Supplement to "Specification testing in random coefficient models"

(*Quantitative Economics*, Vol. 9, No. 3, November 2018, 1371–1417)

Christoph Breunig
School of Business and Economics, Humboldt-Universität zu Berlin

Stefan Hoderlein
Department of Economics, Boston College

This Supplemental Material consists of additional finite sample results and extensions to the random coefficient models considered in the main text. Below, we show that our testing procedures can be extended to two different models. First, we consider the class of heterogeneous binary response models. Second, we discuss an extension of linear random coefficient models to system of equations. In both cases, we again discuss testing functional form restrictions and testing degeneracy of some random coefficients separately.

1. Additional finite sample results

In the following, we present finite sample results using alternative weight functions. Instead of using the standard normal p.d.f. as in the simulations in the main text we consider in the following the density of the uniform distribution as weight function.

Testing functional form restrictions

The data is generated as described in Section 3.1. We also implement the test statistic as described there but use the weight function $\varpi(t) = 0.25 \cdot 1_{[-2,2]}(t)$. Due to the bounded support, the test results are more sensitive with respect to the choice of the variance. Here, we choose the absolute value of the boundary of the support to coincide with the variance of X_1.

Overall our findings are in line with the literature, where test involving conditional characteristic functions are typically found insensitive w.r.t. the associated weighting function; see, for instance, Chen and Hong (2010). Indeed, as we see from Table 1, even using a limited support weight function, such as the uniform density, leads to overall similar empirical rejection probabilities (see Table 1). We also see from Table 1, that the empirical rejection probabilities are somewhat lower and in some cases the finite sample coverage is not accurate (see row 13). We also emphasize that the finite sample

Christoph Breunig: christoph.breunig@hu-berlin.de
Stefan Hoderlein: stefan_hoderlein@yahoo.com

Available at http://qeconomics.org. https://doi.org/10.3982/QE757

TABLE 1. Rows 1, 2, 7, 8, 13, 14, 19, 20 depict the empirical rejection probabilities if H_{mod} holds true, the rows 3–6, 9–12, 15–18, 21–24 show the finite sample power of our tests against various alternatives. The first column states the null model while the second shows the alternative model and is left empty if the null model is the correct model. Column 3 specifies the noise level of the data generating process. Column 4 depicts the values of the varying dimension parameters k_n. Columns 5–7 depict the empirical rejection probabilities for the nominal level 0.05 and $\varpi(t) = 0.25 \cdot 1_{[-2,2]}(t)$.

	Null Model	Alt. Model			Empirical Rejection Probabilities Using		
Rows	H_{mod}	True DGP	η	k_n	$m_n = 9$	$m_n = 12$	$m_n = 15$
1	(3.1)		0.7	5	0.013	0.004	0.001
2	(3.2)			7	0.070	0.018	0.007
3	(3.1)	(3.2)		5	0.935	0.785	0.635
4	(3.1)	(3.3)			0.448	0.173	0.098
5	(3.2)	(3.1)		7	0.837	0.653	0.450
6	(3.2)	(3.3)			0.944	0.758	0.551
7	(3.1)		1	4	0.037	0.005	0.005
8	(3.2)			7	0.220	0.090	0.030
9	(3.1)	(3.2)		5	0.741	0.512	0.341
10	(3.1)	(3.3)			0.290	0.091	0.035
11	(3.2)	(3.1)		7	0.981	0.860	0.676
12	(3.2)	(3.3)			0.982	0.877	0.718
13	(3.1)		0.7	6	0.010	0.000	0.001
14	(3.2)			9	0.064	0.020	0.007
15	(3.1)	(3.2)		6	0.792	0.501	0.351
16	(3.1)	(3.3)			0.335	0.108	0.068
17	(3.2)	(3.1)		9	0.873	0.613	0.437
18	(3.2)	(3.3)			0.931	0.776	0.562
19	(3.1)		1	6	0.031	0.003	0.005
20	(3.2)			9	0.215	0.087	0.037
21	(3.1)	(3.2)		6	0.678	0.387	0.231
22	(3.1)	(3.3)			0.197	0.071	0.020
23	(3.2)	(3.1)		9	0.972	0.867	0.694
24	(3.2)	(3.3)			0.985	0.876	0.683

properties of the test statistic are more sensitive to the choice of the support of $[-2, 2]$. In contrast, when choosing normal density weights, the values of our standardized tests is very insensitive to the choice of the variance of the normal distribution.

