A Test for Jumps in Metric-Space Conditional Means

With Applications to Compositional and Network Data

David Van Dijcke

Department of Economics, University of Michigan

July 21, 2025

Abstract

Standard methods for detecting discontinuities in conditional means are not applicable to outcomes that are complex, non-Euclidean objects like distributions, networks, or covariance matrices. This article develops a nonparametric test for jumps in conditional means when outcomes lie in a non-Euclidean metric space. Using local Fréchet regression, the method estimates a mean path on either side of a candidate cutoff. This extends existing k-sample tests to a non-parametric regression setting with metric-space valued outcomes. I establish the asymptotic distribution of the test and its consistency against contiguous alternatives. For this, I derive a central limit theorem for the local estimator of the conditional Fréchet variance and a consistent estimator of its asymptotic variance. Simulations confirm nominal size control and robust power in finite samples. Two empirical illustrations demonstrate the method's ability to reveal discontinuities missed by scalar-based tests. I find sharp changes in (i) work-from-home compositions at an income threshold for non-compete enforceability and (ii) national input-output networks following the loss of preferential U.S. trade access. These findings show the value of analyzing regression outcomes in their native metric spaces.

Keywords: Fréchet regression, regression discontinuity, changepoint, metric-space data, local polynomial regression, structural break

I thank Zach Brown, Florian Gunsilius, Andreas Hagemann, Eva Janssens, Harry Kleyer, Hans-Georg Müller, Mark Polyak, Konstantin Sonin, Kaspar Wuthrich, Yidong Zhou, and participants at the Michigan Econometrics-IO Workshop, and Econometrics, IO, and Labor Lunches for helpful discussion and comments. I gratefully acknowledge support from the Rackham Pre-Doctoral Fellowship at the University of Michigan. This research was supported in part through computational resources and services provided by Advanced Research Computing at the University of Michigan, Ann Arbor. All errors are mine. An R package accompanying this paper is available at https://github.com/Davidvandijcke/frechesTest.

1 Introduction

Testing for a discontinuity in a conditional mean function is a central task in modern econometrics and statistics, forming the basis of the regression discontinuity (RD) design for causal inference (Thistlethwaite & Campbell 1960), changepoint detection (Page 1957), and structural break tests (Chow 1960). Under the assumption that the path of conditional means evolves smoothly, discontinuities can reveal important insights into the data-generating process, such as the presence of regime changes, treatment effects, or structural breaks. While central to applied work, standard methods are predicated on scalar-valued outcomes. This is a significant limitation, as many critical questions in science involve outcomes that are inherently complex, non-Euclidean objects. For instance, how does a negative trade shock alter the *network structure* of a national economy? Or how does crossing a critical temperature threshold trigger an abrupt change in the *distribution* of daily rainfall? Answering such questions requires a methodological framework that can detect jumps or breaks in otherwise smoothly evolving distributions, networks, or other metric-space valued data.

A framework for analyzing such data can be built on the Fréchet mean, which generalizes the concept of mathematical expectation to general metric spaces (Fréchet 1948). In a regression context, interest lies specifically in the conditional Fréchet mean (Petersen & Müller 2019), which describes the central tendency of a metric-valued outcome Y as a function of a scalar covariate X. A direct test on the conditional Fréchet means themselves, however, is generally not feasible. Because these means are elements of the object space, which often lacks additive structure—for example, they are themselves networks or distributions—they cannot be directly compared through differencing. This prevents the use of standard test statistics that rely on simple algebraic comparisons.

This article develops a novel framework that overcomes this challenge. Since conditional

Fréchet means cannot be differenced, my approach instead detects discontinuities by comparing their associated conditional Fréchet variances, which are scalars. This extends the variance-based logic of Dubey & Müller (2019) from a simple k-sample (or group/multi-sample) comparison (Kruskal & Wallis 1952) to the more complex, nonparametric regression setting where conditional means and variances evolve as arbitrary smooth functions. To operationalize this, I adapt the local Fréchet regression estimator of Petersen & Müller (2019) to estimate conditional Fréchet variances on either side of a jump point, as well as by pooling across the jump. The resulting test statistic compares the pooled variance at the jump point to the one-sided variances, providing the first general method for detecting structural breaks in metric-space regression functions. To establish the asymptotic validity and consistency of this test, I derive a new central limit theorem for the local estimator of the conditional Fréchet variance, which generalizes the canonical results of Fan & Yao (1998).

In support of the theoretical results, I report simulation exercises for three metric spaces of interest: univariate density functions, graph Laplacians of networks, and covariance matrices. These results indicate that the test has excellent size control in finite samples and quickly converges to the nominal level. Moreover, it exhibits robust power, reliably detecting discontinuities in all three settings even with modest sample sizes and small jump magnitudes. A comparison to a localized adaptation of the k-sample test of Dubey & Müller (2019) to the regression setting illustrates the usefulness of the local regression approach. While the adaptation of the k-sample test does achieve nominal coverage under a piecewise-constant signal, it severely overrejects the null under a piecewise-smooth one, unlike the proposed test.

To demonstrate the test's practical utility, I present two empirical applications. First, I find a significant structural shift in the composition of work-from-home (WFH) arrangements in

Washington State at the income threshold where non-compete agreements become legally enforceable. This suggests that WFH was an important bargaining margin during the sample period, which covered the main COVID-19 years (2020-2023). Second, I detect an abrupt change in the structure of national input-output networks for countries that lose preferential trade access under the US Generalized System of Preferences after its reinstatement in 2015. This indicates that this relatively modest program, which has been subject to repeated legislative lapses and reauthorizations, nonetheless has a meaningful impact on the production networks of developing countries reliant on it. In both applications, existing tests for jumps in scalar-valued outcomes, derived from the complex objects, fail to detect statistically significant jumps. Thus, these applications illustrate the ability of the proposed method to uncover nuanced structural changes that are invisible to methods that ignore the metric space topology.

This article contributes to several strands of the statistics and econometrics literature. First, it advances the field of Fréchet regression (Petersen & Müller 2019) by providing a novel tool for inference on local Fréchet regression. It extends existing group-comparison and changepoint tests for non-Euclidean objects (Dubey & Müller 2019, 2020, Chen & Friedman 2017, Székely & Rizzo 2013, Wang et al. 2024, Müller et al. 2024, Jiang et al. 2024, Zhang, Zhang, Jones, Basner & Shou 2025, Zhang, Zhu & Shao 2025, Dubey & Zheng 2023) to a fully nonparametric regression setting. Compared to the growing body of inferential work in this area (Chen & Müller 2022, Bhattacharjee & Müller 2023, Kurisu & Otsu 2024, Van Dijcke 2025), the test proposed here generalizes to a larger class of metric spaces, works in the nonparametric regression setting, or both, and does so under milder regularity conditions, allowing for broad applicability.

Second, this article contributes to the large literatures on changepoint detection (Page 1957, Müller 1992), regression discontinuity design (RDD) (Hahn et al. 2001), and structural break estimation (Chow 1960). My contribution is to develop the first unified framework for testing discontinuities in metric-space conditional means, moving beyond functional data to general metric-space valued data such as distributions, networks, or spheres (Berkes et al. 2009). Recent research has developed testing approaches tailored to specific types of metric spaces like networks (Enikeeva & Klopp 2025, Madrid Padilla et al. 2023, Kei et al. 2024, Penaloza & Stevens 2024, Bhattacharjee et al. 2020) or covariance matrices (Dörnemann & Dette 2024, Ryan & Killick 2023, Cho et al. 2025, Avanesov & Buzun 2018, Dette et al. 2022, Aue, Hörmann, Horváth & Reimherr 2009). These methods usually assume specific generative models, such as the stochastic block model for networks, and piecewise-constant signals. By contrast, the proposed test is model-agnostic and allows for piecewise-smooth processes, in addition to being applicable to many different metric spaces.

Finally, though the focus is on testing, this article contributes to the emerging literature on RDD for non-Euclidean outcomes. For the case of distribution-valued outcomes, Van Dijcke (2025) provided the first treatment of this problem, developing bias-corrected inference and establishing the complete inferential framework necessary for applied research. For that particular metric space, the linear structure of univariate quantile functions allowed to move beyond just testing for a jump to also providing point estimates and confidence bands for the jump's magnitude. Conversely, the test developed in the current article is more general and works for metric spaces that do not have such linear structure. Subsequently and independently, Kurisu et al. (2025) proposed a conceptual extension of RDD to general geodesic metric spaces. While their framework encompasses a broad class of metric spaces, it focuses primarily on identification and estimation, leaving the question of inference—and particularly hypothesis testing—unaddressed. The present article contributes to this literature by developing the first formal test for discontinuities in conditional Fréchet means in general metric spaces, providing an inferential tool that helps operationalize metric space RDDs for practical applications. In doing so, I build upon the foundation established in Van Dijcke (2025) while extending the scope to the general metric spaces considered in Kurisu et al. (2025).

