Nonparametric Identification and Estimation with Non-Classical Errors-in-Variables

Kirill S. Evdokimov* Andrei Zeleneev†‡

This version: May 10, 2023

Abstract

This paper considers nonparametric identification and estimation of the regression function when a covariate is mismeasured. The measurement error need not be classical. Employing the small measurement error approximation, we establish nonparametric identification under weak and easy-to-interpret conditions on the instrumental variable. The paper also provides nonparametric estimators of the regression function and derives their rates of convergence.

*Universitat Pompeu Fabra and Barcelona School of Economics: kirill.evdokimov@upf.edu.

†University College London: a.zeleneev@ucl.ac.uk.

‡Evdokimov gratefully acknowledges the support from the Spanish MCIN/AEI via grants RYC2020-030623-I, PID2019-107352GB-I00, and Severo Ochoa Programme CEX2019-000915-S.

1 Introduction

Regression is a key tool for empirical analysis. Errors-in-Variables (EIV) are a widespread problem in many applications. Mismeasurement of a covariate, when not accounted for, may lead to biased estimates and invalid inferences.

The goal of this paper is to study the nonparametric identification and estimation of the regression function when a covariate is mismeasured. Importantly, the measurement error need not be classical and can be correlated with the mismeasured covariate. In this paper, we adopt the small measurement error approximation, which allows us to provide a simple nonparametric characterization of the problem. Then we provide transparent and constructive identification analysis under weak and easy-to-interpret conditions on the instrumental variable.

In Section 2, we focus on the Weakly Classical Measurement Error (WCME) model, where the measurement error is uncorrelated with the true covariate but generally is not independent from it. We show that the skedastic function of the measurement errors, together with its derivative, plays a key role in determining the bias of the naive regression estimator. We also show how the EIV skedastic function can be recovered from the distribution of the observables using a (possibly discrete) instrument, and how one can construct a bias-corrected estimator of the regression function. We derive its rate of convergence and provide conditions under which the approximation error becomes negligible compared to the errors arising from the nonparametric estimation of unknown functions in large samples.

In Section 3, we examine the general Non-Classical Measurement Error (NCME) model, which allows for a very broad form of EIV. In particular, the measurement error can be correlated with the true covariate. Despite the increased generality compared to the WCME model, we demonstrate how the results from Section 2 can be utilized to establish identification of the general NCME model.

Importantly, our approach only requires an instrumental variable, which can be discrete. This allows for a broader range of applications than methods requiring a continuously distributed instrument or the availability of repeated (multiple) measurements of the true covariate (since a repeated measurement can be used as an instrument). In Section 2, we discuss in detail the exclusion and relevance conditions that the instrumental variable needs to satisfy, and consider some examples.

Carroll, Ruppert, Stefanski, and Crainiceanu (2006); Chen, Hong, and Nekipelov

(2011); and Schennach (2020, 2022) provide excellent literature overviews.

Our main focus is on settings where the distribution of measurement error is unknown. Nonparametric analysis of the EIV problem in such settings requires additional information to separate the true covariate from the measurement error. In Economics, most commonly instrumental variables are used for this purpose (Hausman, Ichimura, Newey, and Powell, 1991; Hausman, Newey, and Powell, 1995; Newey, 2001; Schennach, 2007; Hu and Schennach, 2008, among others). Repeated measurements can also be utilized (Hausman et al., 1991; Li and Vuong, 1998; Schennach, 2004, among others) but are less frequently available. Note that repeated measurements can serve as valid instruments.

Stimulated by the seminal work of Wolter and Fuller (1982), the small measurement error (SME) approximation has been widely employed in Statistics and Econometrics to study the effect of EIV on various estimators and to bias-correct them (e.g., Carroll and Stefanski, 1990; Chesher, 1991, 2000; Carroll et al., 2006; Chesher and Schluter, 2002, among others). Bound, Brown, and Mathiowetz (2001) document that the EIV in applied work are typically relatively small although are often non-classical. This paper differs from the previous literature in two ways. First, it appears to be the first paper to study the nonparametric SME approximation with non-classical EIV. Second, with the exception of Evdokimov and Zeleneev (2022), the previously developed SME bias reduction techniques assume that the EIV variance is either known or can be directly estimated from the available data such as repeated measurements. In contrast, this paper demonstrates how the whole EIV skedastic function can be identified using a (possibly discrete) instrumental variable.