Testing degeneracy

The test is implemented as for Table 2 but in the following we use the weighting function $\varpi(t) = 0.5 \cdot 1_{[-1,1]}(t)$. As above, the absolute value of the boundary of the support is thus chosen to coincide with the variance of X_1. As above, we see from Table 2 that the empirical rejection probabilities are somewhat smaller than compared to the standard

TABLE 2. The first row depicts the empirical rejection probabilities under degeneracy of the coefficient of X_2, the rows 2–4, 6–8, 10–12, and 14–16 show the finite sample power of our tests against various alternatives. Column 1 depicts the value of κ in the correct and alternative models. Column 2 specifies the covariance of B_1 and B_2 for the alternative models. Column 3 depicts the value of η in the correct model and is empty if the null model is correct. Columns 4–7 depict the empirical rejection probabilities for the nominal level 0.05 and $\varpi(t) = 0.5 \cdot 1_{[-1,1]}(t)$.

				Empirical Rejection Probabilities Using			
				$k_n = 4$		$k_n = 5$	
Rows	κ	ρ	Alt. Model η	$m_n = 16$	$m_n = 20$	$m_n = 20$	$m_n = 25$
1	1	1		0.028	0.020	0.012	0.008
2			0.3	0.230	0.194	0.123	0.118
3			0.5	0.573	0.550	0.489	0.430
4			0.7	0.860	0.836	0.798	0.759
5		1.5		0.000	0.002	0.000	0.000
6			0.3	0.191	0.113	0.012	0.012
7			0.5	0.555	0.427	0.152	0.116
8			0.7	0.866	0.794	0.536	0.437
9	2	1		0.001	0.000	0.000	0.000
10			0.3	0.078	0.056	0.004	0.003
11			0.5	0.320	0.238	0.057	0.032
12			0.7	0.671	0.560	0.310	0.229
13		1.5		0.024	0.025	0.013	0.011
14			0.3	0.322	0.276	0.224	0.199
15			0.5	0.724	0.635	0.612	0.578
16			0.7	0.937	0.900	0.888	0.852

normal weight as depicted in Table 2. In some cases, this leads again to inaccurate finite sample coverage (see row 9). Overall the results are relatively similar even though the weight functions has only limited support.

1.1 *Binary response models*

We consider the binary response model

$$Y = 1\{g(X,B) < Z\}, \tag{1.1}$$

where, besides the dependent variable Y and covariates X, a *special regressor* Z is observed as well. In the following, we assume that (X, Z) is independent of B. In contrast to the previous section, the test in the binary response model is based on the difference of a partial derivative of the conditional success probability $P(Y = 1|X, Z)$ and a restricted transformation of the p.d.f. f_B.

Testing functional form restrictions In the binary response model (1.1), observe that

$$P[Y=1|X=x,Z=z]=\int 1\{z>g(x,b)\}f_B(b)\,db$$
$$=\int_{-\infty}^{z}\int_{P_{x,s}} f_B(b)\,d\nu(b)\,ds,$$

where ν is the Lebesgue measure on the lower dimensional hyperplane $P_{x,s}=\{b: g(x,b)=s\}$. Consequently, it holds

$$\psi(x,z)\equiv\partial_z P[Y=1|X=x,Z=z]=\int_{P_{x,z}} f_B(b)\,d\nu(b).$$

We consider the null hypothesis $H_{\text{mod}}: Y=1\{g(X,B)<Z\}$ for some random coefficient B. By using the above integral representation of ψ, the null hypothesis H_{mod} is equivalent to $(\mathcal{F}\psi(X,\cdot))(t)=(\mathcal{F}_g f_B)(t,X)$ (recall the definition of the integral transform $(\mathcal{F}_g f)(X,t)\equiv\int\exp(itg(X,b))f(b)\,db$). Instead of using the previous equation, we invert the Fourier transform and conclude that equation (2.3) holds true with

$$\varepsilon(X,z)=\psi(X,z)-\big(\mathcal{F}^{-1}\big[(\mathcal{F}_g f_B)(X,\cdot)\big]\big)(z).$$

Due to nonsingularity of the Fourier transform $\varepsilon(X,z)=0$ is indeed equivalent to $(\mathcal{F}\psi(X,\cdot))(t)=(\mathcal{F}_g f_B)(t,X)$. In the case of a linear g, the random coefficient density f_B is thus identified through the Radon transform; see also Gautier and Hoderlein (2015).