The remainder of the article develops the methodology and main theoretical results (Sections 2–3), presents simulation evidence (Section 4), illustrates the method with two empirical applications (Section 5), and concludes.

2 Setup and Motivation

Let (Ω, d) be a bounded metric space, where $d : \Omega \times \Omega \to [0, \infty)$ is a distance function. Let Y be a random object taking values in Ω, and X be a real-valued random variable representing a covariate or conditioning variable. We are interested in the behavior of Y conditional on X. Denote their joint distribution F and the conditional distribution of Y given X as $F_{Y|X}$.

2.1 Conditional Fréchet Means and Variances

The concept of a conditional Fréchet mean arises as a natural generalization of the Euclidean mean. Let $Z \in \mathbb{R}$, then the conditional expectation $E[Z \mid X = x]$ can be defined as the unique minimizer $f(x)$ of the mean squared error,

$$E[Z \mid X = x] := \operatorname*{argmin}_{f(x) \in \mathbb{R}} E[d_E(Z, f(x))^2 \mid X = x],$$

where $d_E(z_1, z_2) := \|z_1 - z_2\|$ the standard Euclidean metric.

The conditional Fréchet mean of $Y \in \Omega$ given $X = x$, where Y is a metric-space valued *random* object, is defined analogously by replacing the Euclidean metric with the more general distance metric d (Petersen & Müller 2019),

$$m_\oplus(x) := \operatorname*{argmin}_{\omega \in \Omega} M_\oplus(\omega, x), \quad \text{where } M_\oplus(\omega, x) := E\left[d^2(Y, \omega) \mid X = x\right], \tag{1}$$

and note that the expectation is taken over $F_{Y|x}$. In words, the conditional Fréchet mean $m_\oplus(x)$ is an element in Ω that minimizes the expected squared distance to Y among observations with $X = x$. In that sense, it generalizes the standard Euclidean conditional expectation, which it includes as a special case. Beyond this natural interpretation as a generalized expectation, an important benefit of working with the Fréchet mean is that it lies in the metric space itself and hence captures its topology (e.g., it is a network Laplacian), in contrast to other scalar-valued statistics that one may derive from the metric space-valued objects (e.g., the average centrality of the networks). Of course, what aspect of its topology it captures is determined by the choice of distance metric d. Nonetheless, the test proposed in this article works for many metrics and hence does not hinge on a specific choice.

The corresponding conditional Fréchet variance is,

$$V_\oplus(x) \coloneqq E\left[d^2(Y, m_\oplus(x)) \mid X = x\right] = M_\oplus(m_\oplus(x), x). \tag{2}$$

This measures the expected squared dispersion of Y around its conditional Fréchet mean $m_\oplus(x)$, analogous to the standard Euclidean variance.

We are interested in a potential discontinuity in the conditional Fréchet mean at a specific point $X = c \in \mathbb{R}$. To formalize this, define the conditional Fréchet means from the left and right of c,

$$m_{\pm,\oplus} \coloneqq \operatorname*{argmin}_{\omega \in \Omega} \left\{ \lim_{x \to c^\pm} M_\oplus(\omega, x) \right\}, \tag{3}$$

These represent the central tendencies of Y as X approaches c from below ($m_{-,\oplus}$) or above ($m_{+,\oplus}$). The corresponding "one-sided" conditional Fréchet variances are,

$$V_{\pm,\oplus} \coloneqq \lim_{x \to c^\pm} E\left[d^2(Y, m_\oplus(x)) \mid X = x\right]. \tag{4}$$

Assuming these limits exist, one can also relate them to underlying "counterfactual" Fréchet means at $X = c$ for a connection to causal inference, specifically the regression discontinuity design (Van Dijcke 2025, Kurisu et al. 2025). Since the focus of this article is on testing, I do not discuss this further here.

The null hypothesis of no discontinuity in the conditional Fréchet mean at c is,

$$H_0^{\text{mean}} : m_{+,\oplus} = m_{-,\oplus},$$

where $Y_1 = Y_2$ for $Y_1, Y_2 \in \Omega$ is defined as $d(Y_1, Y_2) = 0$. The null hypothesis of no discontinuity in the conditional Fréchet variance is,

$$H_0^{\text{var}} : V_{+,\oplus} = V_{-,\oplus}.$$

Similar to the classical analysis of variance test, the proposed test will assess H_0^{mean} and H_0^{var} jointly, that is,

$$H_0 : H_0^{\text{mean}} \cap H_0^{\text{var}},$$

while the alternative, H_1, is that at least one of these conditions does not hold.

2.2 Motivating Settings

To motivate the problem further, I now discuss several classes of problems and specific applications where detecting jumps in the conditional mean path of a metric-space valued outcome can be highly valuable. Below, I further motivate the test with two empirical illustrations on real-world data.

Example 2.1 (Regression Discontinuity Design (RDD))**.** *The RDD is a powerful quasi-experimental design for causal inference (Hahn et al. 2001, Thistlethwaite & Campbell 1960), hinging on the estimation of a discontinuity in $E[Z|X = x]$ for $Z \in \mathbb{R}$ at a known threshold c of a running variable X. The proposed framework allows the outcome to be a complex, non-Euclidean object.*

- Income Distribution Response to Democratic Governors. *Consider a "close-election" RDD which exploits the fact that when a Democratic governor's vote share (X) exceeds*

50% of the two-party vote, they become elected (T). The outcome Y_i for state i is the distribution of family income in that state. A suitable metric space for this setting is the space of cumulative distribution functions equipped with the 2-Wasserstein distance. This is the application in Van Dijcke (2025), which found negative effects of Democratic governorships on the top of the income distribution but not elsewhere. As mentioned, for this particular metric space, one can also construct tests using the uniform confidence bands developed in that article.

The test proposed here allows for formally assessing if $m_{+,\oplus} \neq m_{-,\oplus}$ in such RDD settings. See also Kurisu et al. (2025) for further examples of metric-space valued RD designs.

Example 2.2 (Changepoint Detection in Metric-Space Valued Processes)**.** *A central task in many disciplines is to identify changepoints where the underlying data-generating process for observations Y_x (indexed by a continuous variable X, such as time or a spatial coordinate) alters its central tendency (Aue, Gabrys, Horváth & Kokoszka 2009). A statistically significant jump in $m_\oplus(x)$ at $X = c$ indicates that such a changepoint exists.*

- Neuroscience: Brain Development Trajectories: *Let X be the chronological age of subjects. Y_x could represent the structural brain network of an individual of age x, perhaps derived from diffusion tensor imaging (DTI) and viewed as a graph object. Using a graph metric (e.g., based on the Frobenius norm or spectral graph distances), a jump in the Fréchet mean brain network structure at a critical developmental age c could highlight a significant maturational shift in the topology of typical brain connectivity (Fair et al. 2009, Dubey & Müller 2019).*

Example 2.3 (Threshold Effects in Complex Systems)**.** *More broadly, the framework applies to detecting threshold effects where the central tendency of a metric-space outcome Y changes abruptly when a continuous underlying factor X crosses a critical value c. This generalizes the search for non-linearities and critical transitions.*

- Ecology: Community Composition along Environmental Gradients: *Let X be a continuous environmental gradient (e.g., soil pH, salinity). Y_x is the ecological community composition (e.g., a vector of species abundances, or a set of species presence/absence) at a site with gradient value x. Using a suitable distance metric such as the Aitchison or geodesic distance, a jump in the Fréchet mean community composition at a specific gradient level c would indicate an ecotone or an abrupt shift in the typical community structure (Killick et al. 2012).*

In all these settings, the objects Y_i possess a complex structure. The core statistical question is whether the conditional Fréchet mean path $m_\oplus(x)$, "$E[Y|X=x]$", exhibits a significant jump at $X=c$. The subsequent sections develop the formal tools to test this hypothesis.

2.3 One-Sided Local Fréchet Regression

To estimate the conditional means and variances from the left and right, $m_{\pm,\oplus}, V_{\pm,\oplus}$, I rely on a one-sided version of local Fréchet regression (Petersen & Müller 2019). Assume one observes i.i.d. data $\{(Y_i, X_i)\}_{i=1}^n \sim F$. Then, the one-sided estimators are defined as,

$$\hat{l}_{\pm,\oplus} := \underset{\omega\in\Omega}{\operatorname{argmin}} \left\{ \hat{L}_{\pm,n}(\omega) := n^{-1} \sum_{i=1}^{n} s_{\pm,in}(c, h_m) d^2\left(Y_i, \omega\right) \right\}, \tag{5}$$

where h_m is a bandwidth and $s_{\pm,in}(c,h)$ are standard one-sided local linear regression weights that use data only to the left or right of the cutoff, respectively,

$$s_{+,in}(c,h) = 1(X_i \geq c) \frac{K_h(X_i - c)}{\hat{\sigma}^2_{+,0}} \left[\hat{\mu}_{+,2} - \hat{\mu}_{+,1}(X_i - c)\right],$$

where $K_h(u) = K(u/h)/h$ for a kernel function K, and the components are sample moments calculated using only data to the right of the cutoff,

$$\hat{\mu}_{+,j} = n^{-1} \sum_{k=1}^{n} 1(X_k \geq c) K_h(X_k - c)(X_k - c)^j \quad \text{for } j = 0, 1, 2,$$

$$\hat{\sigma}^2_{+,0} = \hat{\mu}_{+,0}\hat{\mu}_{+,2} - \hat{\mu}^2_{+,1}.$$

The weights for the left-sided estimator, $s_{-,in}(c,h)$, are defined analogously using observations where $X_i < c$ and $X_k < c$.

