Two sample inference in functional linear models A simulation study

Lajos Horváth* Piotr Kokoszka† Matthew Reimherr‡

For details regarding the discussed testing procedure please see (Horváth, Kokoszka, Reimherr. *Two sample inference in functional linear models*). To illustrate the performance of our test in a generic setting, we consider a fully functional linear model with integral kernels of the form

$$\psi(s,t) = c \min\{s,t\} \qquad \psi^*(s,t) = c^* \min\{s,t\},$$

where c and c^* are constants. We set $N = M = 100$, and use 5 FPC's for the regressors variables, and 3 FPC's for the responses.

We use standard Brownian motions as error terms, and consider the regressors of the following four types:

(A) Standard Brownian motions in both samples (Gaussian processes, equal covariances, Table 1).

(B) For the first sample the explanatory functions are standard Brownian motions and for the second sample they are Brownian bridges (Gaussian processes, different covariances, Table 2).

(C) For both sets of explanatory functions we use

$$X(t) = n^{-1/2} \sum_{k=1}^{\lfloor nt \rfloor} \frac{T_i}{\sqrt{\mathrm{var}(T_i)}},$$

where $\{T_i\}$ are iid t-distributed random variables with 6 degrees of freedom and $n = 200$ (heavy–tailed distribution, equal covariances, Table 3).

(D) The first set of explanatory functions are defined as in (C). For the second set we consider

$$X^*(t) = X(t)[X(1) - X(t)],$$

where $X(t)$ is defined in (C) (heavy–tailed distribution, different covariances, Table 4).

*Department of Mathematics, University of Utah, Salt Lake City, UT 84112 (E-mail: horvath@math.utah.edu)

†Department of Mathematics and Statistics, Utah State University, Logan, UT 84322 (E-mail: Piotr Kokoszka@usu.edu)

‡Department of Statistics, University of Chicago, Chicago, IL 60637 (E-mail: mreimherr@uchicago.edu)

In all settings we used the test based on the most general statistic. The results are based on 100 replications. As we can see from the tables, the method works fairly well. The empirical sizes are close to the nominal sizes (first two columns of each table), and the power increases with the size of the difference. The power is smaller if the explanatory functions do not have a common distribution, and/or are heavy–tailed.

Table 1: Empirical Rejection Levels (A)

$\alpha/(c,c^*)$	(0,0)	(1,1)	(1,0)	(1.5,0)	(2,0)
0.1	0.14	0.08	0.50	0.90	0.98
0.05	0.09	0.03	0.40	0.81	0.98
0.01	0.03	0.00	0.18	0.63	0.92

Table 2: Empirical Rejection Levels (B)

$\alpha/(c,c^*)$	(0,0)	(1,1)	(1,0)	(1.5,0)	(2,0)
0.1	0.14	0.09	0.45	0.90	0.98
0.05	0.08	0.06	0.28	0.80	0.95
0.01	0.00	0.01	0.14	0.60	0.93

Table 3: Empirical Rejection Levels (C)

$\alpha/(c,c^*)$	(0,0)	(1,1)	(1,0)	(1.5,0)	(2,0)
0.1	0.11	0.10	0.47	0.85	0.99
0.05	0.04	0.05	0.32	0.78	0.95
0.01	0.02	0.02	0.18	0.60	0.87

Table 4: Empirical Rejection Levels (D)

$\alpha/(c,c^*)$	(0,0)	(1,1)	(1,0)	(1.5,0)	(2,0)	(2.5,0)	(3,0)	(3.5,0)
0.1	0.12	0.09	0.27	0.33	0.49	0.77	0.87	0.93
0.05	0.04	0.05	0.17	0.22	0.37	0.63	0.78	0.89
0.01	0.01	0.02	0.05	0.07	0.23	0.48	0.53	0.70