SUPPLEMENTARY MATERIAL FOR "SEMIPARAMETRIC ESTIMATION OF STRUCTURAL FUNCTIONS IN NONSEPARABLE TRIANGULAR MODELS"

VICTOR CHERNOZHUKOV[†], IVÁN FERNÁNDEZ-VAL[§], WHITNEY NEWEY[‡], SAMI STOULI[¶], AND FRANCIS VELLA[|]

1. Summary

In the Supplementary Material we first report results from a sensitivity analysis we carried out to check the robustness of our empirical results. In Section 2.1, we report additional QSF estimates obtained for different regions of interest, with grids of values of X of varying cardinality and length, as well as with additional quantile levels. We also compare ASF estimates obtained by least-squares projection, as described in Remark 9 in the main text, to those obtained by QR. In Section 2.2 we report more flexible QSF estimates including additional powers of the control variable $\Phi^{-1}(V)^k$ as well as interaction terms $X \cdot \Phi^{-1}(V)^k$, $Z_1 \cdot \Phi^{-1}(V)^k$, and $X \cdot Z_1 \cdot \Phi^{-1}(V)^k$. The selection of these additional terms is investigated for the ASF by means of a least-squares cross-validation procedure. In Section 2.3, we exploit knowledge of the control function distribution and implement a simulation-based integration procedure as an alternative to sample averaging over the estimated control function. Finally in Section 3 we perform several Monte Carlo simulations in order to assess the finite sample performance of our estimators. We compare the DR and QR estimators of structural functions with several designs calibrated to the empirical application. We give a detailed description of our calibration procedure for the Monte Carlo simulations. Overall, our robustness checks show that our empirical results are robust to the modelling, estimation and integration choices, and our additional simulation results confirm the main findings for the ASF discussed in the main text.

Date: November 5, 2018.

[†] Department of Economics, MIT, vchern@mit.edu.
§ Department of Economics, Boston University, ivanf@bu.edu.
‡ Department of Economics, MIT, wnewey@mit.edu.
¶ Department of Economics, University of Bristol, s.stouli@bristol.ac.uk.
| Department of Economics, Georgetown University, Francis.Vella@georgetown.edu.

2. Robustness of Empirical Results

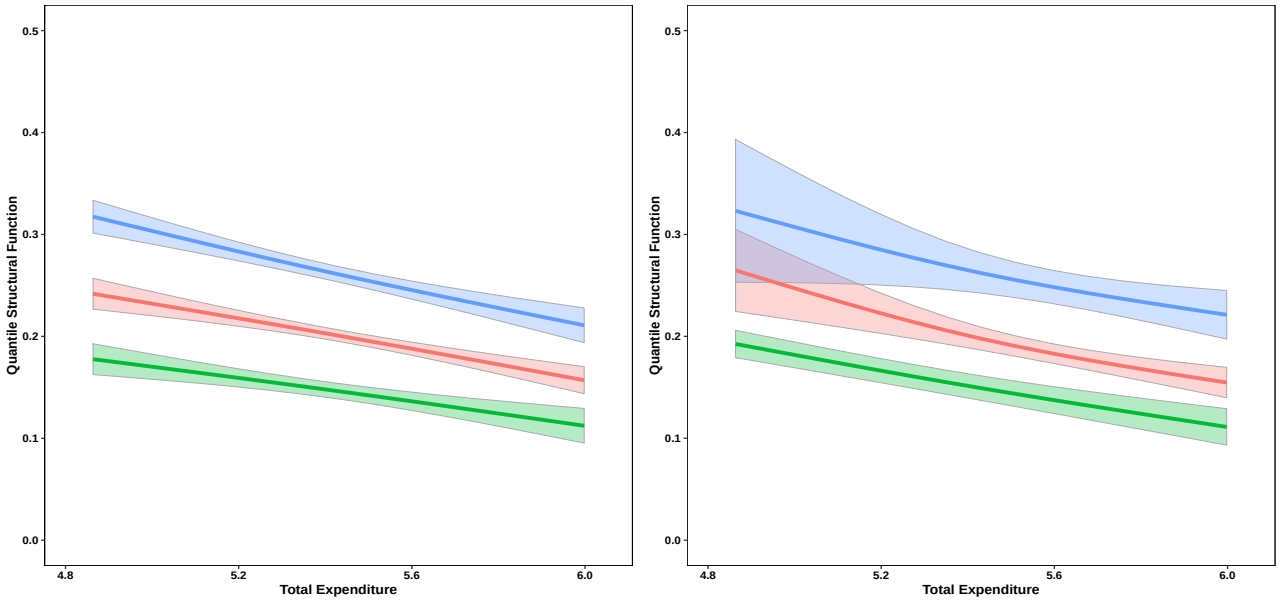
2.1. **Sensitivity Analysis.** To further check the robustness of our empirical findings, we implemented a thorough sensitivity analysis and investigated several alternative specifications. We replicated all results of the empirical application in the main text using a probit specification for DR, as well as enforcing a trimming rule, yielding very similar results which we do not report for brevity. For both QR and DR we report additional QSF estimates with varying grids of values of X, as well as for a different number of quantile levels.

For the equispaced grids $0.1 = t_1 < \cdots < t_K = 0.9$ and $0.15 = t_1^* < \cdots < t_K^* = 0.85$, let $\widetilde{\mathcal{X}}_K = \{\widehat{Q}_X(t_1), \ldots, \widehat{Q}_X(t_K)\}$ and $\widetilde{\mathcal{X}}_K^* = \{\widehat{Q}_X(t_1^*), \ldots, \widehat{Q}_X(t_K^*)\}$. Further let $\widetilde{\mathcal{T}}_3 = \{1/4, 1/2, 3/4\}$ and $\widetilde{\mathcal{T}}_5 = \{1/6, 1/3, 1/2, 2/3, 5/6\}$. Then Figures 2.1-2.4 display QSFs and their uniform confidence bands obtained by setting the regions of interest as follows:

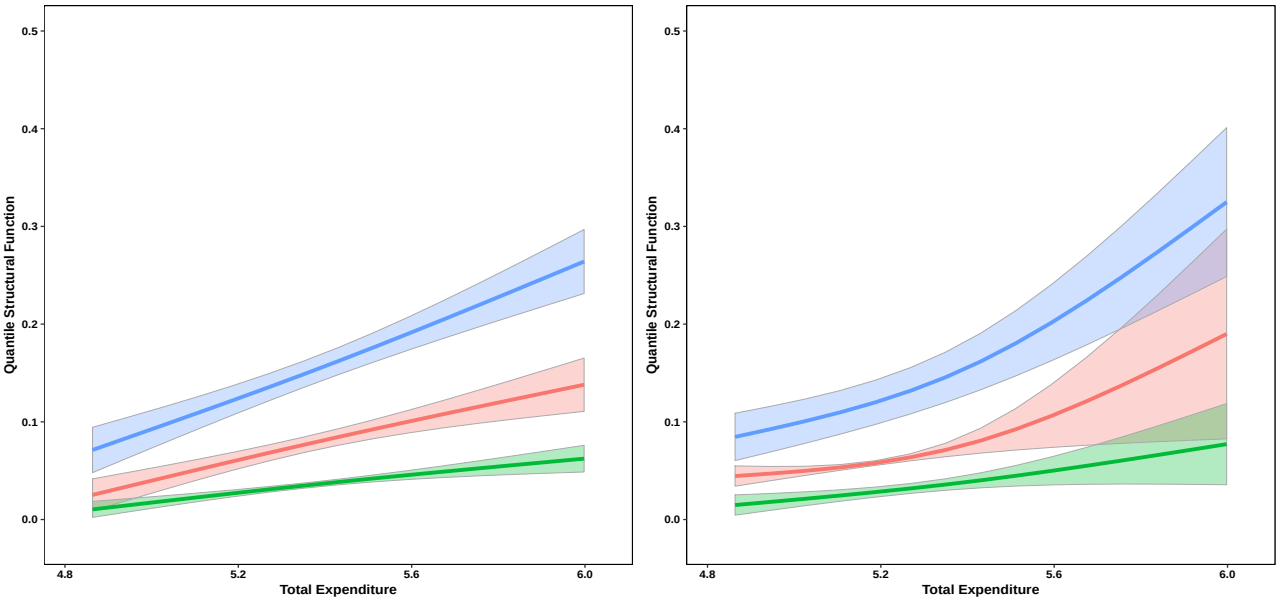
(1) Figure 2.1: we set $\mathcal{I}_Q = \widetilde{\mathcal{T}}_3 \widetilde{\mathcal{X}}_3$,
(2) Figure 2.2: we set $\mathcal{I}_Q = \widetilde{\mathcal{T}}_3 \widetilde{\mathcal{X}}_7$,
(3) Figure 2.3: we set $\mathcal{I}_Q = \widetilde{\mathcal{T}}_5 \widetilde{\mathcal{X}}_5$,
(4) Figure 2.4: we set $\mathcal{I}_Q = \widetilde{\mathcal{T}}_5 \widetilde{\mathcal{X}}_5^*$.