The corresponding population (or pseudo-true) minimizers are denoted $\tilde{l}_{\pm,\oplus}$ with objective function $\tilde{L}_{\pm,\oplus}(\omega)$. The consistency of these estimators is a straightforward extension of the arguments in Petersen & Müller (2019) and Kurisu et al. (2025).

The sample estimators of the conditional Fréchet variances from above ($V_{+,\oplus}$) and below ($V_{-,\oplus}$) are then simply,

$$\hat{V}_{\pm,\oplus} = \frac{1}{n}\sum_{i=1}^{n} s_{\pm,in}(c,h) d^2(Y_i, \hat{l}_{\pm,\oplus}), \tag{6}$$

where I use a different bandwidth h for the variance than for the mean estimation (h_m) to allow for different degrees of smoothing. The asymptotic variance of $d^2(Y, m_{\pm,\oplus})$ conditional on $X = c$ is denoted $\sigma^2_{\pm,V} := \operatorname{var}\left(d^2(Y, m_{\pm,\oplus}) \mid X = c\right)$, and its sample estimator $\hat{\sigma}^2_{\pm,V}$ is constructed as,

$$\hat{\sigma}^2_{\pm,V} = \frac{1}{n}\sum_{i=1}^{n} s_{\pm,in}(c,h) d^4(Y_i, \hat{l}_{\pm,\oplus}) - \left(\frac{1}{n}\sum_{i=1}^{n} s_{\pm,in}(c,h) d^2(Y_i, \hat{l}_{\pm,\oplus})\right)^2,$$

by the variance shortcut formula. The asymptotic normality of $\hat{V}_{\pm,\oplus}$ and consistency of $\hat{\sigma}^2_{\pm,V}$ are established below. Bandwidth selection is done by a cross-validation procedure that aims to minimize the out-of-sample prediction on both sides of the cutoff simultaneously (see Appendix B for more details).

2.4 Pooled Estimators

To construct the ANOVA-style test, I also need pooled estimators that do not distinguish between observations above and below c. The pooled local Fréchet mean estimator is,

$$\hat{l}_{p,\oplus} := \operatorname*{argmin}_{\omega\in\Omega} \frac{1}{n}\sum_{i=1}^{n} s_{p,in}(c,h) d^2(Y_i, \omega),$$

and the pooled conditional Fréchet variance estimator is,

$$\hat{V}_{p,\oplus} := \frac{1}{n}\sum_{i=1}^{n} s_{p,in}(c,h) d^2(Y_i, \hat{l}_{p,\oplus}),$$

where $s_{p,in}(c,h) = \frac{1}{2}s_{+,in}(c,h) + \frac{1}{2}s_{-,in}(c,h)$. The use of uniform weights $\frac{1}{2}$ assumes that the same bandwidth h is used on both sides, which can easily be relaxed. The corresponding pooled population moments are,

$$m_{p,\oplus} := \underset{\omega \in \Omega}{\operatorname{argmin}} \left\{ M_{p,\oplus}(\omega) := \frac{1}{2}M_{+,\oplus}(\omega) + \frac{1}{2}M_{-,\oplus}(\omega) \right\}$$
$$V_{p,\oplus} := \frac{1}{2}M_{+,\oplus}(m_{p,\oplus}) + \frac{1}{2}M_{-,\oplus}(m_{p,\oplus}),$$

i.e., the pooled population estimators pool across the one-sided objective functions. The consistency of these pooled estimators is proved in Proposition D.5 in the Appendix.

2.5 A Jump Test for Conditional Metric-Space Means

To test for a discontinuity, I propose an ANOVA-style test statistic based on these one-sided and pooled local Fréchet regression estimators. Define auxiliary statistics,

$$F_n = \hat{V}_{p,\oplus} - \left(\frac{1}{2}\hat{V}_{+,\oplus} + \frac{1}{2}\hat{V}_{-,\oplus} \right), \tag{7}$$

$$U_n = \frac{\left(\hat{V}_{+,\oplus} - \hat{V}_{-,\oplus} \right)^2}{(S_+ \hat{\sigma}^2_{+,V} / \hat{f}_X(c)) + (S_- \hat{\sigma}^2_{-,V} / \hat{f}_X(c))}. \tag{8}$$

Here, U_n tests for differences in variances (H_0^{var}). Under the null of equal variances (i.e. $U_n \to 0$), any difference between the pooled and one-sided variances reflected in F_n must be driven by differences in means. Hence, F_n tests for differences in means (H_0^{mean}).

The combined test statistic is,

$$T_n = nh_n U_n + \frac{nh_n F_n^2}{(\hat{\sigma}^2_{+,V} S_+ / \hat{f}_X(c)) + (\hat{\sigma}^2_{-,V} S_- / \hat{f}_X(c))}. \tag{9}$$

Here, the F_n term is squared and scaled by its variance estimate so that it has the same variability as the U_n term.

3 Main Results

3.1 Assumptions

I require the following assumptions for the asymptotic results. Throughout, I assume that the metric space Ω is bounded.

L1 (Sampling). *$\{(Y_i, X_i)\}_{i=1}^n$ are i.i.d. copies of a random element (Y, X) defined on a probability space $(\Xi, \mathcal{F}, P)$. $Y \in \Omega$, $X \in \mathbb{R}$.*

L2 (Densities and Continuity). *The following hold for some $\varepsilon > 0$. (a) The density f_X is twice continuously differentiable on $(c-\varepsilon, c+\varepsilon)$, bounded away from 0 and ∞, and $f_X(c) > 0$. (b) For every $y \in \Omega$ the function $g_y(x) = \partial_x F_{X|Y}(x \mid y)$ is twice continuously differentiable on $(c-\varepsilon, c+\varepsilon)$, and $\sup_{y\in\Omega} \max_{k=0,1,2} \sup_{x\in(c-\varepsilon,c+\varepsilon)} |g_y^{(k)}(x)| < \infty$. (c) For every open set $U \subset \Omega$ with $F_Y(\partial U) = 0$ the map $x \longmapsto \int_U dF_{Y|X=x}(y)$ is twice continuously differentiable on the open intervals $(c-\varepsilon, c)$ and $(c, c+\varepsilon)$.*

For each side separately, this assumption is identical to the one in Petersen & Müller (2019), and generalizes the standard distributional assumptions imposed in local polynomial regression.

L3 (Uniqueness of Fréchet Means). *The Fréchet means $m_{\pm,\oplus}$ and $m_{p,\oplus}$ exist and are unique. For all n, their estimators $\tilde{l}_{\pm,\oplus}$, $\hat{l}_{\pm,\oplus}$, $\tilde{l}_{p,\oplus}$, $\hat{l}_{p,\oplus}$ exist and are unique (the hatted estimators almost surely). Additionally, for any $\varepsilon > 0$: $\inf_{d(\omega,m_{\pm,\oplus})>\varepsilon} \{M_{\pm,\oplus}(\omega) - M_{\pm,\oplus}(m_{\pm,\oplus})\} > 0$, $\liminf_n \inf_{d(\omega,\tilde{l}_{\pm,\oplus})>\varepsilon} \left\{\tilde{L}_{\pm,n}(\omega) - \tilde{L}_{\pm,n}\left(\tilde{l}_{\pm,\oplus}\right)\right\} > 0$, $\liminf_n \inf_{d(\omega,\hat{l}_{\pm,\oplus})>\varepsilon} \left\{\hat{L}_{\pm,n}(\omega) - \hat{L}_{\pm,n}\left(\hat{l}_{\pm,\oplus}\right)\right\} > 0$ (a.s.). Analogous conditions hold for the pooled estimators $\tilde{l}_{p,\oplus}$ and $\hat{l}_{p,\oplus}$ with respect to $M_{p,\oplus}(\omega)$, denoted as Assumption L3-p.*

L4 (Curvature Conditions). *(i) Provided $d(\omega, m_{\pm,\oplus}) < \eta_1$, there exist $\eta_1 > 0, C_1 > 0$ and*

$\beta_1 > 1$ such that,

$$M_{\pm,\oplus}(\omega) - M_{\pm,\oplus}\left(m_{\pm,\oplus}\right) \geq C_1 d\left(\omega, m_{\pm,\oplus}\right)^{\beta_1}.$$

(ii) Provided $d\left(\omega, \tilde{l}_{\pm,\oplus}\right) < \eta_2$, there exists $\eta_2 > 0, C_2 > 0$ and $\beta_2 > 1$ such that:

$$\liminf_n \left[\tilde{L}_{\pm,n}(\omega) - \tilde{L}_{\pm,n}\left(\tilde{l}_{\pm,\oplus}\right)\right] \geq C_2 d\left(\omega, \tilde{l}_{\pm,\oplus}\right)^{\beta_2}.$$

Analogous conditions hold for $M_{p,\oplus}(\omega)$ and $\tilde{L}_{p,n}(\omega)$, denote these as Assumption L4-p.

