Proposition S1 verifies Assumption 3.2 whenever the empirical matrix cross-moments satisfy its temporal condition.

S.3.4 Compositional data

For the spherical representation based on the square-root transformation, let $\mathbf{z} = \sqrt{\mathbf{y}} \in \mathcal{S}_+^{d-1}$. Let $\zeta_t = (\zeta_{1,t}, \dots, \zeta_{J+1,t})$ be a joint innovation vector, with $\zeta_{j,t} \in \mathbb{R}^d$. Assume that the induced loss process is stationary and ergodic, or that ζ_t and the loss expectations satisfy the mixing regime above, and define $v_{j,t}(\mathbf{z}) = (\mathbf{I}_d - \mathbf{z}\mathbf{z}^\top)\zeta_{j,t} \in T_{\mathbf{z}}\mathcal{S}^{d-1}$. The corresponding perturbation is

$$\mathcal{P}_{j,t}(\mathbf{z}) = \text{Exp}_{\mathbf{z}}(av_{j,t}(\mathbf{z})).$$

Assume that the latent objects satisfy $\inf_{1 \leq j \leq J+1} d(U_j, \partial\mathcal{S}_+^{d-1}) \geq \eta$ for some $\eta > 0$, that $\|\zeta_{j,t}\| \leq M$ almost surely, and that $aM < \eta$. Since $\|v_{j,t}(U_j)\| \leq M$, these conditions ensure that $\mathcal{P}_{j,t}(U_j) \in \mathcal{S}_+^{d-1}$ almost surely. Assume that there are a center $z_c \in \mathcal{S}_+^{d-1}$ and radii $0 < \rho < r < \pi/4$ such that all unperturbed outcomes $g_t(U_j)$ lie in the closed spherical ball of radius ρ centered at z_c . Suppose also that the maps g_t are uniformly Lipschitz on the relevant range, $\|\zeta_{j,t}\|$ is uniformly bounded, and a is small enough that all perturbed outcomes $g_t(\mathcal{P}_{j,t}(U_j))$ lie in the concentric closed ball K of radius r . The perturbed compositions satisfy $d(\mathcal{P}_{j,t}(\mathbf{z}), \mathbf{z}) \leq Ca$, and their weighted Fréchet means remain in K . On K , the squared-distance objectives are uniformly strongly geodesically convex. Continuity of the weighted squared-distance objective, compactness of K , and uniqueness of its minimizer imply joint measurability of $m_t(\mathbf{w})$ by the measurable maximum theorem. Proposition S2 therefore verifies Assumption 3.2 under either temporal regime above.

S.3.5 SPD matrices

Let $\mathcal{M} = \text{Sym}_m^+$ be the space of $m \times m$ SPD matrices. Under the Log-Euclidean metric, perturbations can be introduced additively after applying the matrix logarithm. Specifically, let $\Xi_t = (\Xi_{1,t}, \dots, \Xi_{J+1,t})$ be a joint innovation vector with symmetric-matrix components. Assume that the induced coordinate-outcome process is stationary and ergodic, or that Ξ_t

and the cross-moment expectations satisfy the mixing regime above, and define

$$\mathcal{P}_{j,t}(A) = \exp\{\log A + a\Xi_{j,t}\}, \quad A \in \text{Sym}_m^+.$$

The map is well defined and $d_{\text{LE}}(\mathcal{P}_{j,t}(A), A) = a\|\Xi_{j,t}\|_F$.

More generally, suppose an SPD metric is represented as a Hilbert distance after an invertible transformation Φ . Let $\Xi_{j,t}$ take values in the corresponding coordinate space, and assume that $\Phi(A) + a\Xi_{j,t}$ belongs to $\Phi(\text{Sym}_m^+)$ almost surely. Then set

$$\mathcal{P}_{j,t}(A) = \Phi^{-1}(\Phi(A) + a\Xi_{j,t}).$$

For the Frobenius, power, and Log-Euclidean metrics, one may take $\Phi(A) = A$, A^p , and $\log A$, respectively. For the Log-Cholesky metric, Φ consists of the strictly lower triangular Cholesky entries and logarithms of its diagonal entries. For the identity and power transformations, membership in the transformed SPD space follows, for example, from a uniform eigenvalue margin and a sufficiently small upper bound on a . If the transformed latent and reference matrices and $\Xi_{j,t}$ are uniformly bounded and a ranges over such a bounded interval, the standing bounded-range condition holds and the perturbations have size $O(a)$ in the corresponding metric. Proposition S1 then verifies Assumption 3.2 under its cross-moment condition.

S.3.6 Univariate probability distributions

Consider univariate probability distributions on $\mathbb{R}$ with finite second moments, equipped with the 2-Wasserstein metric. Let $\xi_t = (\xi_{1,t}, \dots, \xi_{J+1,t})$ be a joint random-function innovation vector. Assume that the induced coordinate-outcome process is stationary and ergodic, or that ξ_t and the cross-moment expectations satisfy the mixing regime above. Assume also that, almost surely, $T_{j,t,a}(x) = x + a\xi_{j,t}(x)$ is nondecreasing and Borel measurable and that $T_{j,t,a}(X)$ has a finite second moment whenever X does. Define $\mathcal{P}_{j,t}(\mu) = (T_{j,t,a})_{\#}\mu$.