QSF estimates across varying regions of interest confirm the results of the empirical application in the main text. For QR, varying the number of grid points has very little effect on the QSF estimates and confidence bands. For DR, QSF estimates are also almost identical across specifications, and only the shape of confidence bands varies according to K. For both goods, and both DR and QR methods, all specifications capture the features emphasized in the main text: for both goods QSF estimates display heteroskedasticity, and estimates for leisure display asymmetry. These features are especially apparent in Figure 2.3 which shows the QSF at 5 different quantile levels. Finally, comparing Figures 2.1-2.4 shows that the length of confidence bands over $\widetilde{\mathcal{X}}$ is affected by the choice of end-points for $\widetilde{\mathcal{X}}_K$, especially so for DR estimates, but is robust to the choice of K.

For QR we also check the robustness of our ASF estimates by comparing them to those obtained based on the least-squares projection characterization of the ASF for the QR baseline given in Remark 9 in the main text. Figure 2.5 shows that the two estimates are very similar for both food and leisure share expenditure.

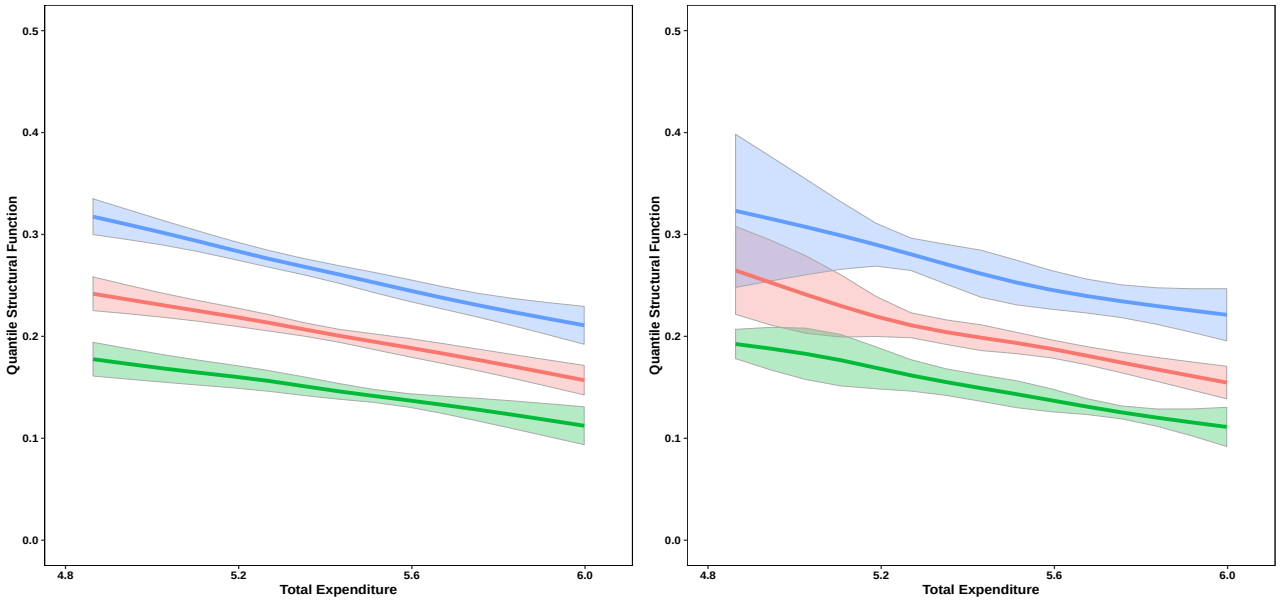


(A) Food.

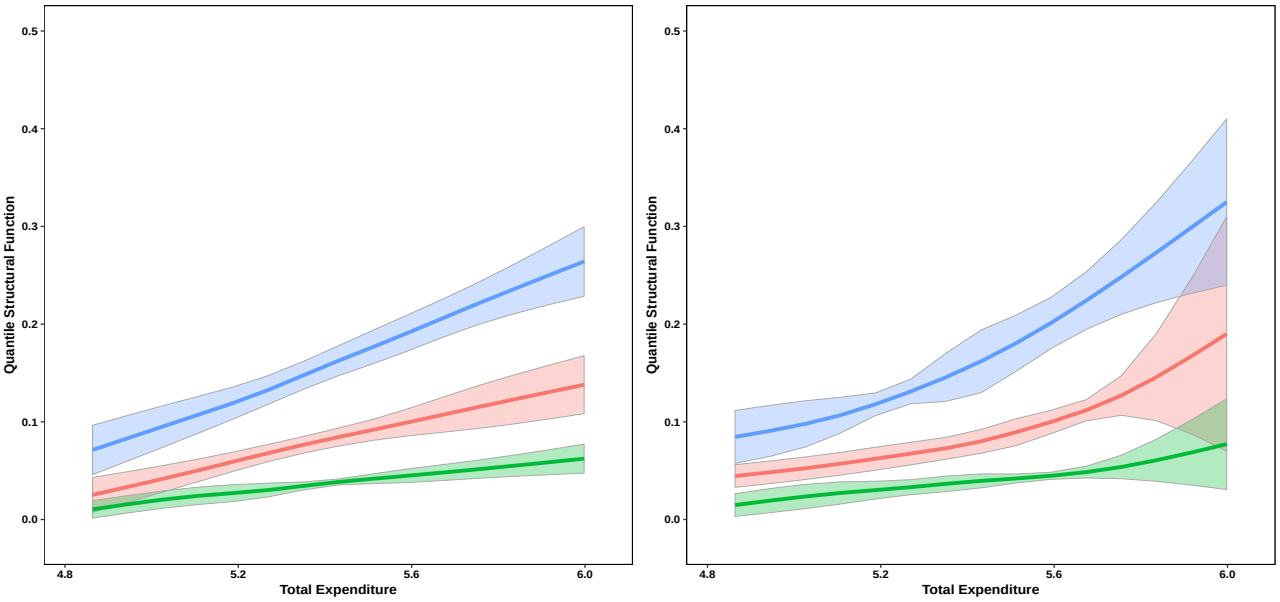


(B) Leisure.

FIGURE 2.1. QSF over $\widetilde{\mathcal{T}}_3\widetilde{\mathcal{X}}_3$. Quantile (left) and distribution regression (right).