To estimate the function ε, we replace ψ by a series least squares estimator. Let us introduce the matrix $\mathbf{W}_n=(p_{\underline{m_n}}(X_1,Z_1),\dots,p_{\underline{m_n}}(X_n,Z_n))'$ where the basis function p_l, $l\geq 1$, are assumed to be differentiable with respect to the (d_x+1)th entry. We estimate ψ by

$$\widehat{\psi}_n(x,z)=\partial_z p_{\underline{m_n}}(x,z)'\big(\mathbf{W}_n'\mathbf{W}_n\big)^{-1}\mathbf{Y}_n,$$

where $\mathbf{Y}_n=(Y_1,\dots,Y_n)'$. Consequently, we replace the function ε by

$$\widehat{\varepsilon}_n(X_j,z)=\widehat{\psi}_n(X_j,z)-\big(\mathcal{F}^{-1}\big[(\mathcal{F}_g\widehat{f}_{Bn})(X_j,\cdot)\big]\big)(z),$$

where $\widehat{f}_{Bn}$ is the sieve minimum distance estimator given by

$$\widehat{f}_{Bn}\in\operatorname*{arg\,min}_{f\in\mathcal{B}_n}\left\{\sum_{j=1}^{n}\int\big|\widehat{\psi}_n(X_j,z)-\big(\mathcal{F}^{-1}\big[(\mathcal{F}_g f)(X_j,\cdot)\big]\big)(z)\big|^2\varpi(z)\,dz\right\}\tag{1.2}$$

and $\mathcal{B}_n=\{\phi(\cdot)=\sum_{l=1}^{k_n}\beta_l q_l(\cdot)\}$. Our test statistic is $S_n=n^{-1}\sum_{j=1}^{n}\int|\widehat{\varepsilon}_n(X_j,z)|^2\varpi(z)\,dz$ where, in this section, ϖ is an integrable weighting function on the support of Z.

We introduce an m_n dimensional linear sieve space $\Psi_n\equiv\{\phi:\phi(x,z)=\sum_{l=1}^{m_n}\beta_l p_l(x,z)\}$. Let $p_{\underline{m_n}}(X,Z)$ be a tensor-product of vectors of basis functions $p_{\underline{m_{n_1}}}(X)$ and $p_{\underline{m_{n_2}}}(Z)$ for integers m_{n_1} and m_{n_2} with $m_n=m_{n_1}\cdot m_{n_2}$. We assume that $\partial_z p_{\underline{m_{n_2}}}(z)=(p_0(z),2p_1(z),\dots,m_{n_2}p_{m_{n_2}-1}(z))'$. Further, let τ_l denote the squared integer that is associated with $\partial_z p_l$. In Definition 1, $p_l(X)$ has to be replaced by $\tau_l p_l(X,Z)$. Let

$B_n = \int E[\overline{(\mathcal{F}^{-1}[(\mathcal{F}_g q_{\underline{k_n}})(X,\cdot)](z)}(\mathcal{F}^{-1}[(\mathcal{F}_g q_{\underline{k_n}})(X,\cdot)](z)']\varpi(z)\,dz$, which is denoted by $\widehat{B}_n$ when the expectation is replaced by the sample mean. In contrast to Assumption 4, we assume in the following that $\|B_n^-\|$ is bounded from above.

Assumption 1.

(i) *The random vector* (X, Z) *is independent of* B.

(ii) *For any p.d.f.* f_B *satisfying* $\mathcal{F}\psi = \mathcal{F}_g f_B$ *there exists* $\Pi_{k_n} f_B \in \Psi_n$ *such that* $n\|\psi - \mathcal{F}^{-1}\mathcal{F}_g \Pi_{k_n} f_B\|_\varpi^2 = o(\sqrt{m_n})$.