The analysis of this paper complements the existing "large" measurement error literature. By focusing on a narrower range of settings, the paper provides simpler characterizations of the problem and estimators, which are valid under very weak and easy-to-interpret conditions on the instrumental variable. In particular, our identification results do not rely on the completeness conditions, and estimation does not involve solving ill-posed inverse problems or deconvolution.

2 Weakly Classical Measurement Errors

We consider the regression model

$$\rho(x) \equiv E[Y_i | X_i^* = x], \tag{1}$$

where $Y_i \in \mathbb{R}$ is the outcome variable, and $X_i^* \in \mathbb{R}$ is the true value of the covariate for individual i. The researcher observed a mismeasured version of X_i^*:

$$X_i = X_i^* + \varepsilon_i.$$

where ε_i is the measurement error. The researcher has a random sample of (Y_i, X_i, Z_i), where Z_i are instrumental variables that are used to identify the model and will be discussed later. It is straightforward to also include correctly measured covariates into the model, see Remark 4 for details.

In this section we consider the Weakly Classical Measurement Error (WCME) model:

Assumption WCME. *$X_i = X_i^* + \varepsilon_i$ and $E[\varepsilon_i | X_i^*] = 0$.*

The measurement error ε_i is uncorrelated with the true covariate X_i^*. Assumption WCME is significantly weaker than the (Strongly) Classical Measurement Error (CME) assumption, since ε_i need not be independent from X_i^*. For example, the measurement error can be conditionally heteroskedastic, i.e., its conditional variance

$$v(x) \equiv V[\varepsilon_i | X_i^* = x].$$

may depend on x. Function $v(x^*)$ is usually unknown.

Example (WCME-LIN-RC)**.** *Suppose $X_i = \psi_{i1} + \psi_{i2} X_i^*$, where $(\psi_{i1}, \psi_{i2}) \perp X_i^*$. Assumption WCME is satisfied if $E[(\psi_{i1}, \psi_{i2})] = (0, 1)$. Here $v(x) = \sigma_{\psi_1}^2 + \sigma_{\psi_2}^2 x^2 + 2\sigma_{\psi_1\psi_2} x$.*

In this paper we use the Small Measurement Error (SME) approximation (e.g., Wolter and Fuller, 1982) for the analysis, i.e., we will consider the approximations of the model when $v(x^*)$ and the higher conditional moments of ε_i are small.

Specifically, we model the measurement error as $\varepsilon_i = \tau \xi_i$ where the distribution of ξ_i is fixed, and τ is a non-stochastic parameter. Assumption WCME requires

$E[\xi_i|X_i^*] = 0$. The conditional variance of ε_i is given by $v(x) = \tau^2 V[\xi_i|X_i^* = x] = O(\tau^2)$.

We study the properties of the model when $\tau \to 0$. Under some smoothness conditions,

$$E\left[Y_i|X_i = x\right] = \rho\left(x\right) + O\left(\tau^2\right).$$

Thus, a naive regression estimator of ρ that ignores the presence of the measurement errors in X_i has a bias of order $O\left(\tau^2\right)$, e.g., see Chesher (1991).

The goal of the small measurement error analysis is to provide a function $\widetilde{\rho}\left(x\right)$ that has a smaller bias, i.e., satisfies

$$\widetilde{\rho}\left(x\right) = \rho\left(x\right) + O\left(\tau^p\right), \tag{2}$$

for some $p \geq 3$.

To identify the model we will rely on an observed instrumental variable (instrument) Z_i that satisfies the following exogeneity assumption.

Assumption 2.1. $E\left[Y_i|X_i^*, Z_i\right] = E\left[Y_i|X_i^*\right]$.