As mentioned, the uniqueness and curvature assumptions are satisfied for all specific metric spaces considered in this article (Petersen & Müller 2019, Zhou & Müller 2022). More generally, Kimura & Bondell (2025) characterize a class of metric spaces that have unique conditional Fréchet means.

The following are generalizations of standard local polynomial regression assumptions,

K1 (Kernel)**.** *The kernel $K : \mathbb{R} \to [0, \infty)$ is a continuous probability density function, symmetric around zero, with compact support.*

K2 (Bandwidth)**.** *(a) The bandwidth sequence $h = h(n)$, satisfies $h \to 0$, $h_m \to 0$, $nh \to \infty$, $nh_m \to \infty$. (b) $nhh_m^{4/(\beta_1-1)} \to 0$. (c) $h_m(n) = \varrho_m n^{-\theta}$ for some $\varrho_m > 0$, $\theta > 0$, and $h(n) = \varrho n^{-\gamma}$ for some $\varrho > 0$, $\gamma > \theta$, $\gamma > \frac{1}{5}$. (d) $(nh)^{1/2}(nh_m)^{-1/(2(\beta_2-1))} \to 0$, where β_1, β_2 control the curvature of the Fréchet mean estimators in Assumption L4.*

First, the variance bandwidth h must undersmooth the variance estimator ($\gamma > 1/5$) in (c) to eliminate its own asymptotic bias. Second, the variance bandwidth h is assumed to converge faster than the mean bandwidth h_m ($\gamma > \theta$). This baseline relationship ensures that for sufficiently regular problems ($1 < \beta_2 \leq 2$), the stochastic error from the first stage is automatically controlled. The relative rate condition in (b) which handles the bias component of the first-stage error and essentially requires that the mean bandwidth h_m cannot converge too slowly. This ensures the mean estimator $\hat{l}_\oplus$ does not itself need to be undersmoothed—an adaptive property that generalizes the findings of Fan & Yao (1998). Finally, the condition in (d) handles the stochastic component for less regular problems ($\beta_2 > 2$), demanding a more aggressive separation between the bandwidths as the Fréchet mean's objective function becomes less smooth. Note that $\beta_2 \leq 2$ for all metric spaces considered in this article, see Petersen & Müller (2019), Zhou & Müller (2022). These explicit conditions are necessary because, unlike in the Euclidean case, there are no algebraic cancellations to automatically suppress the first-stage error in a general metric space.

Finally, let $N(\epsilon, S, d)$ be the ϵ-covering number of a set $S \subseteq \Omega$. The entropy integral for a δ-ball $B_\delta(\omega_0)$ around $\omega_0 \in \Omega$ is $J(\delta, \omega_0) := \int_0^1 \sqrt{1 + \log N\left(\delta\epsilon, B_\delta(\omega_0), d\right)} \mathrm{d}\epsilon$.

T1 (Local Entropy Condition)**.** *For any $\omega_0 \in \{m_{-,\oplus}, m_{+,\oplus}, m_{p,\oplus}\}$, $\delta J(\delta, \omega_0) \to 0$ as $\delta \to 0$.*

T2 (Global Entropy Condition)**.** *The entropy integral of Ω is finite: $\int_0^1 \sqrt{1 + \log N(\epsilon, \Omega, d)} \mathrm{d}\epsilon < \infty$. This is a stronger condition, used for uniform consistency results (e.g., Theorem 3.6).*

Note that for most results in this article, I only need the weaker local entropy condition T1, which is also weaker than its equivalent in Dubey & Müller (2019) since the latter imposes it for all $\omega \in \Omega$. The stronger global entropy condition T2 is only required for the power results.

3.2 Central Limit Theorem for Conditional Fréchet Variance

First, I establish the asymptotic distribution of the estimated conditional Fréchet variances $\hat{V}_{\pm,\oplus}$. Denote convergence in distribution with $\xrightarrow{d}$.

Theorem 3.1 (CLT for Conditional Fréchet Variance)**.** *Let Assumptions L1-T1 hold. Let $\hat{f}_X(c)$ be a $\sqrt{nh}$-consistent estimator for $f_X(c)$. Then,*

$$\sqrt{nh}\left(\hat{V}_{\pm,\oplus} - V_{\pm,\oplus}\right) / \left(\sqrt{\frac{S_\pm}{\hat{f}_X(c)}} \hat{\sigma}_{\pm,V}\right) \xrightarrow{d} N\left(0, 1\right),$$

where

$$S_{\pm} := \frac{\int_0^{\infty}(K_{\pm,12} - uK_{\pm,11})^2 K^2(u)\,\mathrm{d}u}{(K_{\pm,12}K_{\pm,10} - K_{\pm,11}^2)^2}.$$

The proof, detailed in Appendix D.2, follows by decomposing the estimation error into a Bahadur-style representation. The leading stochastic term is shown to be asymptotically normal via a Lyapunov CLT, while the plug-in and bias terms are asymptotically negligible under our bandwidth assumptions. The key intuition behind why one can establish a classical CLT is that all the geometry of the complex objects is absorbed into the squared distance $d^2(Y_i, m_{\pm,\oplus})$. Further, because the metric space (Ω, d) is assumed to be bounded, those distances possess uniformly bounded moments of every order, delivering the necessary control over the asymptotic distribution.

3.3 Asymptotic Distribution of the Test

The following results determine the asymptotics of the test statistics in Eqs. (7) and (8) under the respective null hypotheses,

Proposition 3.2. *Under $H_0^{mean} : m_{+,\oplus} = m_{-,\oplus}$ and Assumptions L1-T1 (including L3-p, L4-p for pooled estimators),*

$$\sqrt{nh_n}F_n = o_p(1).$$

Proposition 3.3. *Under $H_0^{var} : V_{+,\oplus} = V_{-,\oplus}$, and Assumptions L1-T1,*

$$nh_nU_n \xrightarrow{d} \chi_1^2,$$

where χ_1^2 is the chi-squared distribution with 1 degree of freedom.

Proofs are in Appendix D.2. The result for U_n follows straightforwardly from the central limit theorem derived in the previous section. The result for F_n is intuitive: under the null, the variance estimators converge to the same quantity, and their difference converges faster to 0 than each separate term does.

For the combined test statistic in (9), the following result obtains.

Corollary 3.4. *Under $H_0 : m_{+,\oplus} = m_{-,\oplus}$ and $V_{+,\oplus} = V_{-,\oplus}$, and Assumptions L1-T1 (including -p versions),*

$$T_n \to_d \chi^2_1.$$

This follows because, under the joint null, the contribution of the F_n term vanishes and hence the asymptotic distribution of the U_n term dominates. Based on this corollary, the rejection region for a test of size α is $R_{n,\alpha} = \{T_n > \chi^2_{1,\alpha}\}$. Since this is an asymptotic result, one may worry that the F_n term can lead to size distortions in finite sample. While the simulations in Section 4 do reveal slight overcoverage in very small samples, the test's acceptance probability rapidly converges to the nominal level.

3.4 Consistency of the Test

To study the power of the test, define the population counterparts,

$$F_{pop} = V_{p,\oplus} - \left(\frac{1}{2}V_{+,\oplus} + \frac{1}{2}V_{-,\oplus}\right), \quad U_{pop} = \frac{(V_{+,\oplus} - V_{-,\oplus})^2}{\left(S_+\sigma^2_{+,V}/f_X(c)\right) + \left(S_-\sigma^2_{-,V}/f_X(c)\right)}.$$

Proposition 3.5. *Under Assumptions L1-T1 (including -p versions) and T2, as $n \to \infty$, $|F_n - F_{pop}| = o_p(1)$. $F_{pop} \geq 0$, and $F_{pop} = 0$ if and only if $m_{p,\oplus} = m_{+,\oplus} = m_{-,\oplus}$. Also, $|U_n - U_{pop}| = o_p(1)$, and $U_{pop} \geq 0$, with $U_{pop} = 0$ if and only if $V_{+,\oplus} = V_{-,\oplus}$.*

This proposition provides the crucial link between the test statistics and the hypotheses, showing that F_n and U_n are consistent for population measures of mean and variance differences, respectively. These population measures are zero if and only if no discontinuity exists. This equivalence is immediate for U_n, while for F_n, it follows from the uniqueness of the Fréchet means.

Finally, consider sequences of alternatives $H_A^{(n)}$ where $F_{pop} \geq a_n$ or $U_{pop} \geq b_n$ for sequences $a_n, b_n \geq 0$ where either $a_n > 0$ or $b_n > 0$. The power function is $\beta_{H_A^{(n)}} =$

$\inf_{(F_{pop},U_{pop})\in H_A^{(n)}} P(T_n > \chi^2_{1,\alpha})$.