(iii) *There exists* $\Pi_{m_n}\psi \in \Psi_n$ *such that* $n\|\Pi_{m_n}\psi - \psi\|_\varpi^2 = o(\sqrt{m_n})$.

(iv) *It holds* $\|B_n^-\| = O(1)$ *and* $P(\operatorname{rank}(B_n) = \operatorname{rank}(\widehat{B}_n)) = 1 + o(1)$.

(v) *It holds* $k_n \log n = o(\sqrt{m_n})$ *and* $m_n^2(\log n)\sum_{l=1}^{m_n}\tau_l = o(n\lambda_n)$.

Assumption 1 is similar to Assumption 4. Note that due to the partial derivatives of the basis functions we need to be more restrictive about the dimension parameter m_n, which is captured in Assumption 1(iv). The following result establishes the asymptotic distribution of our test statistic under H_{mod} in the binary response model (1.1).

Proposition 1. *Let Assumptions* 2, 3, *and* 1 *hold with* $\delta(Y, X, Z) = Y - \int 1\{Z > X'b\} f_B(b)\,db$. *Then, under* H_{mod} *we have*

$$(\sqrt{2}\varsigma_{m_n})^{-1}(nS_n - \mu_{m_n}) \xrightarrow{d} \mathcal{N}(0,1).$$

The critical values can be estimated as in Remark 2.1 but where now $\delta_n(Y, X, Z) = Y - \int 1\{Z \geq X'b\}\widehat{f}_{Bn}(b)\,db$ with the estimator $\widehat{f}_{Bn}$ given in (1.2).

Testing degeneracy To keep the presentation simple, we only consider the linear case in the following. Under H_{lin}, the binary response model (1.1) simplifies to

$$Y = 1\{X'B < Z\}. \tag{1.3}$$

$$\int \exp(itz)\psi(X,z)\,dz = \int \exp(itz)\psi(X_1, z - X_2'b_2)\,dz.$$

By nonsingularity of the Fourier transform, we conclude that H_{deg} is equivalent to equation (2.3) where

$$\varepsilon(X,z) = \psi(X,z) - \psi(X_1, z - X_2'b_2).$$

If ψ only depends on X_1, we consider the estimator $\widehat{\psi}_{1n}(x_1, z) = \partial_z p_{\underline{k_n}}(x_1, z)'(\mathbf{W}_{n1}'\mathbf{W}_{n1})^{-1}\mathbf{Y}_n$, where $\mathbf{W}_{n1} = (p_{\underline{k_n}}(X_{11}, Z_1), \ldots, p_{\underline{k_n}}(X_{1n}, Z_n))'$. We propose a minimum distance estimator of b_2 given by

$$\widehat{b}_{2n} = \operatorname*{arg\,min}_{\beta\in\mathcal{B}} \sum_{j=1}^n \int \left|\widehat{\psi}_n(X_j, t) - \widehat{\psi}_{1n}(X_{1j}, t - \beta' X_2)\right|^2 \varpi(t)\,dt. \tag{1.4}$$

Consequently, we estimate the function ε by $\widehat{\varepsilon}_n(X_j, z) = \widehat{\psi}_n(X_j, z) - \widehat{\psi}_{1n}(X_{1j}, z - \widehat{b}_{2n} X_{2j})$.

Proposition 2. *Let Assumptions* 2, 3, 1(i), (iii), (v) *with* $\delta(Y, X, Z) = Y - P(Y = 1|X_1, Z - X_2' b_2)$, *and* H_{lin} *hold true. Assume that* $n \int E|(\Pi_{k_n}\psi)(X_1, z) - \psi(X_1, z)|^2 \varpi(z)\, dz = o(\sqrt{m_n})$. *Then, under* H_{deg} *we have*

$$(\sqrt{2}\varsigma_{m_n})^{-1}(nS_n - \mu_{m_n}) \xrightarrow{d} \mathcal{N}(0, 1).$$

The critical values can be estimated as in Remark 2.1 by replacing $P(Y = 1|X_1, Z - X_2' b_2)$ by a series least squares estimator.