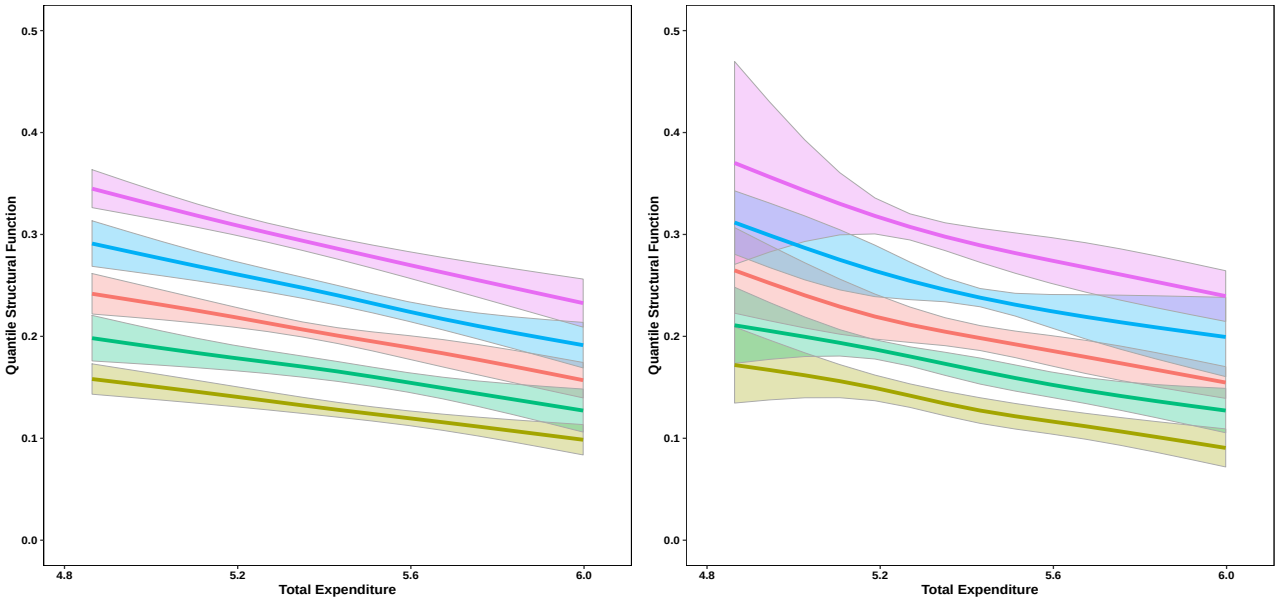


(A) Food.

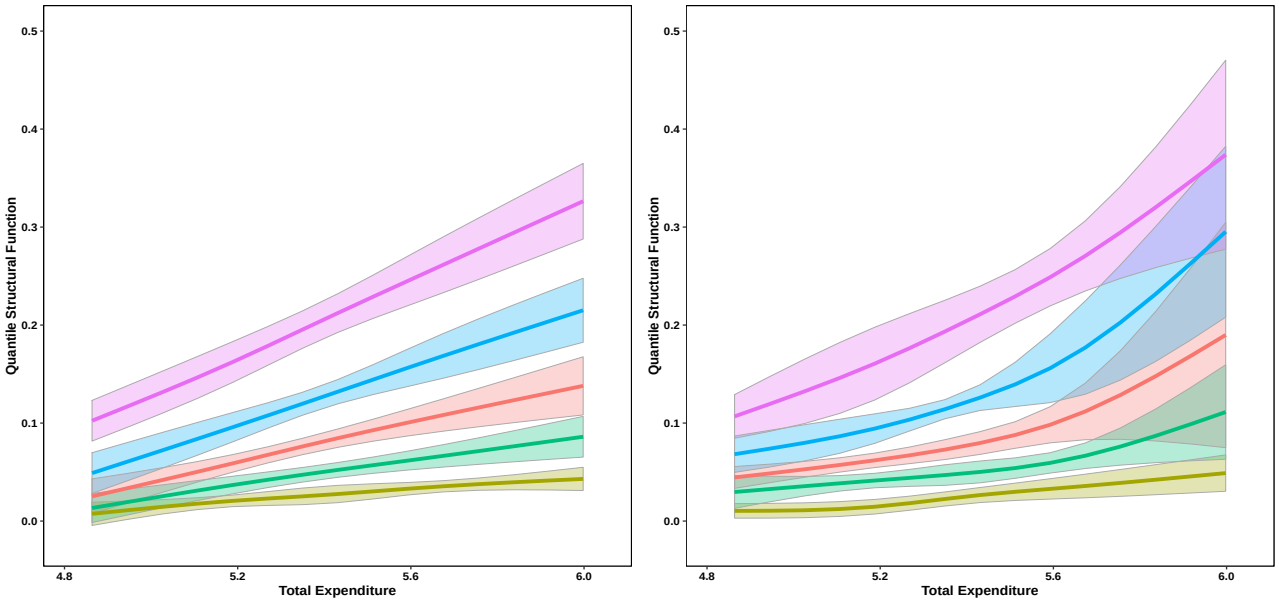


(B) Leisure.

FIGURE 2.2. QSF over $\widetilde{\mathcal{T}}_3\widetilde{\mathcal{X}}_7$. Quantile (left) and distribution regression (right).

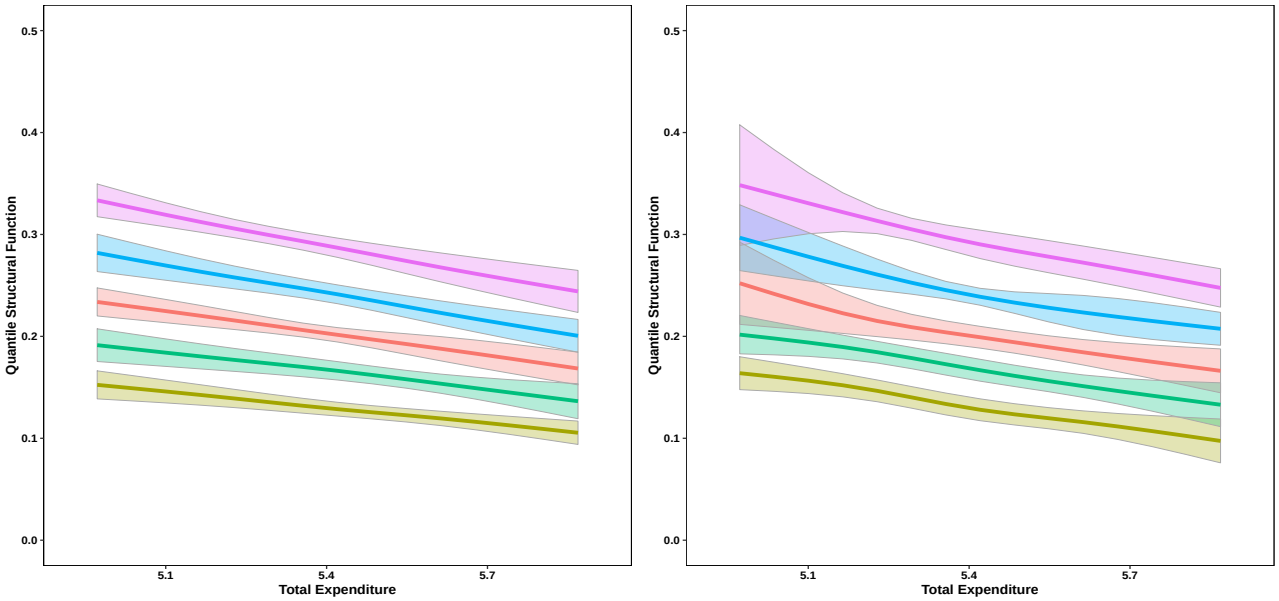


(A) Food.