Theorem 3.6 (Consistency of the Test)**.** *Under Assumptions L1-T2 (including -p versions), for sequences of alternatives $H_A^{(n)}$ and any $\alpha > 0$,*

(a) If $F_{pop} \geq a_n$ with $(nh_n)^{1/2}a_n \to \infty$, then $\beta_{H_A^{(n)}} \to 1$.

(b) If $U_{pop} \geq b_n$ with $nh_nb_n \to \infty$, then $\beta_{H_A^{(n)}} \to 1$.

This shows the test will pick up any jump whose size does not shrink faster than the sampling noise inside the bandwidth window, i.e., the test is consistent against contiguous alternatives.

4 Simulations

To evaluate the test's finite-sample performance, I carry out Monte Carlo simulations in three metric spaces: probability densities, covariance matrices, and graph Laplacians of networks. I consider both piecewise-smooth and piecewise-constant data generating processes (DGPs) to assess the usefulness of the nonparametric regression approach (see Appendix C for full details). Though there are no general metric-space tests that directly compete with the proposed one, I adapt the k-sample test of Dubey & Müller (2019) to compare the two means in a shrinking window around the cutoff to illustrate the value of the nonparametric regression approach. Since this is essentially equivalent to a local constant regression approach, we should expect it to do well on the piecewise-constant signal but overreject on the piecewise-smooth one.

The results confirm the asymptotic theory. Table 1 shows that the proposed test exhibits excellent size control, with empirical rejection rates under the null quickly converging to the nominal 5% level as the sample size increases. The test also demonstrates robust power across all settings. Figure 1 highlights the key advantage of the regression-based approach.

Metric Space	N	Size	Power
Covariance	200	0.031	0.997
	500	0.046	0.997
	1000	0.048	1.000
Density	200	0.038	0.999
	500	0.043	1.000
	1000	0.043	1.000
Network	200	0.021	1.000
	500	0.062	1.000
	1000	0.055	1.000

Table 1: Size and Power by Sample Size and Metric Space

Note: Empirical rejection rates of the proposed Fréchet jump test from 1000 Monte Carlo simulations at a nominal $\alpha = 0.05$ level. 'Size' is the rejection rate under the null hypothesis of no jump (H_0). 'Power' is the rejection rate under a fixed alternative (H_1) with a jump introduced at $c = 0.5$. The Data Generating Process is piecewise-smooth in all cases. The fixed jump magnitudes used to calculate 'Power' are: $\delta_D = 1.5$ for Densities, $\beta_C = 1.5$ for Covariance matrices, and $\delta_N = 0.25$ for Networks.

While both tests perform well under a simple piecewise-constant signal (right panels) – with the k-sample test expectedly converging faster due to its parametric rate –, the localized k-sample test severely over-rejects the null under the piecewise-smooth DGPs (left panels). My proposed test, by contrast, maintains correct size while reliably detecting true jumps, confirming its utility for a broader and more realistic class of problems.

5 Empirical Illustrations

5.1 Non-Compete Agreements and the Composition of Work from Home

Question. Here, I apply the test to investigate how the enforceability of non-compete agreements (NCAs) affects the composition of work-from-home (WFH) arrangements. In

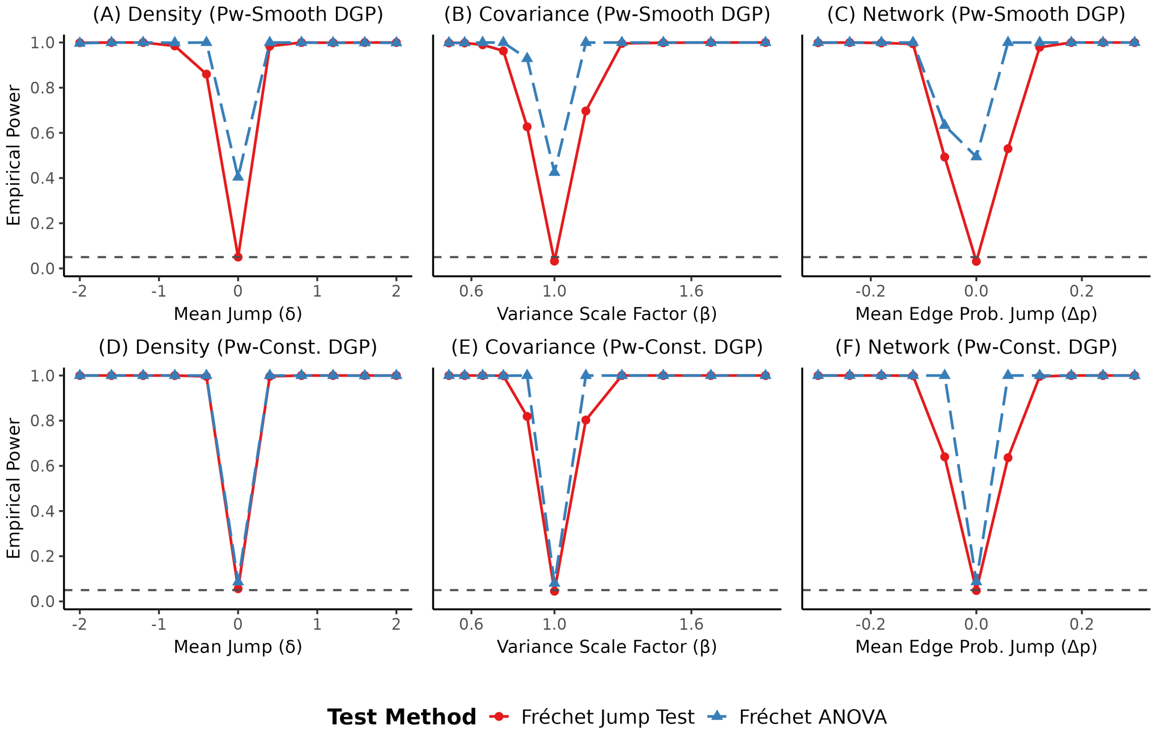

Figure 1: Power Curves

Note: Figure shows empirical power curves comparing the proposed Fréchet jump test and a localized adaptation of the two-sample test from Dubey & Müller (2019). x-axis represents the jump magnitude: additive mean jump for densities (δ_D), multiplicative scale factor for covariance matrices (β_C), and additive edge probability jump for networks (δ_N). Columns correspond to the metric space; rows to the DGP (Piecewise-Smooth vs. Piecewise-Constant). Results based on 1000 simulations with $N = 200$. Dotted horizontal line indicates the nominal test level of $\alpha = 0.05$.

2020, Washington state enacted a law (RCW 49.62) that makes NCAs enforceable only for employees whose annual earnings exceed a specific, inflation-adjusted threshold of approximately \$100,000. This policy creates a sharp discontinuity where non-compete agreements are unenforceable below the income threshold but become enforceable when an employee's salary crosses it. Unlike a traditional RD design, however, I cannot rule out manipulation around the threshold. Indeed, the object of interest is precisely how firms and workers strategically adjust employment terms when the legal landscape of post-employment mobility changes.

My central hypothesis is that WFH arrangements become a key bargaining margin, particularly during the sample period which covered the COVID-19 pandemic years when WFH became ubiquitous. When an employee's salary crosses the threshold, rendering an NCA enforceable, their ability to switch to a local competitor is curtailed. One might expect a worker to respond by seeking employment with an out-of-state firm to escape the policy's reach. However, my results below do not reveal any significant jump in the probability of working for an out-of-state employer at the threshold. This suggests that adjustments are more likely to occur within the current employment relationship. In particular, faced with a binding NCA, a high-earning employee may demand non-monetary compensation for their reduced future mobility. From the firm's perspective, granting greater WFH flexibility could be a valuable, low-cost concession to attract these employees. Therefore, I test for an abrupt shift in the average WFH composition at the earnings threshold for newly hired employees.

Previous literature has found that only 10% of employees bargain over non-compete agreements (Starr et al. 2021), though this was estimated prior to the spread of non-compete laws such as Washington's, and likely increases for higher-income employees like the ones in my sample. Lipsitz & Starr (2022) found that non-competes negatively affect workers by reducing their wages and job-to-job mobility, creating cause for bargaining over the agreement. Further, the emerging literature on the effects of remote work have documented WFH's role as a job amenity that is substitutable for higher income and leads to less attrition (Mas & Pallais 2017, Barrero et al. 2022, Bloom et al. 2022, Cullen et al. 2025).

Data and Empirical Approach. I use data from the US Census Bureau's Survey of Income and Program Participation (SIPP) for the years 2020 through 2023 to construct a panel of individual job spells for employees in Washington state. The running variable is the employee's annual salary, centered around the relevant year-specific legal threshold.