1.2 *Application to systems of equations*

In this subsection, we apply our testing procedure to systems of equations, that is, situations in which the endogenous variable is not a scalar, but a vector. For simplicity, we consider in the following only the bivariate case. Formally, we consider the model

$$Y = g(X, B), \tag{1.5}$$

for some function g and $Y \in \mathbb{R}^2$. Again the vector of random coefficients $B = (B_0, B_1, B_2, B_3)$ is assumed to be independent of the covariates X.

Testing functional form restrictions Null hypothesis H_{mod} is equivalent to equation (2.3) with

$$\varepsilon(X, t) = E\big[\exp(it'Y) - \exp(it' g(X, B))|X\big]$$

for some $t \in \mathbb{R}^2$. Our test of H_{mod} is now based on $S_n \equiv n^{-1} \sum_{j=1}^{n} \int |\widehat{\varepsilon}_n(X_j, t)|^2 \varpi(t)\, dt$ where $\widehat{\varepsilon}_n$ is the estimator of ε introduced in Example 1 but with a multivariate index t and ϖ being a weighting function on $\mathbb{R}^2$. Under a slight modification of assumptions required for Theorem 2.1, asymptotic normality of the standardized test statistic S_n follows under H_{mod}.

Testing degeneracy In the partially linear case (i.e., $H_{\text{part-lin}}$ holds), the random coefficient model (1.5) simplifies to

$$\begin{aligned} Y_1 &= B_0 + B_{11}' X_1 + B_{12}' X_2, \\ Y_2 &= B_2 + B_{31}' X_1 + g_2(X_2, B_{32}). \end{aligned}$$

This model is identified if B_{32} is degenerate (see Hoderlein, Holzmann, and Meister (2014)). A test for degeneracy of $H_{\text{deg}} : B_{32} = b$, for some nonstochastic vector b, uses only the second equation, that is,

$$E\big[\exp(itY_2)|X\big] = E\big[\exp\big(it\big(B_2 + B_{31}' X_1\big)\big)|X_1\big] \exp\big(itg_2(X_{32}, b)\big).$$

We can consequently use the testing methodology developed in Section 2.3.4.

1.3 *Proofs of Section 3*

In the following, we make use of the notation $\widehat{\alpha}_n \equiv (n\widehat{R}_n)^{-1}\sum_j Y_j p_{\underline{m_n}}(X_j, Z_j)$ where $\widehat{R}_n = n^{-1}\sum_j p_{\underline{m_n}}(X_j, Z_j) p_{\underline{m_n}}(X_j, Z_j)'$. The Kronecker product for matrices is denoted by $\otimes$.

Proof of Proposition 1. We make use of the decomposition

$$
\begin{aligned}
nS_n = &\sum_j \int \big|\partial_z p_{\underline{m_n}}(X_j, z)'\big(\widehat{\alpha}_n - E\big[1\big\{Z > g(X,B)\big\} p_{\underline{m_n}}(X,Z)\big]\big)\big|^2 \varpi(z)\,dz \\
&+ 2\sum_j \int \partial_z p_{\underline{m_n}}(X_j, z)'\big(\widehat{\alpha}_n - E\big[1\big\{Z > g(X,B)\big\} p_{\underline{m_n}}(X,Z)\big]\big) \\
&\times \big(\partial_z p_{\underline{m_n}}(X_j, z)' E\big[1\big\{Z > g(X,B)\big\} p_{\underline{m_n}}(X,Z)\big] \\
&- \big(\mathcal{F}^{-1}\big[(\mathcal{F}_g \widehat{f}_{Bn})(X_j,\cdot)\big]\big)(z)\big)\varpi(z)\,dz \\
&+ \sum_j \int \big|\partial_z p_{\underline{m_n}}(X_j, z)' E\big[1\big\{Z > g(X,B)\big\} p_{\underline{m_n}}(X,Z)\big] \\
&- \big(\mathcal{F}^{-1}\big[(\mathcal{F}_g \widehat{f}_{Bn})(X_j,\cdot)\big]\big)(z)\big|^2 \varpi(z)\,dz \\
= &\, I_n + 2II_n + III_n \quad \text{(say)}.
\end{aligned}
$$