Its quantile function is

$$F_{\mathcal{P}_{j,t}(\mu)}^{-1}(s) = T_{j,t,a}(F_{\mu}^{-1}(s)), \quad \text{for Lebesgue-almost every } s \in (0, 1).$$

Moreover,

$$d_{\mathcal{W}}(\mathcal{P}_{j,t}(\mu), \mu) \leq a \left[\int |\xi_{j,t}(x)|^2 d\mu(x) \right]^{1/2}.$$

Uniform bounds on the right-hand side therefore give perturbations of order a . A common bounded support for the latent and reference distributions, together with a uniform bound on $\xi_{j,t}$ over that support and bounded a , gives the standing bounded-range condition. The quantile map $\Phi(\mu) = F_{\mu}^{-1}$ embeds this space isometrically into a convex subset of $L^2(0, 1)$, so Proposition S1 verifies Assumption 3.2 under its cross-moment condition.

S.3.7 Functional data

Let $\mathcal{M} = L^2(\mathcal{T})$, where $\mathcal{T}$ is a compact interval, equipped with the usual L^2 metric. Let $\xi_t = (\xi_{1,t}, \dots, \xi_{J+1,t})$ be a joint $L^2(\mathcal{T})$ -valued innovation vector. Assume that the induced coordinate-outcome process is stationary and ergodic, or that ξ_t and the cross-moment expectations satisfy the mixing regime above, and define

$$\mathcal{P}_{j,t}(f) = f + a\xi_{j,t}, \quad f \in L^2(\mathcal{T}).$$

Then $d_{L^2}(\mathcal{P}_{j,t}(f), f) = a\|\xi_{j,t}\|_{L^2}$. Uniformly bounded latent and reference L^2 norms, almost surely bounded innovation norms, and bounded a give the standing bounded-range condition; uniformly bounded first moments suffice for the order- a expectation bound used in Corollary S1. Taking Φ to be the identity, Proposition S1 verifies Assumption 3.2 under its cross-moment condition.

S.3.8 Summary of the verification

Propositions S1 and S2 verify every component of Assumption 3.2: pointwise criterion convergence, continuity and existence of the possibly nonunique limiting minimizer set,

separation from that set, existence of a measurable empirical minimizer, and the uniform local stability condition. The temporal assumptions are used specifically to establish the law of large numbers in Assumption 3.2(i); the geometric arguments establish Assumption 3.2(ii).

For completeness, the same constructions also provide primitive conditions for Assumption S1(ii). In the Hilbert-coordinate cases, for donor tuples $x = (x_2, \dots, x_{J+1})$ and $x' = (x'_2, \dots, x'_{J+1})$,

$$d(m^{(w)}(x), m^{(w)}(x')) \leq \max_{2 \leq j \leq J+1} d(x_j, x'_j).$$

If the maps g_t are uniformly Lipschitz on the relevant ranges, the order- a perturbation bounds therefore give Assumption S1(ii) with $\delta_{T_0} = a_{T_0}$. For spherical compositions, the same uniform Lipschitz condition on g_t , together with the quadratic-growth comparison, gives an order- $a_{T_0}^{1/2}$ bound, so one may take $\delta_{T_0} = a_{T_0}^{1/2}$; a uniform Lipschitz sensitivity bound for the Fréchet means gives the sharper order a_{T_0} . The latent balancing and persistence requirements used in Assumption S1(i) are treated in Section S.2.

Table S1: Sufficient conditions verifying the standing bounded-range condition and Assumption 3.2. The temporal condition in the last column is imposed on the induced process or the full innovation vector across units, as specified in Section S.3.2.

Outcome	Range and geometry	Verification of Assumption 3.2
Networks	Uniformly bounded latent and reference edge weights and bounded edge innovations; Frobenius coordinates.	Proposition S1; stationary-ergodic or mixing convergence of the matrix cross-moments.
Compositions	A common closed spherical ball of radius less than $\pi/4$ in the interior of the positive orthant; uniformly bounded tangent perturbations.	Proposition S2; stationary-ergodic or mixing law of large numbers for the bounded loss process.
SPD matrices	Uniformly bounded coordinates under the identity, power, logarithm, or Log-Cholesky transformation, as appropriate; bounded coordinate innovations that remain in the transformed SPD space.	Proposition S1; stationary-ergodic or mixing convergence of the transformed-matrix cross-moments.
Distributions	Common bounded support, or uniformly bounded quantile functions in $L^2(0, 1)$; perturbations remain in the corresponding bounded set.	Proposition S1 with quantile coordinates; stationary-ergodic or mixing convergence of the quantile cross-moments.
Functional data	Uniformly bounded latent and reference L^2 norms and almost surely bounded innovation norms.	Proposition S1 with identity coordinates; stationary-ergodic or mixing convergence of the L^2 cross-moments.