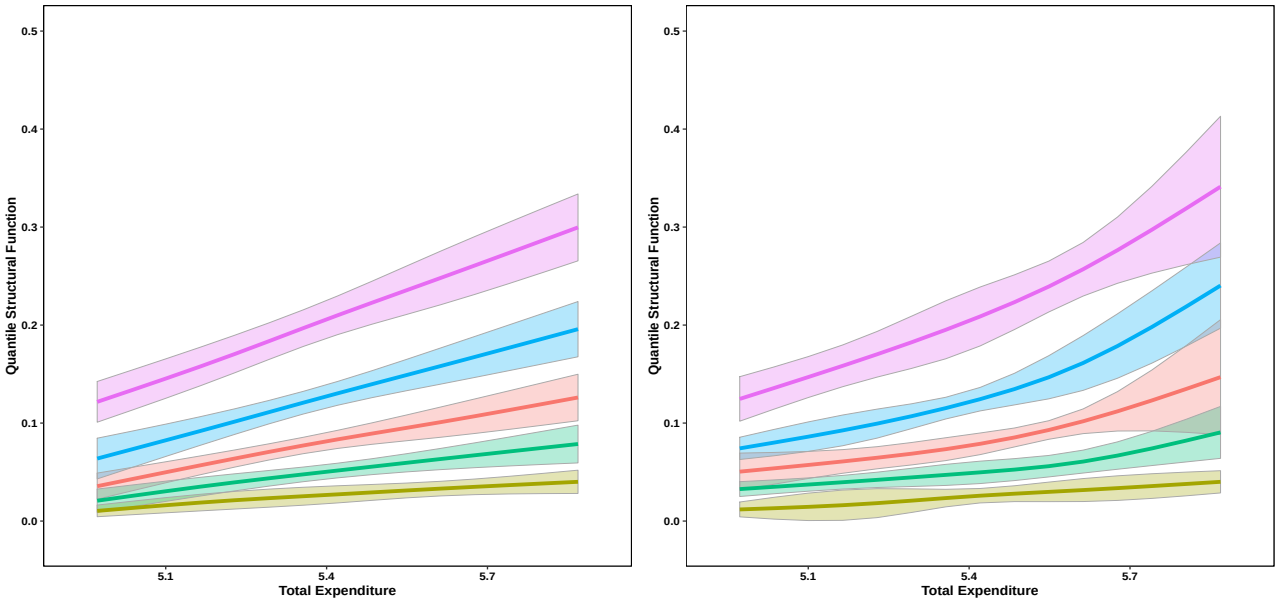


(B) Leisure.

FIGURE 2.3. QSF over $\widetilde{\mathcal{T}}_5\widetilde{\mathcal{X}}_5$. Quantile (left) and distribution regression (right).



(A) Food.



(B) Leisure.

FIGURE 2.4. QSF over $\widetilde{\mathcal{T}}_5\widetilde{\mathcal{X}}_5^*$. Quantile (left) and distribution regression (right).

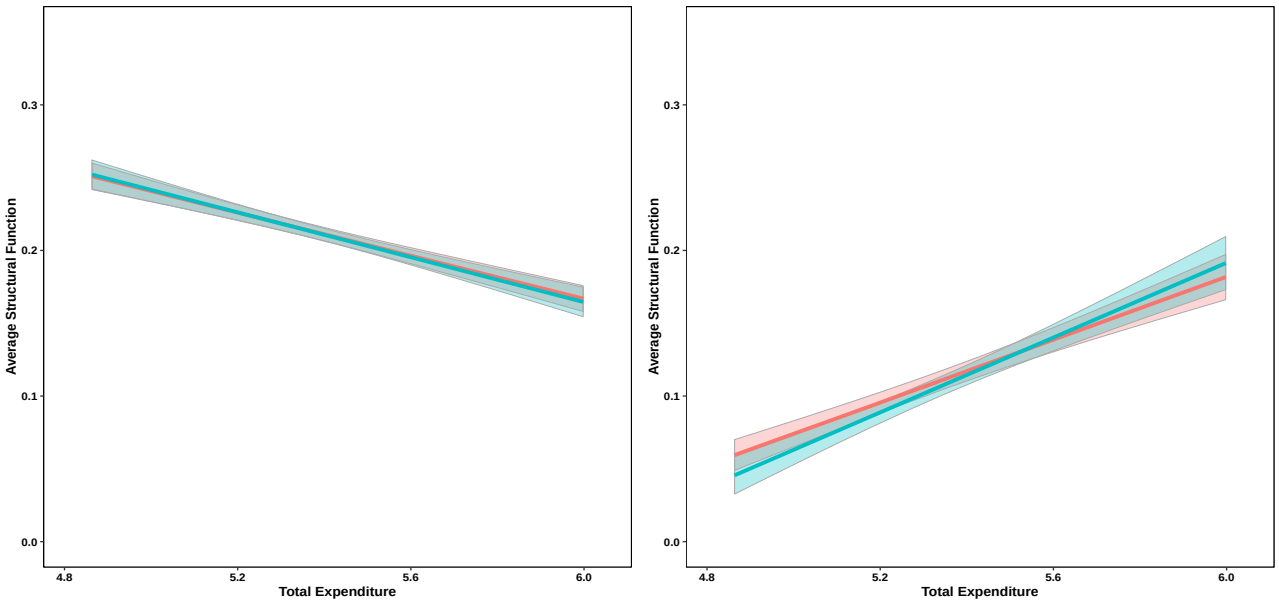


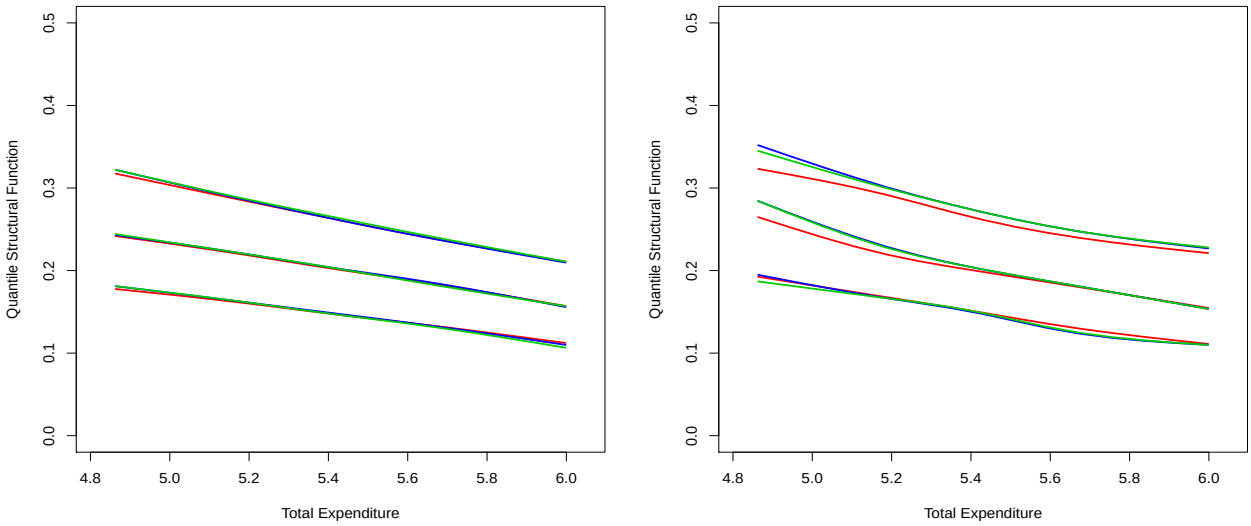
FIGURE 2.5. Comparison of ASF estimates. Food (left) and leisure (right); QR (red) and OLS (blue).

2.2. **Flexible Specifications.** Our estimators can also easily accommodate additional powers of the control function $\Phi^{-1}(V)^k$ as well as interaction terms $X \cdot \Phi^{-1}(V)^k$, $Z_1 \cdot \Phi^{-1}(V)^k$, and $X \cdot Z_1 \cdot \Phi^{-1}(V)^k$. We consider augmenting our baseline specifications by adding quadratic and both quadratic and cubic transformations of the control function and associated interactions terms. The control function and its powers are interacted with total expenditure (X), the children variable (Z_1), and their interaction $(X \cdot Z_1)$.