The outcome of interest is the composition of an employee's WFH schedule, defined as a 6-dimensional vector representing the average proportion of time during the job spell that the employee worked 0 to 5 days from home. Since the components of this compositional vector are non-negative and sum to one, I formally test for a discontinuity by applying a square-root transformation so the vector lies on the positive orthant of a unit 5-sphere ($\mathbb{S}^5 \cap \mathbb{R}^6_+$) and adopting the geodesic metric, $d_g(y_1, y_2) = \arccos\left(y_1^\top y_2\right)$ (Bhattacharjee & Müller 2023, Supp. S.4.2).

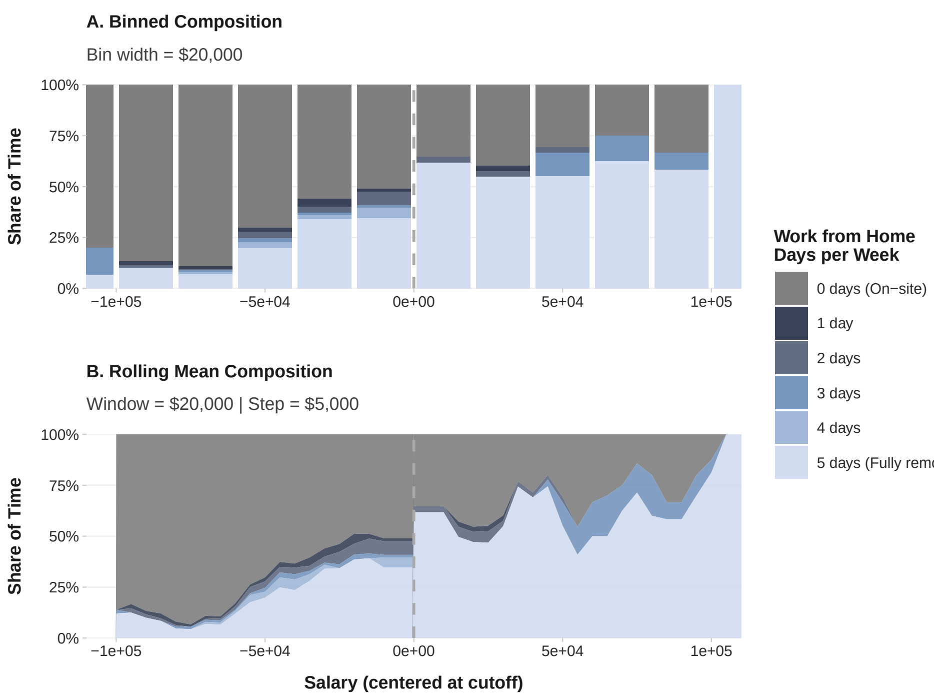

Figure 2: Work from Home Composition by Income: Binned and Smoothed

Note: figures show 100% stacked bar chart and 100% stacked area chart indicating percentage of time spent working from home 0–5 days a week throughout a given employment spell (y axis), by annual salary of the employee (x axis). Annual salary is centered around the Washington non-compete enforceability cutoff in each respective year (2020-2023), indicated by dashed vertical line. Darker colors indicate fewer days worked from home.

Results. Figure 2 depicts the evolution of employees' WFH composition during their job spell (y axis) by their annual salary (x axis), using both a binned stacked bar chart and one smoothed in a rolling window. The annual salary variable was centered around the enforceability cutoff of approximately \$100K annual income. In both panels of the figure, one can see a stark jump in the share of days worked fully remotely (light blue color) once the non-compete agreements become enforceable, while the share of time worked fully in person or partly at home diminishes. This suggests that employees, when faced with the possibility of a non-compete, used work from home as a bargaining margin during the pandemic.

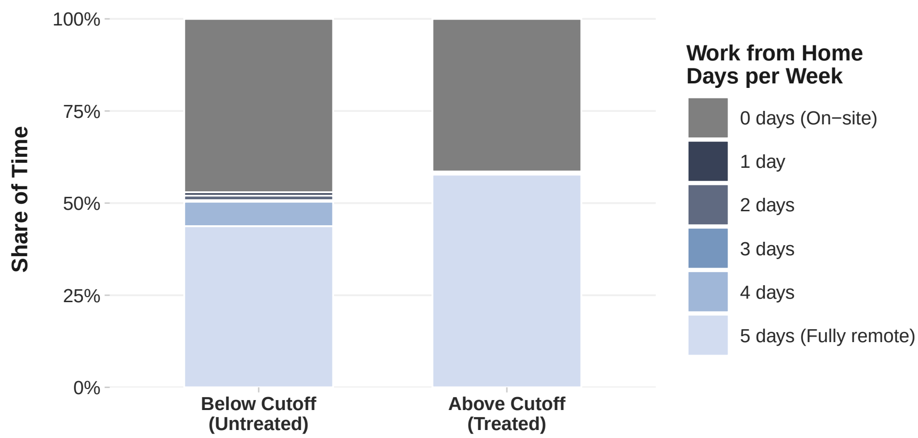

Figure 3: Non-Compete Enforceability and Work-From Home Job Composition: Fréchet Mean Estimates

Note: figure shows estimates of Fréchet mean of employees' WFH composition to the left and right of the non-compete enforceability cutoff in Washington state (2020-2023). Darker colors indicate fewer days worked from home.

These descriptive results are echoed in the local Fréchet mean estimates to the left and right of the cutoff in Figure 3. The mean WFH composition below the cutoff has approximately 15% less time spent working fully remotely than the composition above the cutoff. Partly remote work vanishes completely, while fully in-person work diminishes by approximately 5%. The test for this difference between the two composition vectors has a p-value of 0.059, significant at the 10% level. By contrast, Table E-1 reports the coefficients of a classical local polynomial regression on each of the components of the vector separately (share of time worked 0–5 days at home, as well as whether any or all time was spent working from home). None of these coefficients are significant at the 10% level, illustrating the additional power the test can provide by exploiting the information in the compositional structure of the data. As mentioned, I do not find evidence that employees above the cutoff were more likely to work or move out of state. Furthermore, although one may expect employees to respond strategically to the cut-off point, the RD estimates do not suggest any covariate imbalance in age, gender, or education of the employees. These findings indicate that the enforceability of non-competes induced a structural shift in WFH arrangements, which served as a key bargaining margin for affected employees during the pandemic.

5.2 Preferential Tariff Loss and Countries' Input-Output Networks

Question. Next, I investigate how national economies restructure in response to losing preferential trade status with the US. To that end, I exploit a unique policy event following the reauthorization of the US Generalized System of Preferences (GSP) on June 29, 2015. The GSP program, which grants duty-free access for thousands of products, had been inactive for nearly two years prior to this date. The 2015 reboot created a clear, forward-looking timeline for firms and governments: a country's GSP eligibility for the year 2017 would be determined by its Gross National Income (GNI) per capita from the year 2015.

This setup creates a regression discontinuity design for studying the effect of preferential tariffs on a country's production network. The running variable is a country's 2015 GNI per capita. The treatment is the loss of GSP benefits on January 1, 2017, for countries whose 2015 GNI crossed the high-income threshold. The question is whether this negative trade shock induced a structural change in a country's national input-output (I-O) network as export demand dropped and firms adapted their supply chains.

A potential concern is that the World Bank's high-income threshold might trigger other policies concurrently with the loss of GSP benefits. However, the most significant of these—the loss of Official Development Assistance (ODA) and graduation from the World Bank's IBRD lending—operate on longer timelines. ODA eligibility ends only after a country remains high-income for three consecutive years (Organisation for Economic Co-operation and Development 2023), while IBRD graduation is a gradual, case-by-case process that unfolds over several years (Heckelman et al. 2011). In contrast, US GSP status is revoked without exception on a sharp two-year trigger (Jones 2017). This ensures that at the point of GSP expiration, other major policy changes linked to the income threshold have not yet occurred, isolating the GSP effect.

The international trade literature has documented extensive direct and indirect links between foreign exports, tariffs, and domestic production networks (Baqaee & Farhi 2024, Dhyne et al. 2021, Amiti & Konings 2007). Specifically for the GSP, Özden & Reinhardt (2005) found that GSP graduation correlates with affected countries implementing more liberal trade policies, while Hakobyan (2020) found that GSP expiration led to sizeable declines in exports to the US of around 3%. However, to my knowledge, there exist no empirical methods to study the impact of trade shocks directly on the full structure of production networks. This underscores the usefulness of the proposed test for empirically validating predictions from trade network models and evaluating the impact of trade shocks.

Data and Empirical Approach. The running variable, GNI per capita, is sourced from the World Bank Development Indicators (WDI), with the cutoff set at the official 2015 high-income threshold. The primary outcomes are domestic input-output (I-O) networks for 2016 (capturing anticipatory effects) and 2017 (capturing implementation effects), derived from the EORA26 global database (Lenzen et al. 2012). The analysis is constrained to these years because the GSP program lapsed again at the end of 2017, and the subsequent period is further confounded by the COVID-19 pandemic. For each of the 140 countries in the final dataset, I extract the domestic I-O table from the EORA26 database (Lenzen et al. 2012) and aggregate its 26 sectors into 8 broader categories (see Table E-2). The outcome objects for the test are the Leontief Inverse (or total requirements) matrices derived from these 8×8 I-O matrices. A Leontief Inverse captures the total (direct and indirect) output from each sector required to satisfy one unit of final demand, thus representing the economy's full production multipliers.