S.4 Proofs

S.4.1 Proof of Theorem 3.1

Proof of (i). For $t \leq T_0$, $Y_{j,t} = Y_{j,t}(N)$ for every unit, and write $m_t(\mathbf{w}) = \hat{g}_{2:J+1,t}^{(\mathbf{w})}$. The standing bounded-range condition in Section 2 places the treated and donor outcomes in the closed ball of radius D centered at ν_0 . By optimality of the weighted Fréchet mean,

$$\sum_{j=2}^{J+1} w_j d^2(Y_{j,t}(N), m_t(\mathbf{w})) \leq \sum_{j=2}^{J+1} w_j d^2(Y_{j,t}(N), \nu_0) \leq D^2.$$

The triangle and Cauchy–Schwarz inequalities therefore give

$$d(\nu_0, m_t(\mathbf{w})) \leq D + \sum_{j=2}^{J+1} w_j d(Y_{j,t}(N), m_t(\mathbf{w})) \leq 2D.$$

Consequently, $d(Y_{1,t}(N), m_t(\mathbf{w})) \leq 3D$ for every $\mathbf{w} \in \Delta^{J-1}$. For any $\mathbf{w}_1, \mathbf{w}_2 \in \Delta^{J-1}$,

$$|\widehat{Q}(\mathbf{w}_1) - \widehat{Q}(\mathbf{w}_2)| \leq \frac{6D}{T_0} \sum_{t=1}^{T_0} d(m_t(\mathbf{w}_1), m_t(\mathbf{w}_2)). \quad (\text{S8})$$

Assumption 3.2(ii) and (S8) imply stochastic equicontinuity of $\widehat{Q}$ on Δ^{J-1} .

Pointwise convergence in Assumption 3.2(i) gives finite-dimensional convergence of $\widehat{Q}$ to the deterministic process $\widetilde{Q}$ and pointwise asymptotic tightness. Moreover, the simplex is compact and hence totally bounded under the ℓ_1 metric, and continuity of $\widetilde{Q}$ implies uniform continuity. Stochastic equicontinuity and Theorem 1.5.7 of [van der Vaart and Wellner \(2023\)](#) give asymptotic tightness of $\widehat{Q}$. Since the deterministic continuous process $\widetilde{Q}$ is a tight element of $\ell^\infty(\Delta^{J-1})$, finite-dimensional convergence and Theorem 1.5.4 of [van der Vaart and Wellner \(2023\)](#) then give

$$\widehat{Q} \rightsquigarrow \widetilde{Q} \quad \text{in } \ell^\infty(\Delta^{J-1}).$$

Applying the continuous mapping theorem (Theorem 1.3.6 of [van der Vaart and Wellner \(2023\)](#)) to the continuous map $f \mapsto \|f - \widetilde{Q}\|_\infty$, and using that the limit is deterministic, yields

$$\sup_{\mathbf{w} \in \Delta^{J-1}} |\widehat{Q}(\mathbf{w}) - \widetilde{Q}(\mathbf{w})| = o_p(1). \quad (\text{S9})$$

We next adapt the M-estimator consistency argument in Corollary 3.2.3 of [van der Vaart and Wellner \(2023\)](#) to the possibly non-singleton minimizer set. Fix $\varepsilon > 0$. If every $\mathbf{w} \in \Delta^{J-1}$ satisfies $d_1(\mathbf{w}, \widetilde{\mathcal{W}}) < \varepsilon$, the desired conclusion is immediate. Otherwise, the separation condition in Assumption 3.2(i) gives

$$c_\varepsilon = \inf_{\{\mathbf{w} \in \Delta^{J-1} : d_1(\mathbf{w}, \widetilde{\mathcal{W}}) \geq \varepsilon\}} \widetilde{Q}(\mathbf{w}) - \min_{\mathbf{w} \in \Delta^{J-1}} \widetilde{Q}(\mathbf{w}) > 0.$$

Let $q_0 = \min_{\mathbf{w} \in \Delta^{J-1}} \widetilde{Q}(\mathbf{w})$ and fix $\mathbf{w}_0 \in \widetilde{\mathcal{W}}$. On the event that the supremum in (S9) is smaller than $c_\varepsilon/3$, any $\mathbf{w}$ with $d_1(\mathbf{w}, \widetilde{\mathcal{W}}) \geq \varepsilon$ satisfies $\widehat{Q}(\mathbf{w}) > q_0 + 2c_\varepsilon/3$, whereas

$\widehat{Q}(\mathbf{w}_0) < q_0 + c_\varepsilon/3$. An empirical minimizer therefore cannot be at distance at least ε from $\widetilde{\mathcal{W}}$. Hence $d_1(\widehat{\mathbf{w}}, \widetilde{\mathcal{W}}) = o_p(1)$.

The nearest-point correspondence onto the deterministic compact set $\widetilde{\mathcal{W}}$ has nonempty compact values and a measurable selection. Let $\widetilde{\mathbf{w}}$ be the selection in the theorem, and let $L_{T_0} = O_p(1)$ denote the supremum in the post-treatment stability condition. At $t = T_0 + h$,

$$d(\widehat{g}_{2:J+1,t}^{(\widehat{\mathbf{w}})}, \widehat{g}_{2:J+1,t}^{(\widetilde{\mathbf{w}})}) \leq L_{T_0} \|\widehat{\mathbf{w}} - \widetilde{\mathbf{w}}\|_1^\kappa = o_p(1).$$

The triangle inequality now gives the error bound stated in Theorem 3.1(i). $\square$