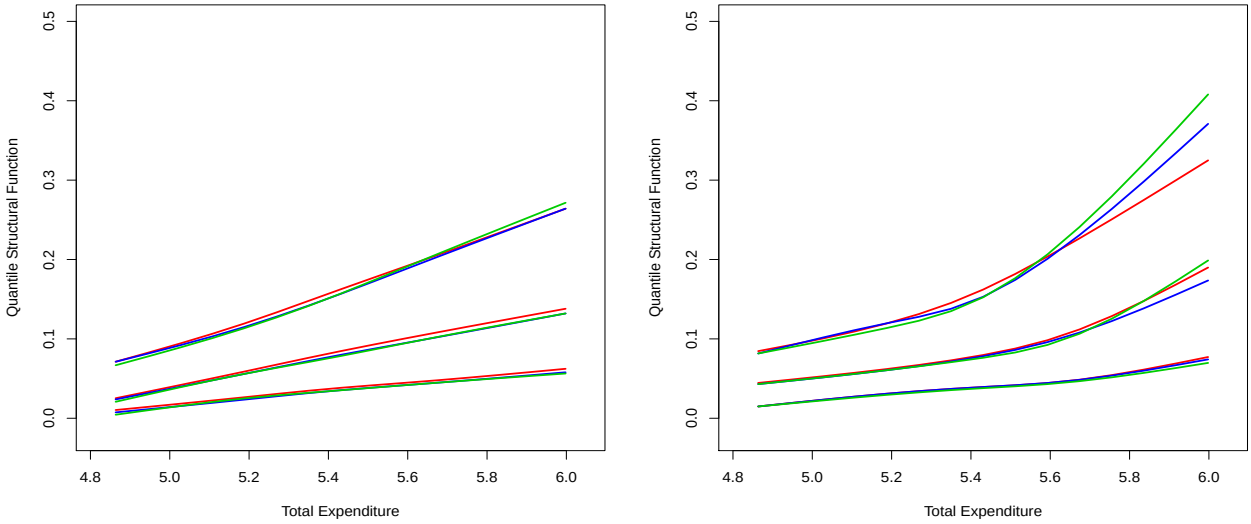
In Figure 2.6, we display the corresponding QSF for $k = 2, 3$, for both methods. The main difference with the baseline specifications is the increased curvature in the DR-based 0.75-QSFs in the right panel of Figures 2.6(A)-(B). Augmenting the model provides further evidence that our QSF estimates are robust are robust to the inclusion of higher-order terms in the control function.

To investigate further the selection of higher-order control function terms, we implement a leave-one-out cross-validation (CV) procedure for the ASF. For conditional mean specification of the QR baseline model, powers of the control function $\Phi^{-1}(V)^k$ as well as interaction terms $X \cdot \Phi^{-1}(V)^k$, $Z_1 \cdot \Phi^{-1}(V)^k$, and $X \cdot Z_1 \cdot \Phi^{-1}(V)^k$, $k = 2, \ldots, 5$, are added in increasing order to the specification of $E[Y \mid X, Z_1, V]$. Figure 2.7 displays the CV criterion values for each specification. The results confirm that adding additional powers of the control function does not improve the model fit markedly for the quantile regression specification of the ASF, for both food and leisure.

The extensions considered in this section further illustrate the complementarity of our estimation methods and their relevance for empirical work.



(A) Food.



(B) Leisure.

FIGURE 2.6. QSF including additional powers of the control function. Baseline (red): $\{p(X) \otimes r_1(Z_1)\} \cdot \Phi^{-1}(V)$; Quadratic specification (blue): baseline spec. + $\{p(X) \otimes r_1(Z_1)\} \cdot \Phi^{-1}(V)^2$; cubic specification (green): quadratic spec. + $\{p(X) \otimes r_1(Z_1)\} \cdot \Phi^{-1}(V)^3$. Quantile (left) and distribution (right) regression.

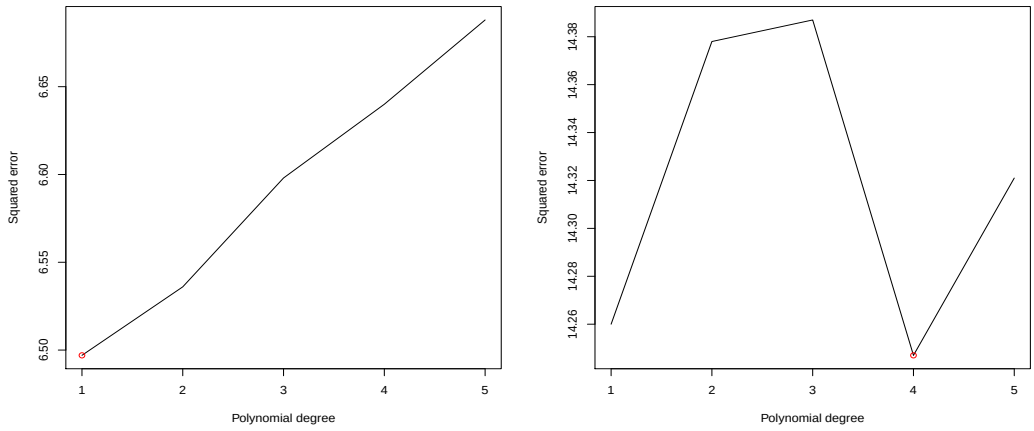


FIGURE 2.7. Least-squares CV criterion $\times 1000$. Food (left) and leisure (right).

2.3. **Robustness to Integration Method.** In this section, we compare the QSF obtained by replacing integration over the control function by sample averaging in the construction of the DSF by integration based on simulation of the control function. This can be done since the control function has a known distribution. Let $\widetilde{\mathcal{Y}}_S$, $\widetilde{\mathcal{X}}_K$ and $\widetilde{\mathcal{T}}_N$ be defined as in Algorithm 1 in the main text. The next algorithm describes estimation of structural functions with control function integration by simulation.

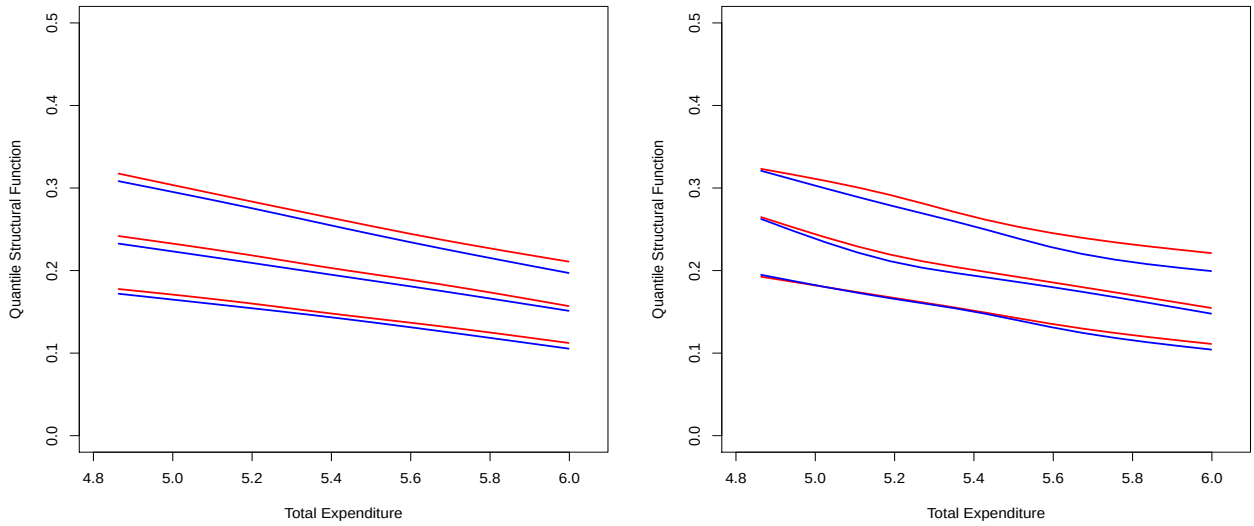
Algorithm 1 ***Estimation of Structural Functions - Integration by Simulation.***

First and Second Stages. *Repeat the first and second stages in Algorithm 1.*
Third Stage. *[DSF via integration by simulation]*