Consider I_n. For all $l \geq 1$, the derivative of a basis function p_l is given by $l p_{l-1}$. Since p_l forms an orthonormal basis in $L^2_{\varpi}(\mathbb{R})$ holds

$$
\begin{aligned}
I_n = &\frac{n}{\lambda_n}\big(\widehat{\beta}_{m_n} - E\big[1\big\{Z > g(X,B)\big\} p_{\underline{m_n}}(X,Z)\big]\big)' \\
&\times (I_{m_{1n}} \otimes T_n)\big(\widehat{\beta}_{m_n} - E\big[1\big\{Z > g(X,B)\big\} p_{\underline{m_n}}(X,Z)\big]\big) + o_p(\sqrt{m_n}),
\end{aligned}
$$

where T_n is a $m_{2n} \times m_{2n}$ diagonal matrix with lth diagonal element is given by $(l-1)^2$. It holds

$$
I_n = \lambda_n^{-1} \sum_{l=1}^{m_n} \tau_l \Big| n^{-1/2} \sum_j \Big(Y_j - \int 1\big\{Z_j > g_2(X_j, b)\big\} f_B(b)\,db \Big) p_l(X_j, Z_j) \Big|^2 + o_p(1).
$$

Thus, Lemma A.2 yields $(\sqrt{2}\varsigma_{m_n})^{-1}(I_n - \mu_{m_n}) \xrightarrow{d} \mathcal{N}(0,1)$. Consider III_n. We have

$$
\begin{aligned}
III_n \lesssim &\sum_j \int \big|(\Pi_{m_n}\psi)(X_j, z) - \psi(X_j, z)\big|^2 \varpi(z)\,dz \\
&+ \sum_j \int \big|\big(\mathcal{F}^{-1}\big[\big(\mathcal{F}_g(\widehat{f}_{Bn} - f_B)\big)(X_j,\cdot)\big]\big)(z)\big|^2 \varpi(z)\,dz \\
= &\, A_{n1} + A_{n2}.
\end{aligned}
$$

We have $A_{n1} = O_p(n\|\Pi_{m_n}\psi - \psi\|_{\varpi}^2) = o_p(\sqrt{m_n})$ and

$$A_{n2} \lesssim \sum_j \int \big|\big(\mathcal{F}^{-1}\big[\big(\mathcal{F}_g(\widehat{f}_{Bn} - \Pi_{k_n} f_B)\big)(X_j,\cdot)\big]\big)(z)\big|^2 \varpi(z)\,dz$$
$$+ \sum_j \int \big|\big(\mathcal{F}^{-1}\big[\big(\mathcal{F}_g(\Pi_{k_n} f_B - f_B)\big)(X_j,\cdot)\big]\big)(z)\big|^2 \varpi(z)\,dz,$$

where the second summand on the right-hand side is of the order $o_p(\sqrt{m_n})$. Further,

$$\sum_j \int \big|\big(\mathcal{F}^{-1}\big[\big(\mathcal{F}_g(\widehat{f}_{Bn} - \Pi_{k_n} f_B)\big)(X_j,\cdot)\big]\big)(z)\big|^2 \varpi(z)\,dz$$
$$= (\widehat{\beta}_n - \beta_n)'$$
$$\times \sum_j \int \overline{\big(\mathcal{F}^{-1}\big[(\mathcal{F}_g q_{\underline{k_n}})(X_j,\cdot)\big]\big)(z)}\big(\mathcal{F}^{-1}\big[(\mathcal{F}_g q_{\underline{k_n}})(X_j,\cdot)\big]\big)(z)'\varpi(z)\,dz(\widehat{\beta}_n - \beta_n)$$

and thus, following the proof of Theorem 2.1 we obtain $A_{n2} = o_p(\sqrt{m_n})$. Similarly as in the proof of Theorem 2.1, it can be seen that $II_n = o_p(\sqrt{m_n})$, which completes the proof. □