Proof of (ii). For all sufficiently large T_0 , $\widetilde{\mathbf{w}}$ belongs to $\bigcap_{s=1}^{T_0} \mathbb{W}_s^*$ by the additional inclusion in part (ii). Assumption 3.1(ii) then implies $\widetilde{\mathbf{w}} \in \mathbb{W}_{T_0+h}^*$ for every fixed $h \geq 1$, which is exactly the stated equality. $\square$

We next give sufficient conditions for consistent recovery of the untreated counterfactual. Let $\{\mathcal{P}_{j,t}^{(T_0)}\}$ be a triangular array of perturbation maps and write

$$Y_{j,t}^{(T_0)}(N) = g_t(\mathcal{P}_{j,t}^{(T_0)}(U_j)), \quad \widehat{g}_{2:J+1,t,T_0}^{(\mathbf{w})} = \arg \min_{\nu \in \mathcal{M}} \sum_{j=2}^{J+1} w_j d^2(g_t(\mathcal{P}_{j,t}^{(T_0)}(U_j)), \nu).$$

Define the row- T_0 empirical criterion and GSC weights by

$$\widehat{Q}_{T_0}(\mathbf{w}) = \frac{1}{T_0} \sum_{s=1}^{T_0} d^2(Y_{1,s}^{(T_0)}(N), \widehat{g}_{2:J+1,s,T_0}^{(\mathbf{w})}), \quad \widehat{\mathbf{w}}_{T_0} \in \arg \min_{\mathbf{w} \in \Delta^{J-1}} \widehat{Q}_{T_0}(\mathbf{w}).$$

For $t > T_0$, define the row- T_0 GSC prediction by

$$\widehat{Y}_{1,t,T_0}^{(\text{GSC})} = \widehat{g}_{2:J+1,t,T_0}^{(\widehat{\mathbf{w}}_{T_0})}.$$

All notation in Theorem 3.1 is understood row by row for this array. Define the latent finite-sample criterion and its limit by

$$Q_{T_0}^*(\mathbf{w}) = \frac{1}{T_0} \sum_{t=1}^{T_0} d^2(g_t(U_1), g_{2:J+1,t}^{(\mathbf{w})}), \quad Q^*(\mathbf{w}) = \lim_{T_0 \rightarrow \infty} Q_{T_0}^*(\mathbf{w}),$$

and let $\mathcal{W}^* = \arg \min_{\mathbf{w} \in \Delta^{J-1}} Q^*(\mathbf{w})$.

Assumption S1.

- (i) The function Q^* exists as a finite continuous function, $\mathcal{W}^*$ is nonempty, and every limiting oracle weight exactly balances the latent pre-treatment outcomes for all sufficiently large T_0 :

$$\mathcal{W}^* \subset \bigcap_{t=1}^{T_0} \mathbb{W}_t^*.$$

- (ii) There is a sequence $\delta_{T_0} > 0$ with $\delta_{T_0} \rightarrow 0$ such that, for every fixed $H \geq 1$,

$$\begin{aligned} \max_{1 \leq t \leq T_0+H} \mathbb{E} \left[d(g_t(\mathcal{P}_{1,t}^{(T_0)}(U_1)), g_t(U_1)) \right] &= O(\delta_{T_0}), \\ \max_{1 \leq t \leq T_0+H} \mathbb{E} \left[\sup_{\mathbf{w} \in \Delta^{J-1}} d(\widehat{g}_{2:J+1,t,T_0}^{(\mathbf{w})}, g_{2:J+1,t}^{(\mathbf{w})}) \right] &= O(\delta_{T_0}). \end{aligned}$$

Corollary S1. Suppose that Assumptions 3.1, 3.2, and [S1](#) hold for the triangular array above. For a fixed $h \geq 1$, suppose also that, for some $\kappa \in (0, 1]$,

$$\sup_{\substack{\mathbf{w}_1, \mathbf{w}_2 \in \Delta^{J-1} \\ \mathbf{w}_1 \neq \mathbf{w}_2}} \frac{d(\widehat{g}_{2:J+1,T_0+h,T_0}^{(\mathbf{w}_1)}, \widehat{g}_{2:J+1,T_0+h,T_0}^{(\mathbf{w}_2)})}{\|\mathbf{w}_1 - \mathbf{w}_2\|_1^\kappa} = O_p(1).$$

Then

$$d(Y_{1,T_0+h}^{(T_0)}(N), \widehat{Y}_{1,T_0+h,T_0}^{(\text{GSC})}) = o_p(1).$$

No uniqueness of the oracle weights in $\mathcal{W}^*$ is required.

Proof. Let $\widehat{Q}_{T_0}$ denote the empirical perturbed criterion in row T_0 . Under the standing bounded-range condition in Section 2, the argument used in the proof of Theorem 3.1 places every latent and perturbed synthetic object within distance $2D$ of ν_0 . Hence, uniformly over $\mathbf{w} \in \Delta^{J-1}$,

$$\begin{aligned} |\mathbb{E}[\widehat{Q}_{T_0}(\mathbf{w})] - Q_{T_0}^*(\mathbf{w})| &\leq \frac{6D}{T_0} \sum_{t=1}^{T_0} \mathbb{E} \left[d(g_t(\mathcal{P}_{1,t}^{(T_0)}(U_1)), g_t(U_1)) \right. \\ &\quad \left. + \sup_{\mathbf{v} \in \Delta^{J-1}} d(\widehat{g}_{2:J+1,t,T_0}^{(\mathbf{v})}, g_{2:J+1,t}^{(\mathbf{v})}) \right] \\ &= O(\delta_{T_0}). \end{aligned}$$

Taking limits gives $\widetilde{Q}(\mathbf{w}) = Q^*(\mathbf{w})$ for every $\mathbf{w}$, and therefore $\widetilde{\mathcal{W}} = \mathcal{W}^*$. Assumption S1(i), together with Assumption 3.1(ii), satisfies the conditions of Theorem 3.1(ii), so the latent-signal term in the bound in Theorem 3.1(i) is zero at $t = T_0 + h$.