This assumption states that Z_i is an "excluded" variable: given X_i^*, instrument Z_i has no effect on the conditional mean of Y_i. Without loss of generality, we can assume that Z_i is discrete. (The instrument also needs to satisfy a "relevance"condition: it needs to affect the conditional distribution $f_{X^*|Z}\left(x|z\right)$. This condition will appear in Theorem 1.)

Assumption 2.2. $E\left[Y_i|X_i^*, Z_i, X_i\right] = E\left[Y_i|X_i^*, Z_i\right]$.

Assumption 2.2 says that the measurement error ε_i is nondifferential: conditional on (X_i^*, Z_i), X_i provides no additional information about (the conditional mean of) Y_i. This assumption can be equivalently stated as $E\left[Y_i|X_i^*, Z_i, \varepsilon_i\right] = E\left[Y_i|X_i^*, Z_i\right] = E\left[Y_i|X_i^*\right]$.

Assumption 2.3. *$\varepsilon_i = \tau\xi_i$, where $\tau \geq 0$ is non-random, $E[\xi_i|X_i^*] = 0$, and $f_{\xi|X^*Z}\left(u|x, z\right) = f_{\xi|X^*}\left(u|x\right)\ \forall u, x, z$.*

Assumption 2.3 and the smoothness conditions below are stated using the auxiliary variable ξ_i, whose variance does not shrink. This allows formulating the smoothness conditions in the conventional form. For example, we will assume that the density $f_{\xi|X^*}(u|x)$ is bounded. In contrast, the density of $\varepsilon_i = \tau\xi_i$ is $f_{\varepsilon|X^*}(e|x) = \frac{1}{\tau} f_{\xi|X^*}\left(\frac{e}{\tau}\middle| x\right)$ and is not bounded as $\tau \to 0$. Assumption 2.3 also implies that $\varepsilon_i \perp Z_i | X_i^*$.

Finally, the following two assumptions are smoothness conditions.

Assumption 2.4. *Function $\rho(x)$ and the conditional densities $f_{X^*|Z}(x|z)$ and $f_{\xi|X^*}(u|x)$ are bounded functions with $m \geq p$ bounded derivatives with respect to x, for some integer $p \geq 3$.*

Assumption 2.5. *$\int |u|^m \sup_{\tilde{x}\in\mathcal{S}_X} \left|\nabla_x^\ell f_{\xi|X^*}(u|\tilde{x})\right| du < \infty$ for $\ell \in \{0,\ldots,m\}$ for some closed convex set $\mathcal{S}_X \subseteq \mathbb{R}$ containing the supports of X_i^* and X_i.*

Assumption 2.5 is a weak restriction imposed on the conditional moments of ξ_i. Appendix B provides a set of primitive conditions that guarantee that Assumption 2.5 holds. Also notice that Assumption 2.5 would automatically hold if the support of ξ_i is bounded, since $f_{\xi|X^*}(u|x)$ and its derivatives are uniformly bounded under Assumption 2.4.

We can now state the first main result of the paper. Let

$$
\begin{aligned}
q(x,z) &\equiv E[Y_i|X_i = x, Z_i = z], \qquad q(x) \equiv E[Y_i|X_i = x], \\
s_{X|Z}(x|z) &\equiv \frac{f'_{X|Z}(x|z)}{f_{X|Z}(x|z)}, \qquad s_{X^*|Z}(x|z) \equiv \frac{f'_{X^*|Z}(x|z)}{f_{X^*|Z}(x|z)}, \\
\widetilde{v}(x) &\equiv \frac{q(x,z_1) - q(x,z_2)}{q'(x)\left[s_{X|Z}(x|z_1) - s_{X|Z}(x|z_2)\right]}, \qquad (3) \\
\widetilde{\rho}(x,z) &\equiv q(x,z) - \widetilde{v}(x)\left[q'(x)\, s_{X|Z}(x|z) + \tfrac{1}{2}q''(x)\right] - q'(x)\,\widetilde{v}'(x). \qquad (4)
\end{aligned}
$$

Let $\mathcal{S}_{X^*}(z)$ denote the conditional support of $X_i^*|Z_i = z$. Consider any two values z_1 and z_2 the instrument can take.