Proof of Proposition 2. We decompose our test statistic as

$$nS_n = \sum_j \int \big|\partial_z p_{\underline{m_n}}(X_j,z)'\big(\widehat{\alpha}_n - E\big[Y p_{\underline{m_n}}(X,Z)\big]\big)\big|^2 \varpi(z)\,dz$$
$$+ 2\sum_j \int \big(\partial_z p_{\underline{m_n}}(X_j,z)'\widehat{\alpha}_n - (\Pi_{m_n}\psi)(X_j,z)\big)$$
$$\times \big((\Pi_{m_n}\psi)(X_j,z) - \widehat{\psi}_{1n}\big(X_{1j}, z - X_{2j}'\widehat{b}_{2n}\big)\big)\varpi(z)\,dz$$
$$+ \sum_j \int \big|(\Pi_{m_n}\psi)(X_j,z) - \widehat{\psi}_{1n}\big(X_{1j}, z - X_{2j}'\widehat{b}_{2n}\big)\big|^2 \varpi(z)\,dz$$
$$= I_n + 2II_n + III_n \quad \text{(say)}.$$

Consider I_n. As in the proof of Proposition 1, we obtain

$$I_n = \lambda_n^{-1} \sum_{l=1}^{m_n} \tau_l \Big| n^{-1/2} \sum_j \Big(Y_j - \int 1\{Z_j \geq X_1'b_1 + X_2'b_2\} f_{B_1}(b_1)\,db_1\Big) p_l(X_j,Z_j)\Big|^2$$
$$+ o_p(1),$$

and thus, Lemma A.2 yields $(\sqrt{2}\varsigma_{m_n})^{-1}(I_n - \mu_{m_n}) \xrightarrow{d} \mathcal{N}(0,1)$. Concerning III_n, we calculate

$$III_n \lesssim \sum_j \int \big|(\Pi_{m_n}\psi)(X_j,z) - \psi(X_j,z)\big|^2 \varpi(z)\,dz$$

$$+\sum_j \int \left|\psi\big(X_{1j}, z - X_{2j}'b_2\big) - \widehat{\psi}_n(X_j, z)\right|^2 \varpi(z)\,dz$$
$$+\sum_j \int \left|\widehat{\psi}_n(X_j, z) - \widehat{\psi}_{1n}\big(X_{1j}, z - X_{2j}'\widehat{b}_{2n}\big)\right|^2 \varpi(z)\,dz.$$

The definition of the estimator $\widehat{b}_2$ in (1.4) yields

$$\begin{aligned}
&\sum_j \int \left|\widehat{\psi}_n(X_j, z) - \widehat{\psi}_{1n}\big(X_{1j}, z - X_{2j}'\widehat{b}_{2n}\big)\right|^2 \varpi(z)\,dz \\
&\quad \le \sum_j \int \left|\widehat{\psi}_n(X_j, z) - \widehat{\psi}_{1n}\big(X_{1j}, z - X_{2j}'b_2\big)\right|^2 \varpi(z)\,dz \\
&\quad \lesssim \sum_j \int \left|\widehat{\psi}_n(X_j, z) - \psi(X_j, z)\right|^2 \varpi(z)\,dz \\
&\qquad + \sum_j \int \left|\widehat{\psi}_{1n}\big(X_{1j}, z - X_{2j}'b_2\big) - \psi\big(X_{1j}, z - X_{2j}'b_2\big)\right|^2 \varpi(z)\,dz \\
&\quad = o_p(\sqrt{m_n}).
\end{aligned}$$

It thus follows $III_n = o_p(\sqrt{m_n})$. Similarly as in the proof of Theorem 2.3 it can be shown that $II_n = o_p(\sqrt{m_n})$. □

References

Chen, B. and Y. Hong (2010), "Characteristic function-based testing for multifactor continuous-time Markov models via nonparametric regression." *Econometric Theory*, 26 (04), 1115–1179.[1]

Gautier, E. and S. Hoderlein (2015), "A triangular treatment effect model with random coefficients in the selection equation." Technical report, Toulouse School of Economics. [4]

Hoderlein, S., H. Holzmann, and A. Meister (2014), "The triangular model with random coefficients." Technical report, Boston College. [6]

Co-editor Rosa L. Matzkin handled this manuscript.

Manuscript received 15 August, 2016; final version accepted 12 February, 2018; available online 5 April, 2018.