Let $\widetilde{\mathbf{w}} \in \mathcal{W}^*$ be the nearest-point selection used in Theorem 3.1. Markov's inequality and Assumption S1(ii) give

$$d(g_t(\mathcal{P}_{1,t}^{(T_0)}(U_1)), g_t(U_1)) = O_p(\delta_{T_0}),$$

$$d(\widehat{g}_{2:J+1,t,T_0}^{(\widetilde{\mathbf{w}})}, g_{2:J+1,t}^{(\widetilde{\mathbf{w}})}) = O_p(\delta_{T_0})$$

at $t = T_0 + h$. Applying the error bound in Theorem 3.1(i) completes the proof because $\delta_{T_0} \rightarrow 0$. $\square$

S.4.2 Proof of Proposition 4.1

Condition on the training-data σ -field $\mathcal{F}_{\text{tr}}$, so that $\widehat{\mathbf{w}}^{\text{tr}}$ is fixed. For $z = (z_1, \dots, z_{J+1}) \in \mathcal{M}^{J+1}$, let

$$m(z_2, \dots, z_{J+1}) = \arg \min_{\nu \in \mathcal{M}} \sum_{j=2}^{J+1} \widehat{w}_j^{\text{tr}} d^2(\nu, g(z_j)), \quad r(z) = d(g(z_1), m(z_2, \dots, z_{J+1})).$$

By the assumed uniqueness and measurability of the weighted Fréchet mean map, r is measurable. If $g_s = g$, then $R_s = r(Z_s)$. If $g_s = h_s \circ g$, isometric equivariance and uniqueness of the weighted Fréchet mean give

$$\widehat{Y}_{1,s}^{(\text{GSC}, \text{tr})} = h_s(m(Z_{2,s}, \dots, Z_{J+1,s})).$$

Therefore

$$R_s = d(h_s(g(Z_{1,s})), h_s(m(Z_{2,s}, \dots, Z_{J+1,s}))) = r(Z_s).$$

In either case, every score is the same measurable scalar function of the corresponding perturbation vector. Conditional exchangeability of the Z_s therefore implies conditional exchangeability of the scores, which gives Assumption 4.1. No independence among the components of Z_s is used.

For the global procedure, each block score is the same measurable function of the corresponding equal-length block of perturbation vectors. Exchangeability of these block vectors therefore implies exchangeability of the block scores in Assumption 4.2. Conditional i.i.d. sampling of the individual Z_s over time is a sufficient special case for both conclusions.

S.4.3 Proof of Theorem 4.1

For each $s \in \mathcal{T}_{\text{cal}}$, since $s \leq T_0$, the treated unit is untreated during the calibration period and hence $Y_{1,s} = Y_{1,s}(N)$. Therefore the calibration scores satisfy

$$R_s = d\left(Y_{1,s}, \hat{Y}_{1,s}^{(\text{GSC}, \text{tr})}\right) = d\left(Y_{1,s}(N), \hat{Y}_{1,s}^{(\text{GSC}, \text{tr})}\right).$$

By Assumption 4.1, conditional on the data used to estimate the GSC training weights, the collection $\{R_s : s \in \mathcal{T}_{\text{cal}}\} \cup \{R_t\}$ is exchangeable, where $R_t = d\left(Y_{1,t}(N), \hat{Y}_{1,t}^{(\text{GSC}, \text{tr})}\right)$.

If $q_{1-\alpha} = +\infty$, then $\mathcal{C}_t^{(\text{GSC})}(1-\alpha) = \mathcal{M}$, and the coverage claim is immediate. Hence it remains to consider the case $q_{1-\alpha} = R_{(k_\alpha)}$ for $k_\alpha \leq m$. By definition of the prediction region,

$$Y_{1,t}(N) \notin \mathcal{C}_t^{(\text{GSC})}(1-\alpha)$$

if and only if

$$d\left(Y_{1,t}(N), \hat{Y}_{1,t}^{(\text{GSC}, \text{tr})}\right) > q_{1-\alpha},$$

or equivalently, $R_t > R_{(k_\alpha)}$. Since the $m+1$ scores are exchangeable, the rank of R_t among $\{R_s : s \in \mathcal{T}_{\text{cal}}\} \cup \{R_t\}$ is uniformly distributed over $\{1, \dots, m+1\}$ in the absence of ties. In the presence of ties, the use of the strict inequality in the event $\{R_t > R_{(k_\alpha)}\}$ yields conservative coverage by the standard conformal prediction argument. Consequently,

$$\Pr\{R_t > R_{(k_\alpha)}\} \leq \frac{m+1-k_\alpha}{m+1} \leq \alpha.$$

Therefore,

$$\Pr\left\{Y_{1,t}(N) \in \mathcal{C}_t^{(\text{GSC})}(1-\alpha)\right\} \geq 1-\alpha.$$

Under the pointwise null hypothesis

$$H_{0,t} : Y_{1,t}(I) = Y_{1,t}(N),$$

we have

$$Y_{1,t} = Y_{1,t}(I) = Y_{1,t}(N).$$

Hence

$$\Pr_{H_{0,t}}\{\phi_t = 1\} = \Pr_{H_{0,t}}\left\{Y_{1,t}(N) \notin \mathcal{C}_t^{(\text{GSC})}(1 - \alpha)\right\} \leq \alpha,$$

which establishes the finite-sample validity of the test.