Theorem 1. *Suppose that Assumptions WCME and 2.1-2.5 are satisfied. Suppose either (i) $p = 3$, or (ii) $E\left[\xi_i^3|X_i^*\right] = 0$ and $p = 4$. Consider any point $x \in \mathcal{S}_{X^*}(z_1) \cap \mathcal{S}_{X^*}(z_2)$ such that*

$$\rho'(x)\left[s_{X^*|Z}(x|z_1) - s_{X^*|Z}(x|z_2)\right] \neq 0. \qquad (5)$$

Then, as $\tau \to 0$,

$$
\begin{aligned}
\widetilde{v}(x) &= v(x) + O(\tau^p), \quad \textit{and} \\
\widetilde{\rho}(x,z_1) &= \rho(x) + O(\tau^p).
\end{aligned}
$$

Theorem 1 demonstrates that $\widetilde{\rho}(x, z_1)$ identifies $\rho(x)$ up to an error of order $O(\tau^p)$ when $\tau \to 0$. This is a substantial improvement over naive regression $q(x)$ which has a bias of order $O(\tau^2)$. The improvement in the magnitude of the approximation error (from $O(\tau^2)$ to $O(\tau^4)$) is especially noticeable when $E[\xi_i^3|X_i^*] = 0$, e.g., when the measurement error is symmetric.

To establish the desired result, we first characterize the bias of $q(x, z)$ up to an error of order $O(\tau^p)$. The bias of $q(x, z)$ is of order $O(\tau^2)$ and determined by the conditional variance of the measurement error $v(x)$ and its derivative $v'(x)$, which are unknown. Then, we show that $\widetilde{v}(x)$ identifies $v(x)$ up to an error of order $O(\tau^p)$.[1] This allows us to approximate the bias of $q(x, z)$ with a sufficient precising using $\widetilde{v}(x)$ in place of $v(x)$. Finally, we construct $\widetilde{\rho}(x, z_1)$ by bias correcting $q(x, z_1)$ and demonstrate that it approximates the true regression function $\rho(x)$ up to an error of order $O(\tau^p)$.

The idea behind nonparametric identification is that although function $E\left[Y_i|X_i^* = x, Z_i = z\right]$ does not depend on z, function $q\left(x, z\right) \equiv E\left[Y_i|X_i = x, Z_i = z\right]$ does vary with z. The theorem shows how this variation allows recovering $v\left(x\right)$. Specifically, the proof of the Theorem shows that

$$q\left(x, z\right) = \rho\left(x\right) + v\left(x\right)\rho'\left(x\right)s_{X^*|Z}\left(x|z\right) + \frac{1}{2}v\left(x\right)\rho''\left(x\right) + \rho'\left(x\right)v'\left(x\right) + O\left(\tau^p\right). \quad (6)$$

Then it is shown that replacing the derivatives of ρ with those of q and $s_{X^*|Z}$ with $s_{X|Z}$ on the right-hand side in the above equation does not increase the magnitude of the approximation error, i.e., that

$$q\left(x, z\right) = \rho\left(x\right) + v\left(x\right)q'\left(x\right)s_{X|Z}\left(x|z\right) + \frac{1}{2}v\left(x\right)q''\left(x\right) + q'\left(x\right)v'\left(x\right) + O\left(\tau^p\right). \quad (7)$$

Note that only the second term on the right-hand side depends on z. Since q, q', and $s_{X|Z}$ are directly identified from the joint distribution of the observables, considering the differences $q\left(x, z_1\right) - q\left(x, z_2\right)$ then allows identification of $v\left(x\right)$ by $\widetilde{v}\left(x\right)$ up to an error of order $O\left(\tau^p\right)$. This identification approach requires the rank condition that $s_{X|Z}\left(x|z\right)$ depends on z, which is ensured by equation (5). In addition, it is necessary that $q'\left(x\right) \neq 0$, which is also ensured by equation (5). The latter condition is weak: $\rho'\left(x\right) = 0$ for all x only if $\rho\left(x\right)$ is a constant.[2]

[1] We also demonstrate that $\widetilde{v}'(x) = v'(x) + O(\tau^p)$. This is an important step of the proof.