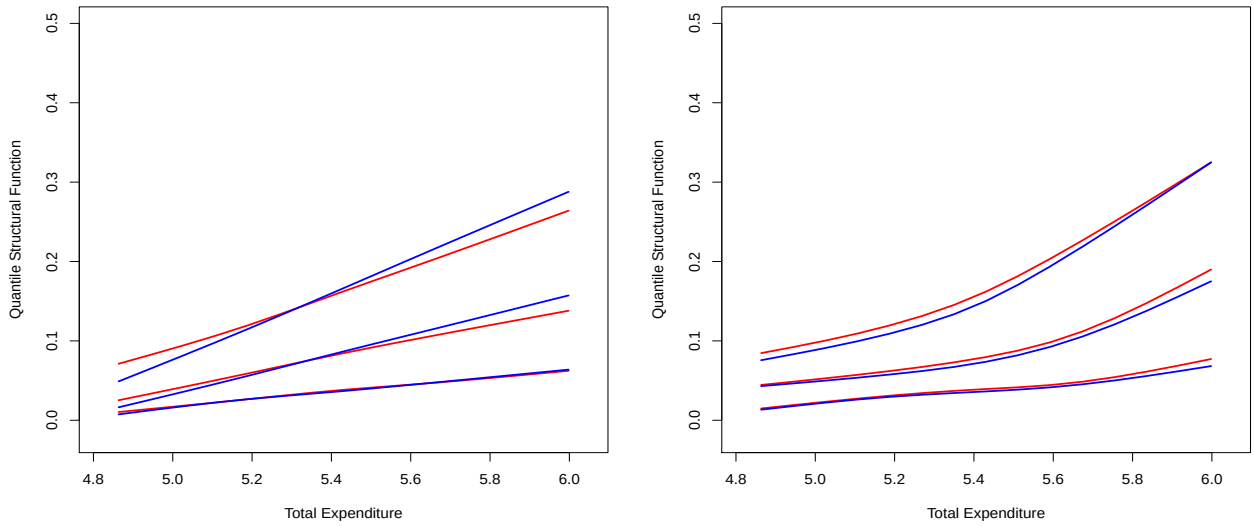
(1) *Draw n realizations $\{\check{V}_i\}_{i=1}^n$ from a Uniform $[0,1]$.*
(2) *For the DSF, set, for $(y,x) \in \widetilde{\mathcal{Y}}_S \widetilde{\mathcal{X}}_K$, $\widehat{G}(y,x) = \sum_{i=1}^n \widehat{F}_Y(y \mid x, \check{V}_i) T_i/n$. For the ASF and QSF, set, for $(\tau,x) \in \widetilde{\mathcal{T}}_N \widetilde{\mathcal{X}}_K$,*

$$\widehat{Q}_S(\tau,x) = \delta \sum_{s=1}^{S} \left[1(y_s \geq 0) - 1\{\widehat{G}(y_s,x) \geq \tau\} \right], \quad \widehat{\mu}_S(x) = \delta \sum_{s=1}^{S} \left[1(y_s \geq 0) - \widehat{G}(y_s,x) \right].$$

Figure 2.8 compares QSF estimates for QR and DR methods, obtained by implementing sample averaging (red) and integration by simulation (blue), for both food and leisure. The obtained ASFs are very similar, with slight differences for the DR based 0.75-QSF for food and the QR based 0.5, 0.75-QSF for leisure which are slightly



(A) Food.



(B) Leisure.

FIGURE 2.8. QSF by sample averaging (red) and integration by simulation (blue). Quantile (left) and distribution regression (right).

steeper. Overall, the shape of the estimated QSF is essentially the same for both integration methods.

3. Numerical Simulations

3.1. **Simulation Results.** To assess the performance of our estimators we implement Monte Carlo experiments based on three different designs, calibrated to the leisure empirical application. The first two experiments are based on Gaussian location-scale and DR triangular models, designed to reflect the respective strengths of the QR and DR estimators. The third experiment is a location triangular model, for which both estimators are consistent for the corresponding structural functions.

Design QR. Our first design is the linear location-scale shift system of equations

$$\begin{aligned} X &= \pi_{11} + \pi_{21} Z + (\pi_{12} + \pi_{22} Z)\eta, \\ Y &= \theta_{11} + \theta_{21} X + (\theta_{12} + \theta_{22} X)\varepsilon. \end{aligned}$$

The ASF and QSF of this model are linear,

$$\mu(x) = \theta_{11} + \theta_{21} x, \quad Q(\tau, x) = \theta_{11} + \theta_{21} x + (\theta_{12} + \theta_{22} x)\Phi^{-1}(\tau).$$

Design DR. Our second design is the nonlinear location-scale shift system of equations

$$\begin{aligned} X &= -\left(\frac{\pi_{11} + \pi_{12} Z}{\pi_{21} + \pi_{22} Z}\right) + \left(\frac{1}{\pi_{21} + \pi_{22} Z}\right)\eta, \\ Y &= -\left(\frac{\theta_{11} + \theta_{12} x}{\theta_{21} + \theta_{22} x}\right) + \left(\frac{1}{\theta_{21} + \theta_{22} x}\right)\varepsilon. \end{aligned}$$

The ASF and QSF of this model are nonlinear,

$$\mu(x) = -\left(\frac{\theta_{11} + \theta_{12} x}{\theta_{21} + \theta_{22} x}\right), \quad Q(\tau, x) = -\left(\frac{\theta_{11} + \theta_{12} x}{\theta_{21} + \theta_{22} x}\right) + \left(\frac{1}{\theta_{21} + \theta_{22} x}\right)\Phi^{-1}(\tau).$$

Design LOC. Our third design is the linear location shift system of equations

$$\begin{aligned} X &= \pi_{11} + \pi_{21} Z + \sigma_\eta \eta, \\ Y &= \theta_{11} + \theta_{21} X + \sigma_\varepsilon \varepsilon, \end{aligned}$$

for which the QR and DR models are correctly specified. The ASF and QSF of this model are

$$\mu(x) = \theta_{11} + \theta_{21} x, \quad Q(\tau, x) = \theta_{11} + \theta_{21} x + \sigma_\varepsilon \Phi^{-1}(\tau).$$

For all three experiments, the sample size is set to $n = 1655$, the number of observations in the empirical application, and 500 simulations are performed. For the regions of interest, we use the same $\mathcal{T}_3$ and $\overline{\mathcal{X}}_5$ as in the empirical application. We

Design	QR			LOC			DR		
				$\rho = -0.2$					
	L_1	L_2	L_∞	L_1	L_2	L_∞	L_1	L_2	L_∞
DR	6.9	8.1	15.2	5.2	6.7	8.3	5.9	7.4	10.2
QR	2.7	3.4	3.5	4.7	6.0	6.9	8.2	9.7	15.4
Ratio×100	251.1	237.2	426.8	110.3	111.9	121.2	72.4	76.6	66.2
OLS	10.6	10.9	22.4	14.9	15.4	26.0	15.4	16.0	33.0
				$\rho = -0.9$					
	L_1	L_1	L_∞	L_1	L_2	L_∞	L_1	L_2	L_∞
DR	4.7	6.0	7.7	6.4	7.9	10.5	7.8	9.5	13.6
QR	3.8	4.5	9.4	4.9	6.0	7.3	9.2	10.5	24.4
Ratio×100	123.6	132.7	82.0	131.1	131.9	144.7	84.4	90.4	56.0
OLS	47.2	47.3	100.2	66.2	66.3	117.9	73.2	73.3	152.8

TABLE 1. Average L^p estimation errors of ASF $\times 1000$ for the DR and QR estimators and their ratio $\times 100$, for $p = 1, 2$ and ∞. Average L^p estimation errors of ASF $\times 1000$ for OLS are included as a benchmark.

let (η, ε) be jointly normal scalar random variables with zero means, unit variances and correlation ρ, and assess the performance of our estimators under two different levels of endogeneity by setting $\rho = -0.2$, for low endogeneity, and $\rho = -0.9$, for extreme endogeneity. Accordingly, the DR estimator is implemented with the probit link function. We report simulation results for the ASF and the QSF.

Table 1 reports a first set of results regarding the accuracy of ASF estimates by DR and QR. For comparison purposes, Table 1 also includes ASF estimates by ordinary least-squares (OLS), providing a benchmark with no correction for endogeneity. We report average estimation errors across simulations of QR and DR estimators, and their ratio in percentage terms. Estimation errors are measured in L^p norms $\|\cdot\|_p$, $p = 1, 2$, and ∞, where for a function $f : \mathcal{X} \mapsto \mathbb{R}$, $\|f\|_p = \left\{\int_{\mathbb{R}} |f(s)|^p \, ds\right\}^{1/p}$, and are then averaged over the 500 simulations.