To perform the test, I work on the space of graph Laplacians of these production networks, equipped with the Frobenius norm, which satisfies the necessary metric-space assumptions (Zhou & Müller 2022). For the main analysis, I symmetrize the Laplacians, though results are highly similar when using in-degree or out-degree Laplacians. Critically, because the population Fréchet mean of graph Laplacians under the Frobenius norm is the entrywise average of their matrix elements (Zhou & Müller 2022), the test directly captures changes in the country-average production multiplier network.

Results. The test indicates a statistically significant jump in the conditional Fréchet mean of the Leontief Inverse matrices at the GSP graduation threshold ($p = 0.028$). This suggests that countries losing GSP access experience a structural change in their domestic production networks. Moreover, the test also shows a significant change in the average Leontief Inverse matrices one year before GSP graduation ($p = 0.016$), further suggesting that countries already began to adjust their production networks in anticipation of the policy change. Crucially, a placebo test for 2015—a year when the GSP program was inactive—reveals no statistically significant effect ($p = 0.290$), reinforcing the conclusion that the detected changes are indeed attributable to the GSP graduation event and not to other confounding policies.

Looking at the jump estimates, Figure 4 plots the elementwise difference between the treated and control Fréchet means of the 8×8 Leontief Inverse matrices in 2017 (mapped back from the graph Laplacians). Warmer colors indicate stronger cross-sector multipliers in the treated group while cooler colors indicate weaker ones. The most striking changes appear in the resource-export chain. The effect magnitudes represent the change in total input required per dollar of final demand. The Mining-Manufacturing flows fall by 0.47 and the Mining-Trade ones by 0.36 on average, while Manufacturing's spill-overs into Construction and Trade decline by 0.26 and 0.21 respectively. These contractions indicate that, once preferential tariffs disappear, raw materials are no longer routed through domestic factories and wholesale networks for export processing.

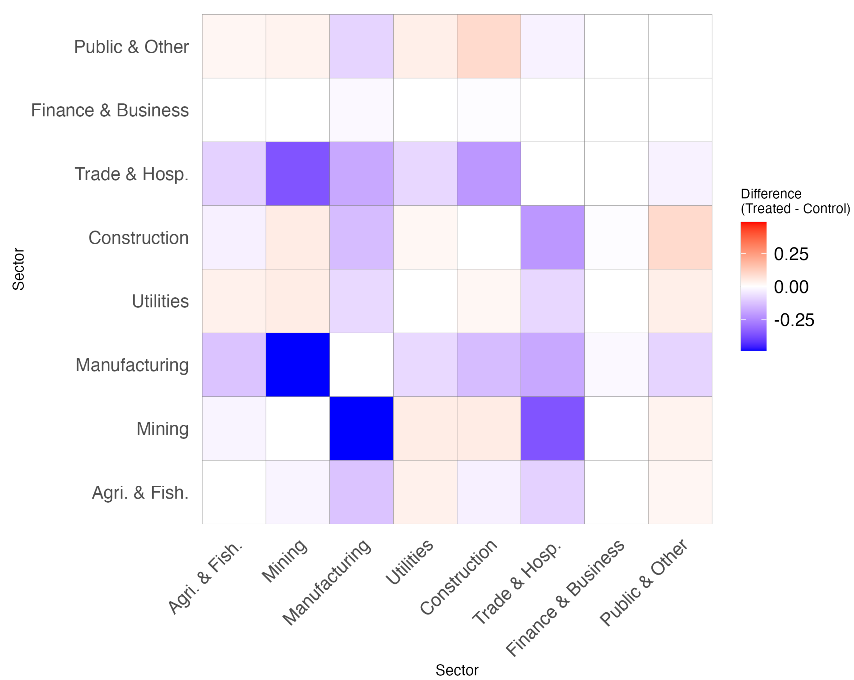

Figure 4: Jump in Input-Output Networks After Preferential Tariff Loss

Note: figure shows estimates of difference in Fréchet means of countries' Leontief Inverse matrices to the left and right of the US GSP eligibility cutoff.

A different pattern emerges on the domestic side of the economy. Construction's linkage with

Public and Other Services rises by 0.11, and Mining's ties with both Construction (0.06) and Utilities (0.04) also increases slightly. The Utilities industry itself deepens its connections with Construction and the public sector (about 0.03 each). Together, these increases suggest that extraction industries redirect output towards home-market infrastructure and energy projects, while construction firms rely more heavily on government demand, likely to offset the export shortfall.

Table E-4 lists the five largest absolute decreases and increases. All major contractions involve Mining or Manufacturing on at least one end, whereas four of the five strengthened links involve Construction or Utilities. The overall heatmap in Figure 4 is light blue along the main diagonal, implying a mild reduction in own-sector multipliers as well. Taken together, these patterns point to a sparser and less tightly coupled production network in economies that lose GSP access—an adjustment that would be invisible in most scalar outcomes such as total trade or sectoral output alone. Indeed, in Table E-3, I report the results of a standard RD analysis on scalar network outcomes such as centrality, complexity, and multiplier effects. None of these show significant changes at the GSP graduation threshold, highlighting the value of the proposed Fréchet jump test for capturing network-wide restructuring that would otherwise be missed.

6 Conclusion

This article has introduced a novel test for detecting discontinuities in conditional Fréchet means within general metric spaces. By leveraging local Fréchet regression techniques, it develops test statistics that compare the conditional Fréchet variances to the left and right of a potential jump location. For that, it establishes a central limit theorem for the conditional Fréchet variance, derived the asymptotic null distribution of the proposed test statistic, and demonstrated its consistency against relevant alternatives. Two empirical applications illustrate the utility of the test by providing novel evidence for the effect of non-competes on work from home job compositions and of preferential tariffs on developing countries' production networks.

The proposed framework is general, making it applicable to a wide range of data types where outcomes are elements of a metric space, such as distributions, networks, or spheres, that evolve nonparametrically in a scalar covariate. The test provides a formal tool for investigating abrupt changes in the central tendency of such complex data. This can be valuable in many fields, including economics, biology, and image analysis, where identifying structural breaks or treatment effects in non-Euclidean data is of increasing importance. Future research could explore finite sample performance, optimal bandwidth selection, and extensions to multivariate conditioning variables or other forms of Fréchet regression.

Data Availability

Replication files for the paper, including all code and data to reproduce the results are hosted at https://github.com/Davidvandijcke/frechesTest_replication.

References

Amiti, M. & Konings, J. (2007), 'Trade liberalization, intermediate inputs, and productivity: Evidence from indonesia', *American Economic Review* **97**(5), 1611–1638.
URL: *https://doi.org/10.1257/aer.97.5.1611*

Aue, A., Gabrys, R., Horváth, L. & Kokoszka, P. (2009), 'Break detection in the covariance structure of multivariate time series models', *The Annals of Statistics* **41**(2), 944–978.

Aue, A., Hörmann, S., Horváth, L. & Reimherr, M. (2009), 'Break detection in the covariance structure of multivariate time series models', *The Annals of Statistics* **37**(6B), 4046–4087. **URL:** *https://doi.org/10.1214/09-AOS707*

Avanesov, V. & Buzun, N. (2018), 'Change-point detection in high-dimensional covariance structure', *Electronic Journal of Statistics* .

Baqaee, D. R. & Farhi, E. (2024), 'Networks, barriers, and trade', *Econometrica* **92**(2), 505–541.

Barrero, J. M., Bloom, N., Davis, S. J., Meyer, B. H. & Mihaylov, E. (2022), The shift to remote work lessens wage-growth pressures, Technical report, National Bureau of Economic Research.

Berkes, I., Gabrys, R., Horváth, L. & Kokoszka, P. (2009), 'Detecting changes in the mean of functional observations', *Journal of the Royal Statistical Society Series B: Statistical Methodology* **71**(5), 927–946.

Bhattacharjee, M., Banerjee, M. & Michailidis, G. (2020), 'Change point estimation in a dynamic stochastic block model', *Journal of machine learning research* **21**(107), 1–59.

Bhattacharjee, S. & Müller, H.-G. (2023), 'Single index fréchet regression', *The Annals of Statistics* **51**(4), 1770–1798.

Billingsley, P. (1995), *Probability and measure*, John Wiley & Sons.

Bloom, N., Han, R. & Liang, J. (2022), How hybrid working from home works out, Technical report, National Bureau of economic research.

Calonico, S., Cattaneo, M. D. & Titiunik, R. (2014), 'Robust nonparametric confidence intervals for regression-discontinuity designs', *Econometrica* **82**(6), 2295–2326.

Chen, H. & Friedman, J. H. (2017), 'A new graph-based two-sample test for multivariate and object data', *Journal of the American statistical association* **112**(517), 397–409.