S.4.4 Proof of Corollary 4.1

Since $(\mathcal{M}, d)$ is uniquely geodesic, for each $\omega \in \mathcal{M}$ there is a unique geodesic $\gamma_{\omega, Y_{1,t}}$. On the event

$$Y_{1,t}(N) \in \mathcal{C}_t^{(\text{GSC})}(1 - \alpha),$$

we have

$$\tau_t = \gamma_{Y_{1,t}(N), Y_{1,t}(I)} = \gamma_{Y_{1,t}(N), Y_{1,t}} \in \left\{\gamma_{\omega, Y_{1,t}} : \omega \in \mathcal{C}_t^{(\text{GSC})}(1 - \alpha)\right\} = \mathcal{G}_t(1 - \alpha),$$

where we used $Y_{1,t} = Y_{1,t}(I)$ for $t > T_0$. Therefore

$$\Pr\{\tau_t \in \mathcal{G}_t(1 - \alpha)\} \geq \Pr\{Y_{1,t}(N) \in \mathcal{C}_t^{(\text{GSC})}(1 - \alpha)\} \geq 1 - \alpha,$$

by Theorem 4.1. This proves the result.

S.4.5 Proof of Corollary 4.2

Since $Y_{1,t} = Y_{1,t}(I)$ for $t > T_0$,

$$\widehat{\delta}_t = d(\widehat{Y}_{1,t}^{(\text{GSC}, \text{tr})}, Y_{1,t}(I)).$$

By the reverse triangle inequality,

$$|\widehat{\delta}_t - \delta_t| = \left|d(\widehat{Y}_{1,t}^{(\text{GSC}, \text{tr})}, Y_{1,t}(I)) - d(Y_{1,t}(N), Y_{1,t}(I))\right| \leq d(\widehat{Y}_{1,t}^{(\text{GSC}, \text{tr})}, Y_{1,t}(N)).$$

On the event $Y_{1,t}(N) \in \mathcal{C}_t^{(\text{GSC})}(1 - \alpha)$,

$$d(\hat{Y}_{1,t}^{(\text{GSC}, \text{tr})}, Y_{1,t}(N)) \leq q_{1-\alpha}.$$

Thus, on this event,

$$\delta_t \in [\hat{\delta}_t - q_{1-\alpha}, \hat{\delta}_t + q_{1-\alpha}].$$

Since $\delta_t \geq 0$, the lower endpoint can be replaced by $\{\hat{\delta}_t - q_{1-\alpha}\}_+$. The coverage statement follows from Theorem 4.1.

S.4.6 Proof of Theorem 4.2

Under H_0^{glob} , Assumption 4.2 implies that, conditional on the data used to estimate the GSC weights, the block scores $S_r(\mathcal{B}_1), \dots, S_r(\mathcal{B}_K), S_r(\mathcal{B}_*)$ are exchangeable.

By definition, $\hat{p}_r$ is the conformal rank p -value associated with the post-treatment block score $S_r(\mathcal{B}_*)$ relative to the calibration block scores $S_r(\mathcal{B}_1), \dots, S_r(\mathcal{B}_K)$.

If there are no ties among the $K + 1$ block scores, exchangeability implies that the rank of $S_r(\mathcal{B}_*)$ among $S_r(\mathcal{B}_1), \dots, S_r(\mathcal{B}_K), S_r(\mathcal{B}_*)$ is uniformly distributed on $\{1, \dots, K + 1\}$. Consequently,

$$\Pr_{H_0^{\text{glob}}} \{\hat{p}_r \leq \alpha\} \leq \alpha.$$

In the presence of ties, the use of the non-strict comparison $S_r(\mathcal{B}_k) \geq S_r(\mathcal{B}_*)$ in the definition of $\hat{p}_r$ yields a conservative conformal p -value. Therefore the same inequality continues to hold:

$$\Pr_{H_0^{\text{glob}}} \{\hat{p}_r \leq \alpha\} \leq \alpha.$$

Since

$$\phi_r^{\text{glob}} = 1\{\hat{p}_r \leq \alpha\},$$

it follows that

$$\Pr_{H_0^{\text{glob}}} \{\phi_r^{\text{glob}} = 1\} = \Pr_{H_0^{\text{glob}}} \{\hat{p}_r \leq \alpha\} \leq \alpha.$$

Hence the global test has finite-sample size at most α .