[2] For an analysis of the role of the conditions such as $\rho'\left(x\right) \neq 0$ in the measurement error literature

Remark 1. *It is easy to check that* $\widetilde{\rho}(x, z_1) = \widetilde{\rho}(x, z_2)$.

Corollary 2 (Classical Measurement Error)**.** *Suppose the hypotheses of Theorem 1 hold, and the measurement error is classical, i.e.,* $\varepsilon_i \perp (X_i^*, Z_i, Y_i)$. *Suppose condition (5) holds for some point* $\dot{x}$. *Then for all* x *and* z *such that* $x \in \mathcal{S}_{\mathcal{X}^*}(z)$, *as* $\tau \to 0$,

$$\widetilde{\rho}_{CME}(x, z) = \rho(x) + O(\tau^p),$$

where

$$\widetilde{\rho}_{CME}(x, z) \equiv q(x, z) - \widetilde{v}(\dot{x})\left[q'(x)\, s_{X|Z}(x|z) + \tfrac{1}{2} q''(x)\right].$$

When the measurement error is classical, $v(x)$ is constant, and hence the term containing $v'(x)$ is absent from equation (6). In addition, Corollary 2 requires only a single point $\dot{x}$ satisfying the rank condition (5) for identification of $\rho(x)$ for all x, since $v(x) = v(\dot{x})$ for all x.[3] In contrast, in the general case of Theorem 1, nonparametric identification of $v(x)$ and $\rho(x)$ for a given x requires condition (5) to hold at that point x. Identification of the Classical Measurement Error model has been previously established in Evdokimov and Zeleneev (2022).

Remark 2 (Examples of IVs)**.** *First, variable* X_i^* *can be caused by* Z_i*; for example,* $X_i^* = q(Z_i, \eta_i)$ *for some unobserved (vector)* η_i *and function* q. *Assumptions 2.1 and 2.2 will be satisfied if* $E[U_i | Z_i, \eta_i, \varepsilon_i] = 0$.

Second, variable Z_i *can be caused by* X_i^**, for example be a second measurement or proxy for* X_i^**:* $Z_i = \chi(X_i^*, \nu_i)$. *For example,* Z_i *can be a second measurement:* $Z_i = \alpha_1 + \alpha_2 X_i^* + \nu_i$. *Assumptions 2.1 and 2.2 will be satisfied if* $E[U_i | X_i^*, \nu_i, \varepsilon_i] = 0$.

Remark 3. *If the skedastic function* $v(x)$ *is known, there is no need in having the instrumental variable* Z_i. *In this case one can use* $\widetilde{\rho}(x)$ *from equation (4) with* $q(x, z)$ *and* $\widetilde{v}(x)$ *replaced by* $q(x)$ *and* $v(x)$*, and the conclusion of Theorem 1 will continue to hold, i.e.,* $\widetilde{\rho}(x) = \rho(x) + O(\tau^p)$.

Remark 4. *It is straightforward to include additional correctly measured covariates* W_i *into the model, and to consider regression function* $\rho(x, w) \equiv E[Y_i | X_i^* = x, W_i = w]$. *The correctly measured covariates* W_i *play no special role, and all of the analysis can be thought of as applying conditionally on* $W_i = w$ *for*

see Evdokimov and Zeleneev (2018).

[3] For the result of Corollary 2 to hold, it is sufficient to require $v(x)$ to be constant, i.e., to assume that ε_i is homoskedastic, instead of requiring $\varepsilon_i \perp (X_i^*, Z_i, Y_i)$.

any given w, i.e., for the stratum with $W_i = w$. Thus, we omit W_i for simplicity of exposition.