For this design, DR and QR-based estimators both perform very well and significantly improve over the OLS benchmark, including for $\rho = -0.2$. As expected, the accuracy of the estimates obtained by each method dominates for the corresponding design. For the QR design, the ratio of average estimation errors ranges from 82 to 426.8. Interestingly, the relative accuracy of DR-based estimates for $\rho = -0.9$ is close to the accuracy of QR estimates, with the ratio of average estimation errors ranging from

82 to 132.7, across norms; this feature is specific to the ASF and does not apply to the QSF. For the DR design, the ratio of average estimation errors ranges from 56 to 90.4. The larger reduction in average errors in L^∞ norm reflects the higher accuracy in estimation of extreme parts of the support where the ASF displays some curvature. Finally, for the LOC design, the performance of both methods is very similar for $\rho = -0.2$, and the QR-based estimator dominates more markedly for $\rho = -0.9$.

Tables 2 report the results regarding the accuracy of DR and QR estimates of the QSF, for quantile levels 0.25, 0.5 and 0.75, respectively. Compared to the results for the ASF, the main feature of the results is the stronger relative performance of the QR-based estimator for all three designs, although the DR-based estimator still dominates for the DR design.

Overall, the simulations show that both DR- and QR-based estimation methods perform well for their respective designs, and yield substantial correction for endogeneity. QR-based estimation dominates for both the QR and LOC designs, but the DR estimator is able to correct for endogeneity in data generating processes displaying nonlinearities in the structural functions. These simulation results illustrate further the complementarity of the two estimation methods introduced in this paper.

3.2. **Calibration.** In this section we give a detailed description of how the three data generating processes used for the Monte Carlo were calibrated to our empirical application.

3.2.1. *Linear Gaussian Location-Scale Model (QR specification).* Consider the heteroscedastic normal system of equations

$$\begin{aligned} X &= \pi_{11} + \pi_{21} Z + (\pi_{12} + \pi_{22} Z)\eta \\ Y &= \theta_{11} + \theta_{21} X + (\theta_{12} + \theta_{22} X)\varepsilon, \end{aligned}$$

where (ξ, ε) are jointly normal with zero means, unit variances and correlation ρ. The reduced form of this system is:

$$\begin{aligned} Q_X(v \mid z) &= \pi_{11} + \pi_{21} z + (\pi_{12} + \pi_{22} z)\Phi^{-1}(v) \\ Q_Y(u \mid x, v) &= \theta_{11} + \theta_{21} x + (\theta_{12} + \theta_{22} x)(\rho\Phi^{-1}(v) + (1-\rho^2)^{1/2}\Phi^{-1}(u)). \end{aligned}$$

This system thus admits the QR representation

$$\begin{aligned} Q_X(v \mid z) &= \pi_1(v) + \pi_2(v) z \\ Q_Y(u \mid x, v) &= \theta_1(u) + \gamma_1(u)\Phi^{-1}(v) + \theta_2(u) x + \gamma_2(u)\Phi^{-1}(v) x, \end{aligned}$$

	Design	QR			LOC			DR		
					$\tau = 0.25$					
		L_1	L_2	L_∞	L_1	L_2	L_∞	L_1	L_2	L_∞
	DR	10.4	11.5	28.2	7.0	8.9	12.6	7.9	9.9	12.5
$\rho = -0.2$	QR	3.4	4.3	5.6	6.0	7.5	8.6	9.1	11.2	15.9
	Ratio×100	302.0	266.1	506.6	117.4	118.3	145.6	86.3	88.7	78.6
	DR	5.7	6.8	11.5	7.4	9.4	15.3	7.7	9.6	14.0
$\rho = -0.9$	QR	2.6	3.3	3.2	5.3	6.7	9.0	8.8	10.6	18.6
	Ratio×100	218.4	207.5	355.9	139.7	140.1	170.1	87.2	91.0	75.2
					$\tau = 0.50$					
		L_1	L_1	L_∞	L_1	L_2	L_∞	L_1	L_2	L_∞
	DR	7.9	9.3	20.0	6.5	8.3	9.9	7.2	9.2	13.2
$\rho = -0.2$	QR	2.7	3.5	4.0	5.6	7.1	7.8	9.4	11.1	16.5
	Ratio×100	289.2	267.8	497.5	116.4	117.1	127.6	77.0	83.0	80.1
	DR	6.1	7.3	17.0	5.4	6.8	7.9	6.2	7.8	10.8
$\rho = -0.9$	QR	2.4	3.0	3.5	4.6	5.7	6.5	8.3	9.6	18.5
	Ratio×100	256.4	246.1	491.6	118.5	118.9	120.9	74.9	81.1	58.7
					$\tau = 0.75$					
		L_1	L_1	L_∞	L_1	L_2	L_∞	L_1	L_2	L_∞
	DR	10.0	11.6	16.5	7.3	9.3	14.3	7.6	9.5	15.7
$\rho = -0.2$	QR	3.4	4.3	5.0	6.1	7.6	8.6	10.3	11.9	21.1
	Ratio×100	297.0	271.9	330.7	119.6	122.1	165.8	73.6	79.5	74.6
	DR	8.1	10.0	20.9	7.0	8.9	15.4	9.2	11.5	24.8
$\rho = -0.9$	QR	2.9	3.6	4.6	5.3	6.6	9.1	10.3	11.7	29.3
	Ratio×100	279.3	276.7	457.0	131.6	135.4	168.7	89.5	98.7	84.6

TABLE 2. Average L^p estimation errors of $\{0.25, 0.5, 0.75\}$-QSF $\times 1000$ for the DR and QR estimators and their ratio $\times 100$, for $p = 1, 2$ and ∞.

with

$$\begin{aligned}
\pi_1(v) &= \pi_{11} + \pi_{12}\Phi^{-1}(v) \\
\pi_2(v) &= \pi_{21} + \pi_{22}\Phi^{-1}(v) \\
\theta_1(u) &= \theta_{11} + \theta_{12}(1-\rho^2)^{1/2}\Phi^{-1}(u) \\
\theta_2(u) &= \theta_{21} + \theta_{22}(1-\rho^2)^{1/2}\Phi^{-1}(u) \\
\gamma_1(u) &= \theta_{12}\rho \\
\gamma_2(u) &= \theta_{22}\rho.
\end{aligned}$$

Define the fine meshes of M values $0.01 = v_1 < \cdots < v_M = 0.99$ and $0.01 = u_1 < \cdots < u_M = 0.99$, with $M = 599$, as in the empirical application. The vectors of parameter values are calibrated following the method suggested in Koenker and Xiao (2002).

For the first stage parameters, we estimate the QR coefficients $(\widehat{\pi}_1(v), \widehat{\pi}_2(v))$ for quantile indexes $v \in \{v_1, \ldots, v_M\}$. The value of the coefficients $\pi_1 = (\pi_{11}, \pi_{12})'$ and $\pi_2 = (\pi_{21}, \pi_{22})'$ are then set to the estimates obtained from linear regression of $(\widehat{\pi}_1(v_m), \widehat{\pi}_2(v_m))$ on $(1, \Phi^{-1}(v_m))$.

For the second stage, we estimate the QR coefficients $(\widehat{\theta}_1(u), \widehat{\theta}_2(u))$ for quantile indexes $u \in \{u_1, \ldots, u_M\}$. The value of the coefficients $\theta_1 = (\theta_{11}, \theta_{12})'$ and $\theta_2 = (\theta_{21}, \theta_{22})'$ are then set to the estimates obtained from linear regression of $(\widehat{\theta}_1(u_m), \widehat{\theta}_2(u_m))$ on $(1, (1 - \widehat{\rho}^2)^{1/2}\Phi^{-1}(u_m))$. The correlation coefficient $\widehat{\rho}$ is calibrated to the correlation between $\widehat{\varepsilon}_i = \Phi^{-1}(\widehat{G}(Y_i, X_i))$ and $\widehat{\eta} = \Phi^{-1}(\widehat{F}_X(X_i \mid Z_i))$, where $\widehat{G}(Y_i, X_i)$ and $\widehat{F}_X(X_i \mid Z_i)$ are the DSF and F_X QR-based estimates evaluated at the n sample points. We find $\hat{\rho} \approx -0.1$.