Chen, Y. & Müller, H.-G. (2022), ‘Uniform convergence of local fréchet regression with applications to locating extrema and time warping for metric space valued trajectories’, *The Annals of Statistics* **50**(3), 1573–1592.

Cho, S., Shin, M., Cho, Y. H. & Lim, J. (2025), ‘Change point detection in high dimensional covariance matrix using pillai’s statistics’, *AStA Advances in Statistical Analysis* **109**(1), 53–84.

Chow, G. C. (1960), ‘Tests of equality between sets of coefficients in two linear regressions’, *Econometrica: Journal of the Econometric Society* pp. 591–605.

Cullen, Z., Pakzad-Hurson, B. & Perez-Truglia, R. (2025), Home sweet home: How much do employees value remote work?, *in* ‘AEA Papers and Proceedings’, Vol. 115, American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203, pp. 276–281.

Dette, H., Pan, G. & Yang, Q. (2022), ‘Estimating a change point in a sequence of very high-dimensional covariance matrices’, *Journal of the American Statistical Association* **117**(537), 444–454.

Dhyne, E., Kikkawa, A. K., Mogstad, M. & Tintelnot, F. (2021), ‘Trade and domestic production networks’, *Review of Economic Studies* **88**(2), 643–668.
URL: *https://doi.org/10.1093/restud/rdaa062*

Dörnemann, N. & Dette, H. (2024), ‘Detecting change points of covariance matrices in high dimensions’, *arXiv preprint arXiv:2409.15588* .

Dubey, P. & Müller, H.-G. (2019), ‘Fréchet analysis of variance for random objects’, *Biometrika* **106**(4), 803–821.

Dubey, P. & Müller, H.-G. (2020), ‘Fréchet change-point detection’, *The Annals of Statistics* **48**(6), 3312–3335.

Dubey, P. & Zheng, M. (2023), ‘Change point detection for random objects using distance profiles’, *arXiv preprint arXiv:2311.16025* .

Enikeeva, F. & Klopp, O. (2025), ‘Change-point detection in dynamic networks with missing links’, *Operations Research* .

Fair, D. A., Cohen, A. L., Power, J. D., Dosenbach, N. U., Church, J. A., Miezin, F. M., Schlaggar, B. L. & Petersen, S. E. (2009), ‘Functional brain networks develop from a “local to distributed” organization’, *PLoS computational biology* **5**(5), e1000381.

Fan, J. & Gijbels, I. (1992), ‘Variable bandwidth and local linear regression smoothers’, *The Annals of Statistics* pp. 2008–2036.

Fan, J. & Gijbels, I. (1996), *Local Polynomial Modelling and Its Applications*, Chapman & Hall, London.

Fan, J. & Yao, Q. (1998), ‘Efficient estimation of conditional variance functions in stochastic regression’, *Biometrika* **85**(3), 645–660.

Fréchet, M. (1948), Les éléments aléatoires de nature quelconque dans un espace distancié, *in* ‘Annales de l’institut Henri Poincaré’, Vol. 10, pp. 215–310.

Hahn, J., Todd, P. & Van der Klaauw, W. (2001), ‘Identification and estimation of treatment effects with a regression-discontinuity design’, *Econometrica* **69**(1), 201–209.

Hakobyan, S. (2020), ‘Gsp expiration and declining exports from developing countries’, *Canadian Journal of Economics/Revue canadienne d’économique* **53**(3), 1132–1161.

Heckelman, J. C., Knack, S. & Rogers, F. H. (2011), Crossing the threshold: An analysis of ibrd graduation policy, Policy Research Working Paper 5531, The World Bank.
URL: *https://openknowledge.worldbank.org/handle/10986/3349*

Imbens, G. W. & Lemieux, T. (2008), ‘Regression discontinuity designs: A guide to practice’, *Journal of econometrics* **142**(2), 615–635.

Jiang, F., Zhu, C. & Shao, X. (2024), ‘Two-sample and change-point inference for non-euclidean valued time series’, *Electronic Journal of Statistics* **18**(1), 848–894.

Jones, V. C. (2017), Generalized system of preferences (gsp): Overview and issues for congress, Technical Report RL33663, Congressional Research Service.
URL: *https://crsreports.congress.gov/product/pdf/RL/RL33663/16*

Kei, Y. L., Li, J., Li, H., Chen, Y. & Padilla, O. H. M. (2024), ‘Change point detection in dynamic graphs with generative model’, *arXiv preprint arXiv:2404.04719* .

Killick, R., Fearnhead, P. & Eckley, I. A. (2012), ‘Optimal detection of changepoints with a linear computational cost’, *Journal of the American Statistical Association* **107**(500), 1590–1598.

Kimura, M. & Bondell, H. (2025), ‘Theoretical and Practical Analysis of Fréchet Regression via Comparison Geometry’, *arXiv preprint arXiv:2502.01995* .

Kruskal, W. H. & Wallis, W. A. (1952), ‘Use of ranks in one-criterion variance analysis’, *Journal of the American statistical Association* **47**(260), 583–621.

Kurisu, D. & Otsu, T. (2024), *Empirical Likelihood for Manifolds*, LSE, STICERD.

Kurisu, D., Zhou, Y., Otsu, T. & Müller, H.-G. (2025), ‘Regression discontinuity designs for functional data and random objects in geodesic spaces’, *arXiv preprint arXiv:2506.18136* .

Lenzen, M., Kanemoto, K., Moran, D. & Geschke, A. (2012), ‘Mapping the structure of the world economy’, *Environmental science & technology* **46**(15), 8374–8381.

Lipsitz, M. & Starr, E. (2022), 'Low-wage workers and the enforceability of noncompete agreements', *Management Science* **68**(1), 143–170.

Madrid Padilla, C. M., Xu, H., Wang, D., MADRID Padilla, O. H. & Yu, Y. (2023), 'Change point detection and inference in multivariate non-parametric models under mixing conditions', *Advances in Neural Information Processing Systems* **36**, 21081–21134.

Mas, A. & Pallais, A. (2017), 'Valuing alternative work arrangements', *American Economic Review* **107**(12), 3722–3759.

Müller, H.-G. (1992), 'Change-points in nonparametric regression analysis', *The Annals of Statistics* pp. 737–761.

Müller, R., Schuhmacher, D. & Mateu, J. (2024), 'Anova for metric spaces, with applications to spatial data', *Statistical Science* **39**(2), 262–285.

Organisation for Economic Co-operation and Development (2023), 'When and why do countries stop being eligible for receiving official development assistance?', OECD Development Co-operation Directorate Blog Post.

Özden, Ç. & Reinhardt, E. (2005), 'The perversity of preferences: Gsp and developing country trade policies, 1976–2000', *Journal of Development Economics* **78**(1), 1–21.

Page, E. (1957), 'On problems in which a change in a parameter occurs at an unknown point', *Biometrika* **44**(1/2), 248–252.

Penaloza, E. & Stevens, N. (2024), 'Changepoint detection in highly-attributed dynamic graphs', *arXiv preprint arXiv:2407.06998* .

Petersen, A. & Müller, H.-G. (2019), 'Fréchet regression for random objects with euclidean predictors', *The Annals of Statistics* **47**(2), 691–719.

Ryan, S. & Killick, R. (2023), ‘Detecting changes in covariance via random matrix theory’, *Technometrics* **65**(4), 480–491.

Starr, E. P., Prescott, J. J. & Bishara, N. D. (2021), ‘Noncompete agreements in the US labor force’, *The Journal of Law and Economics* **64**(1), 53–84.

Székely, G. J. & Rizzo, M. L. (2013), ‘Energy statistics: A class of statistics based on distances’, *Journal of statistical planning and inference* **143**(8), 1249–1272.

Thistlethwaite, D. L. & Campbell, D. T. (1960), ‘Regression-discontinuity analysis: An alternative to the ex post facto experiment.’, *Journal of Educational psychology* **51**(6), 309.

van der Vaart, A. W. & Wellner, J. A. (1996), *Weak Convergence and Empirical Processes: With Applications to Statistics*, Springer, New York.

Van Dijcke, D. (2025), ‘Regression discontinuity design with distribution-valued outcomes’, *arXiv preprint arXiv:2504.03992* .

Wang, X., Zhu, J., Pan, W., Zhu, J. & Zhang, H. (2024), ‘Nonparametric statistical inference via metric distribution function in metric spaces’, *Journal of the American Statistical Association* **119**(548), 2772–2784.

Zhang, J., Zhang, S., Jones, C. W., Basner, M. & Shou, H. (2025), ‘A generalized fréchet test for object data with unequal repeated measurements’, *arXiv preprint arXiv:2503.01514* .

Zhang, Y., Zhu, C. & Shao, X. (2025), ‘Change-point detection for object-valued time series’, *Journal of Business & Economic Statistics* (just-accepted), 1–23.

Zhou, Y. & Müller, H.-G. (2022), ‘Network regression with graph laplacians’, *Journal of Machine Learning Research* **23**(320), 1–41.