Nonparametric Estimation Theorem 1 suggests an analogue estimator of $\widetilde{\rho}$ by replacing functions that appear in equations (3)-(4) with their standard nonparametric estimators (e.g., kernel or sieve), with optimally chosen tuning parameters. Let $\widehat{\rho}\,(x)$ denote this estimator. As an alternative to $\widehat{\rho}\,(x)$, we also consider $\widehat{\rho}^{\text{Naive}}\,(x)$, a naive nonparametric estimator of $\rho(x)$ ignoring the presence of the measurement error.

To approximate the finite sample properties of the studied estimators when τ is small, we consider a triangular asymptotic framework with drifting $\tau = \tau_n$ converging to zero as the sample size $n \to \infty$.

Lemma 3. *Suppose the hypotheses of Theorem 1 hold and Z_i is discrete. Also, suppose $\tau_n = o(1)$, then*

$$\begin{aligned} \widehat{\rho}\,(x) - \rho\,(x) &= O_p\left(n^{-\frac{m-1}{2m+1}} + \tau_n^p\right), \\ \widehat{\rho}^{Naive}\,(x) - \rho\,(x) &= O_p\left(n^{-\frac{m}{2m+1}} + \tau_n^2\right). \end{aligned}$$

Lemma 3 establishes the rates of convergence for the proposed and naive estimators. For each estimator, the rate of convergence is determined by two components: the standard nonparametric learning rate and the EIV (errors-in-variables) bias due to the presence of the measurement error.

The nonparametric learning rate for $\widehat{\rho}(x)$ is slower than for the naive estimator because it involves nonparametric estimation of derivatives such as $q'(x)$ and $s_{X|Z}(x|z)$. However, as Theorem 1 suggests, the EIV bias of the proposed estimator $\widehat{\rho}(x)$ is of order $O(\tau_n^p)$, whereas the naive estimator has a much larger bias of order $O(\tau_n^2)$. Thus, despite the slower nonparametric learning rate, $\widehat{\rho}(x)$ has a faster rate of convergence than $\widehat{\rho}^{\text{Naive}}\,(x)$ unless τ_n is very small, i.e., the measurement error is negligible.

To illustrate this result, suppose the conditions of Theorem 1(ii) hold, $m = p = 4$, and $\tau_n = O\left(n^{-\frac{1}{12}}\right)$. Then $\widehat{\rho}\,(x) - \rho\,(x) = O_p\left(n^{-\frac{1}{3}}\right)$, but the naive estimator has a much slower rate of convergence: $\widehat{\rho}^{\text{Naive}}\,(x) - \rho\,(x) = O_p\left(n^{-\frac{1}{6}}\right)$, because of the EIV bias.

3 General Non-Classical Measurement Error

In this section, we will use notation $\mathcal{X}_i^*$ for the true mismeasured covariate and consider the general measurement model

$$X_i = \boldsymbol{m}\left(\mathcal{X}_i^*, \psi_i\right), \tag{8}$$

where $\boldsymbol{m}$ is an unknown function and ψ_i is a random vector independent from $(Y_i, \mathcal{X}_i^*, Z_i)$. Function $\boldsymbol{m}$ need not be monotone in any of the arguments. The measurement error is non-classical: $X_i - \mathcal{X}_i^*$ and $\mathcal{X}_i^*$ are generally correlated.

As before, we want to identify and estimate the regression function

$$\rho_{\mathcal{X}^*}\left(\varkappa\right) \equiv E\left[Y_i | \mathcal{X}_i^* = \varkappa\right]. \tag{9}$$

We will assume that the measurement is sufficiently informative about the true $\mathcal{X}_i^*$. Define

$$\mu\left(\varkappa^*\right) \equiv E\left[X_i | \, \mathcal{X}_i^* = \varkappa^*\right].$$

Assumption MONOT-MEAS. $\mu\left(\varkappa^*\right)$ *is a strictly increasing function.*

Example (NCME-LIN-RC, continued)**.** *For* $X_i = \psi_{i1} + \psi_{i2}\mathcal{X}_i^*$*, we have* $\mu\left(\varkappa^*\right) = c_{\psi 1} + c_{\psi 2}\varkappa^*$*, and Assumption MONOT-MEAS is satisfied if* $c_{\psi 2} > 0$*. Note that* ψ_{i2} *is allowed to take negative values, which makes* $\boldsymbol{m}$ *a decreasing function of* $\mathcal{X}_i^*$ *for such observations.*