3.2.2. *Nonlinear Gaussian Location-Scale Model (DR specification).* Consider the linear DR system of equations:

$$\begin{aligned} \eta &= (\pi_{11} + \pi_{21}X) + (\pi_{12} + \pi_{22}X)Z \\ \varepsilon &= (\theta_{11} + \theta_{21}Y) + (\theta_{12} + \theta_{22}Y)X, \end{aligned}$$

where (η, ε) are jointly normal with zero means, unit variances and correlation ρ. The reduced form of this system is

$$\begin{aligned} F_X(x \mid z) &= \Phi\left((\pi_{11} + \pi_{21}x) + (\pi_{12} + \pi_{22}x)z\right) \\ F_Y(y \mid x, v) &= \Phi\left(\frac{1}{(1-\rho^2)^{1/2}}(\theta_{11} + \theta_{21}y) + \frac{1}{(1-\rho^2)^{1/2}}(\theta_{12} + \theta_{22}y)x - \frac{\rho}{(1-\rho^2)^{1/2}}\Phi^{-1}(v)\right). \end{aligned}$$

This system thus admits the Gaussian DR representation

$$\begin{aligned} F_X(x \mid z) &= \Phi(\pi_1(x) + \pi_2(x)z) \\ F_Y(y \mid x, v) &= \Phi(\theta_1(y) + \gamma_1(y)\Phi^{-1}(v) + \theta_2(y)x) \end{aligned}$$

with

$$\begin{aligned}
\pi_1(x) &= \pi_{11} + \pi_{12}x \\
\pi_2(x) &= \pi_{21} + \pi_{22}x \\
\theta_1(y) &= \frac{1}{(1-\rho^2)^{1/2}}[\theta_{11} + \theta_{12}y] \\
\theta_2(y) &= \frac{1}{(1-\rho^2)^{1/2}}[\theta_{21} + \theta_{22}y] \\
\gamma_1(y) &= -\frac{\rho}{(1-\rho^2)^{1/2}}.
\end{aligned}$$

Define the fine mesh of M values $0.01 = t_1 < \cdots < t_M = 0.99$, with $M = 599$, as in the empirical application. The vectors of parameter values are calibrated by implementing the DR analog to the method suggested in Koenker and Xiao (2002).

For the first stage parameters, we estimate the DR coefficients $(\widehat{\pi}_1(x), \widehat{\pi}_2(x))$ for a fine mesh of X values $x \in \{\widehat{Q}_X(t_1), \ldots, \widehat{Q}_X(t_M)\}$, setting the link function to the gaussian CDF. The value of the coefficients $\pi_1 = (\pi_{11}, \pi_{12})'$ and $\pi_2 = (\pi_{21}, \pi_{22})'$ are then set to the estimates obtained from linear regression of $(\widehat{\pi}_1(\widehat{Q}_X(t_m)), \widehat{\pi}_2(\widehat{Q}_X(t_m)))$ on $(1, Q_X(t_m))$.

For the second stage, we estimate the DR coefficients $(\widehat{\theta}_1(y), \widehat{\theta}_2(y))$ for the fine mesh of Y values $y \in \{\widehat{Q}_Y(t_1), \ldots, \widehat{Q}_Y(t_M)\}$, setting the link function to the gaussian CDF. The value of the coefficients $\theta_1 = (\theta_{11}, \theta_{12})'$ and $\theta_2 = (\theta_{21}, \theta_{22})'$ are then set to the estimates obtained from linear regression of $(\widehat{\theta}_1(\widehat{Q}_Y(t_m)), \widehat{\theta}_2(\widehat{Q}_Y(t_m)))$ on $((1-\widehat{\rho}^2)^{-1/2}, (1-\widehat{\rho}^2)^{-1/2}\widehat{Q}_Y(t_m))$. The correlation coefficient $\widehat{\rho}$ is calibrated to the correlation between $\widehat{\varepsilon}_i = \Phi^{-1}(\widehat{G}(Y_i, X_i))$ and $\widehat{\eta} = \Phi^{-1}(\widehat{F}_X(X_i \mid Z_i))$, where $\widehat{G}(Y_i, X_i)$ and $\widehat{F}_X(X_i \mid Z_i)$ are the DSF and F_X DR-based estimates evaluated at the n sample points. We find $\hat{\rho} \approx -0.1$.

In practice, the corresponding ASF and QSFs calibrated to our data are very close to being linear over the range of values of X considered, as can be seen from Figure 3.1 which displays the true $\{0.25, 0.5, 0.75\}$-QSFs generated by our calibration (black lines). This yields simulation results favorable to our QR-based estimators. In order to assess our methods for a data generating process which is less favorable to the QR specification, the parameter values for θ_{12} and θ_{22} used in the simulations are set to the parameter values from our initial calibration multiplied by 1.5 and 2.25, respectively. This modification generates some curvature in the QSF as shown in Figure 3.1 which also displays the true $\{0.25, 0.5, 0.75\}$-QSFs generated by our adjusted calibration (red lines).

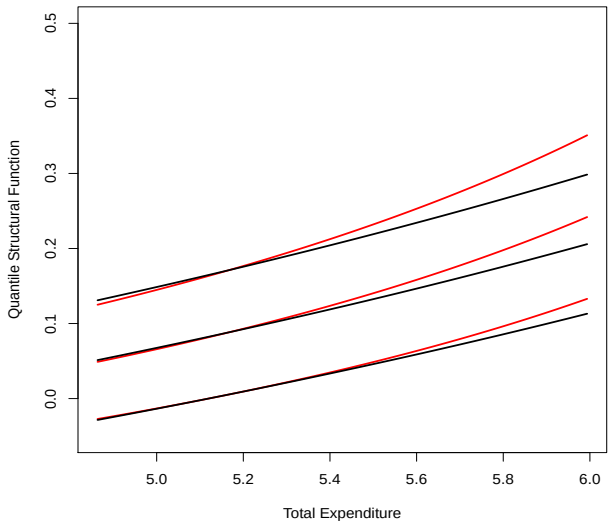


FIGURE 3.1. True QSFs for DR Calibration. Initial calibration (black) and modified calibration used in simulations (red).

3.2.3. *Linear Gaussian Location Model (LOC specification).* The QR and DR specifications coincide for the location shift model

$$\begin{aligned} X &= \pi_{11} + \pi_{21} Z + \sigma_\eta \eta \\ Y &= \theta_{11} + \theta_{21} X + \sigma_\varepsilon \varepsilon. \end{aligned}$$

For this model, both the conditional quantile and distribution functions are linear in the covariate, and so are the structural functions. After substitution, the reduced-form equation for the second stage is

$$\begin{aligned} Y &= \theta_{11} + \theta_{21} X + [\rho\eta + (1-\rho^2)^{1/2}\xi] \\ &= \theta_{11} + \theta_{21} X + \sigma_\varepsilon \rho \eta + \sigma_\varepsilon (1-\rho^2)^{1/2} \xi. \end{aligned}$$

The value of the coefficients $(\pi_{11}, \pi_{21})'$ are then set to the estimates obtained from linear regression of X_i on $(1, Z_i)$, and the scale parameter σ_η is set to the corresponding estimate $\widehat{\sigma}_\eta = [(n-1)^{-1} \sum_{i=1}^n (X_i - \widehat{\pi}_{11} - \widehat{\pi}_{21} Z_i)^2]^{1/2}$. Letting $\widehat{\eta}_i = (X_i - \widehat{\pi}_{11} - \widehat{\pi}_{21} Z_i)/\widehat{\sigma}_\eta$, the value of the coefficients $(\theta_{11}, \theta_{21})'$ are then set to the estimates of the first two coefficients in the linear regression of Y_i on $(1, X_i, \widehat{\eta}_i)$. We then set the scale parameter σ_ε to the corresponding estimate $\widehat{\sigma}_\varepsilon = [(n-1)^{-1} \sum_{i=1}^n (Y_i - \widehat{\theta}_{11} - \widehat{\theta}_{21} X_i)^2]^{1/2}$. The correlation coefficient $\widehat{\rho}$ is calibrated to the correlation between $\widehat{\varepsilon}_i = (Y_i - \widehat{\theta}_{11} - \widehat{\theta}_{21} X_i)/\widehat{\sigma}_\varepsilon$ and $\widehat{\eta}$. We find $\hat{\rho} \approx -0.2$.