S.5 Additional simulation and application results

We report the SPD counterpart to the network simulation in Section 5.2 of the main paper. In Log-Euclidean coordinates, the donor latent objects are 10×10 symmetric matrices with bounded entries, the treated latent object is their convex combination with oracle weights $\mathbf{w}^*$, and the reference object contains an analogous periodic common component. Independent perturbation matrices are formed by symmetrizing matrices of standardized $\text{Unif}[-\sqrt{3}, \sqrt{3}]$ entries and are added in log coordinates at the fixed scale $\sigma = .10$. Exponentiation maps the resulting symmetric matrices to SPD outcomes. All other estimation settings match the network experiment.

Quantity	$T_0 = 20$	$T_0 = 100$	$T_0 = 500$
Post-treatment prediction gap	.048 (<.001)	.022 (<.001)	.010 (<.001)
Realized post-treatment error	.387 (.001)	.386 (.001)	.386 (.001)
$\ \hat{\mathbf{w}} - \tilde{\mathbf{w}}\ _1$	.174 (.002)	.081 (.001)	.036 (<.001)
$\ \tilde{\mathbf{w}} - \mathbf{w}^*\ _1$	.115 (.001)	.118 (.001)	.115 (.001)

Table S2: SPD estimation results for $T_0 \in \{20, 100, 500\}$ and $\sigma = .10$, based on 500 replications. The post-treatment prediction gap compares the post-treatment predictions based on $\hat{\mathbf{w}}$ and $\tilde{\mathbf{w}}$, while the realized post-treatment error compares the untreated post-treatment outcome with the prediction based on $\hat{\mathbf{w}}$. The remaining rows report the ℓ_1 weight-estimation error and oracle-weight gap. Entries are Monte Carlo means with Monte Carlo standard errors in parentheses.

Table S2 shows that the post-treatment prediction gap and weight-estimation error decrease with T_0 , whereas the realized post-treatment error and oracle-weight gap remain stable, consistent with the decomposition in Theorem 3.1. This pattern agrees with the network results.

To separate the effects of pre-treatment length and perturbation magnitude, we repeat the estimation experiment over $T_0 \in \{20, 100, 500\}$ and $\sigma \in \{20^{-1/2}, 100^{-1/2}, 500^{-1/2}\}$. For each perturbation scale σ , let $\tilde{\mathbf{w}}_\sigma$ denote the selected minimizer of $\tilde{Q}$ under the corresponding simulation design. The latent objects and standardized innovations are paired across perturbation scales within each outcome, T_0 , and replication. Table S3 reports the four

diagnostics used in the fixed- σ experiments.

Outcome	T_0	$\sigma = 20^{-1/2}$	$\sigma = 100^{-1/2}$	$\sigma = 500^{-1/2}$
<i>Panel A: Post-treatment prediction gap</i>				
Network	20	.161 (.002)	.068 (.001)	.030 (<.001)
	100	.076 (.001)	.031 (<.001)	.014 (<.001)
	500	.034 (<.001)	.014 (<.001)	.006 (<.001)
SPD	20	.107 (.001)	.048 (<.001)	.022 (<.001)
	100	.050 (<.001)	.022 (<.001)	.010 (<.001)
	500	.022 (<.001)	.010 (<.001)	.005 (<.001)
<i>Panel B: Realized post-treatment error</i>				
Network	20	.846 (.006)	.374 (.003)	.167 (.001)
	100	.844 (.006)	.373 (.003)	.167 (.001)
	500	.830 (.006)	.367 (.003)	.164 (.001)
SPD	20	.861 (.003)	.387 (.001)	.174 (.001)
	100	.859 (.003)	.386 (.001)	.173 (.001)
	500	.859 (.003)	.386 (.001)	.173 (.001)
<i>Panel C: Weight-estimation error $\ \hat{\mathbf{w}} - \tilde{\mathbf{w}}_\sigma\ _1$</i>				
Network	20	.253 (.002)	.120 (.001)	.053 (.001)
	100	.121 (.001)	.055 (.001)	.024 (<.001)
	500	.054 (.001)	.025 (<.001)	.011 (<.001)
SPD	20	.301 (.002)	.174 (.002)	.086 (.001)
	100	.140 (.001)	.081 (.001)	.039 (<.001)
	500	.062 (<.001)	.036 (<.001)	.018 (<.001)
<i>Panel D: Oracle-weight gap $\ \tilde{\mathbf{w}}_\sigma - \mathbf{w}^*\ _1$</i>				
Network	20	.194 (.002)	.063 (.001)	.015 (<.001)
	100	.193 (.002)	.062 (.001)	.014 (<.001)
	500	.194 (.002)	.062 (.001)	.015 (<.001)
SPD	20	.315 (.003)	.115 (.001)	.029 (<.001)
	100	.320 (.003)	.118 (.001)	.030 (<.001)
	500	.316 (.003)	.115 (.001)	.029 (<.001)

Table S3: Network and SPD estimation results for $T_0 \in \{20, 100, 500\}$ and $\sigma \in \{20^{-1/2}, 100^{-1/2}, 500^{-1/2}\}$, based on 500 replications. The post-treatment prediction gap compares the post-treatment predictions based on $\hat{\mathbf{w}}$ and $\tilde{\mathbf{w}}_\sigma$, while the realized post-treatment error compares the untreated post-treatment outcome with the prediction based on $\hat{\mathbf{w}}$. The remaining panels report the ℓ_1 weight-estimation error and oracle-weight gap. Entries are Monte Carlo means with Monte Carlo standard errors in parentheses. Bold entries satisfy $\sigma = T_0^{-1/2}$.