Note that since functions $\boldsymbol{m}$ and $\rho_{\mathcal{X}^*}$ are unrestricted, the model (8)-(9) cannot be identified without some normalization or additional information. Specifically, for any strictly increasing function λ, we can define an observationally equivalent model with $\tilde{\mathcal{X}}_i^* \equiv \lambda\left(\mathcal{X}_i^*\right)$, $\rho_{\tilde{\mathcal{X}}^*}\left(\tilde{\varkappa}^*\right) \equiv \rho_{\mathcal{X}^*}\left(\lambda^{-1}\left(\tilde{\varkappa}^*\right)\right)$ and $\boldsymbol{m}_{\tilde{\mathcal{X}}_i^*}\left(\tilde{\varkappa}^*, \psi\right) \equiv \boldsymbol{m}\left(\lambda^{-1}\left(\tilde{\varkappa}^*\right), \psi\right)$.

Let us *define* random variable

$$X_i^* \equiv \mu\left(\mathcal{X}_i^*\right). \tag{10}$$

Then,

$$E\left[X_i | X_i^*\right] = E\left[X_i | \mu\left(\mathcal{X}_i^*\right)\right] = E\left[X_i | \mathcal{X}_i^*\right] = \mu\left(\mathcal{X}_i^*\right) = X_i^*,$$

where the first equality follows by (10), the second equality follows from the strict monotonicity of $\mu(\cdot)$, the third equality is the definition of $\mu\left(\cdot\right)$, and the last equality follows from (10).

Thus, for the general measurement error model we can consider an observationally equivalent model that defines X_i^* as in equation (10):

$$\begin{aligned}\rho_{X^*}(x) &\equiv E[Y_i|X_i^*=x],\\ X_i &= X_i^*+\varepsilon_i, \qquad E[\varepsilon_i|X_i^*]=0.\end{aligned}$$

Since in this model Assumption WCME holds, we can apply the result of Theorem 1 to identify $\rho_{X^*}(x)$ and $v(x)\equiv E[\varepsilon_i^2|X_i^*=x]$ (up to an error of order $O(\tau^p)$) using $\widetilde{v}(x)$ and $\widetilde{\rho}(x)$ defined in equations (3) and (4), respectively. Notice that $\rho_{\mathcal{X}^*}(\varkappa)=\rho_{X^*}(\mu(\varkappa))$. However, since $\mathcal{X}_i^*$ is not observed, one cannot identify $\mu(\varkappa)$ and $\rho_{\mathcal{X}^*}(\varkappa)$ without some sadditional information.

Suppose for a moment that the marginal distribution $F_{\mathcal{X}^*}$ of $\mathcal{X}_i^*$ is known (for example, from a separate dataset, e.g., administrative records). In this case, we can identify $\rho_{\mathcal{X}^*}(\varkappa)$ up to an error of order $O(\tau^p)$ using

$$\widetilde{\rho}_{\mathcal{X}^*}(\varkappa,z)\equiv\widetilde{\rho}_{X^*}\left(\widetilde{Q}_{X^*}(F_{\mathcal{X}^*}(\varkappa)),z\right), \tag{11}$$

where

$$\widetilde{Q}_{X^*}(s)\equiv Q_X(s)+\frac{1}{2}\left\{s_X(Q_X(s))\,\widetilde{v}(Q_X(s))+\nabla_x\widetilde{v}(Q_X(s))\right\}. \tag{12}$$

Here $F_{\mathcal{X}^*}(\cdot)$ denotes the CDF of $\mathcal{X}_i^*$, and $Q_X(\cdot)$ denotes the quantile function (QF) of X_i. Note that all functions on the right-hand side of equation (12) are identified directly from the observed data.

First, we demonstrate that $\widetilde{Q}_{X^*}(s)=Q_{X^*}(s)+O(\tau^p)$, where $Q_{X^*}(\cdot)$ denotes the quantile function of X_i^*. Combining this with the result of Theorem 1 allows us to establish the desired result formalized by the theorem below.