At each fixed σ , the post-treatment prediction gap and weight-estimation error decrease as T_0 increases, whereas the realized post-treatment error and oracle-weight gap remain stable. Along the diagonal $\sigma = T_0^{-1/2}$, the realized post-treatment errors retain an approximately $T_0^{-1/2}$ pattern: they are .846, .373, and .164 for networks and .861, .386, and

.173 for SPD matrices. The post-treatment prediction gaps decrease more rapidly along the same diagonal, from .161 to .031 and .006 for networks and from .107 to .022 and .005 for SPD matrices, an empirical pattern close to T_0^{-1} in these designs. Thus, in these designs, smaller perturbations and longer pre-treatment periods jointly reduce the post-treatment prediction gap and realized post-treatment error, while the newly realized post-treatment perturbation remains the leading component of the latter.

The SPD inference experiment uses the same training-calibration split, post-treatment length, and effect grid as the network experiment and sets the global block score to $r = 2$. In log coordinates, the imposed treatment satisfies

$$Y_{1,t}(I) = \exp \left\{ \log Y_{1,t}(N) + aI_{10}/\sqrt{10} \right\},$$

so its Log-Euclidean magnitude is exactly a . Under the null, pointwise counterfactual coverage is .954, and pointwise and global rejection probabilities are .046 and .056, respectively. At $a = .25$, pointwise and global power are .726 and .992, and both equal one for $a \geq .50$. Figure S1 displays the rejection probabilities.

For the employment comparison, fitting conventional SCM separately to each of the three shares produces the raw post-treatment predictions summarized in Table S4. The separate fits do not impose a common donor-weight vector and therefore do not preserve the unit-sum constraint. Closing the three predictions is an additional post-processing step that changes the componentwise estimates; the maximum change to a single component is .003642 over the held-out pre-treatment period and .001996 after treatment.

Year	Primary	Secondary	Tertiary	Sum
2011	.055462	.237262	.708622	1.001346
2012	.053831	.237325	.711657	1.002813
2013	.052093	.229994	.716648	.998735

Table S4: Raw predictions from three separately fitted componentwise SCMs in the employment application.

Table S5 gives the complete pointwise inference results for both applications. The 90%

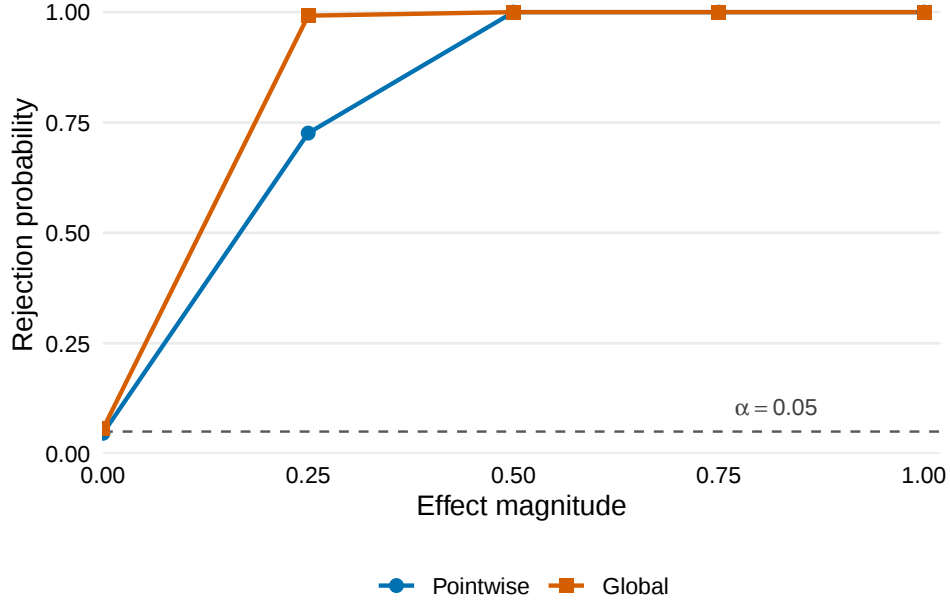


Figure S1: Empirical rejection probabilities for the pointwise and global tests in the SPD design over 500 replications. The dashed line marks the nominal level $\alpha = .05$; values at zero effect report empirical size.

split-conformal intervals use 11 calibration periods for employment and 10 for fertility. At the 95% level, the split-conformal quantile is $+\infty$ in both applications because the calibration samples are too short to attain that rank.

Application	Period	$\widehat{\delta}_t$	90% interval	p -value
Employment	2011	.0421	[.0066, .0776]	1/12
	2012	.0561	[.0206, .0916]	1/12
	2013	.0646	[.0291, .1001]	1/12
Fertility	1972	.0813	[.0207, .1419]	1/11
	1973	.1279	[.0673, .1885]	1/11
	1974	.1544	[.0938, .2150]	1/11
	1975	.1332	[.0726, .1938]	1/11

Table S5: Pointwise effect-magnitude estimates, split-conformal intervals, and conformal p -values for the applications.