Assumption 3.1. $\int\left|\nabla_x^\ell f_{X^*}(x)\right|dx<\infty$ *for* $\ell\in\{1,\dots,p\}$.

Theorem 4. *Suppose that the hypotheses of Theorem 1 are satisfied for* $x=Q_{X^*}(F_{\mathcal{X}^*}(\varkappa))$. *Also, suppose Assumptions MONOT-MEAS and 3.1 hold. Then, as* $\tau\to0$,

$$\widetilde{\rho}_{\mathcal{X}^*}(\varkappa,z_1)=\rho_{\mathcal{X}^*}(\varkappa)+O(\tau^p).$$

Theorem 4 demonstrates that, if the marginal distribution of $\mathcal{X}_i^*$ is given, it is possible to identify $\rho(\varkappa)$ up to an error of order $O(\tau^p)$ in the general NCME model (8) building on the identification results for the WCME model.

Note that obtaining (an estimate of) the marginal distribution $F_{\mathcal{X}^*}$ is a much simpler task than obtaining a validation sample, i.e., the data on $(X_i, \mathcal{X}_i^*)$ jointly. For example, suppose $\mathcal{X}_i^*$ are individual wages, and X_i are self-reported wages in a survey. The marginal distribution $F_{\mathcal{X}^*}$ can be provided by the Social Security Administration or similar tax authorities in other countries. Providing such marginal distribution does not pose any privacy risks. In contrast, obtaining a validation sample that links individual's responses X_i to the individual's social security records $\mathcal{X}_i^*$ is a difficult task that in particular faces major challenges concerning privacy.

If the distribution of $\mathcal{X}_i^*$ is unknown we can still apply Theorem 4 at $\varkappa = Q_{\mathcal{X}^*}(q)$ for any quantile $q \in (0,1)$ to identify $E[Y_i|\mathcal{X}_i^* = Q_{\mathcal{X}^*}(q)] = \rho_{\mathcal{X}^*}(Q_{\mathcal{X}^*}(q))$, where $Q_{\mathcal{X}^*}(\cdot)$ is the (unknown) quantile function of $\mathcal{X}_i^*$:

Corollary 5. *Suppose that the hypotheses of Theorem 4 are satisfied for $x = Q_{X^*}(q)$. Then, as $\tau \to 0$,*

$$\widetilde{\rho}_{X^*}\left(\widetilde{Q}_{X^*}(q), z_1\right) = \widetilde{\rho}_{\mathcal{X}^*}(Q_{\mathcal{X}^*}(q), z_1) = \rho_{\mathcal{X}^*}(Q_{\mathcal{X}^*}(q)) + O(\tau^p).$$

Corollary 5 demonstrates that even if $F_{\mathcal{X}^*}$ is unknown, we can still identify the conditional expectation of Y_i^* given the q'th quantile of $\mathcal{X}_i^*$. Notice that in some applications, the unobserved variable $\mathcal{X}_i^*$, for example an individual's ability, might not even have well-defined economic units. In such settings, identification of $E[Y_i|\mathcal{X}_i^* = Q_{\mathcal{X}^*}(q)]$ is fully exhaustive.

References

Bound, J., C. Brown, and N. Mathiowetz (2001): "Measurement Error in Survey Data," in *Handbook of Econometrics*, Elsevier, 3705–3843.

Carroll, R. J., D. Ruppert, L. A. Stefanski, and C. M. Crainiceanu (2006): *Measurement Error in Nonlinear Models*, Chapman and Hall/CRC.

Carroll, R. J. and L. A. Stefanski (1990): "Approximate quasi-likelihood estimation in models with surrogate predictors," *Journal of the American Statistical Association*, 85, 652–663.

Chen, X., H. Hong, and D. Nekipelov (2011): "Nonlinear Models of Measurement Errors," *Journal of Economic Literature*, 49, 901–937.

Chesher, A. (1991): "The effect of measurement error," *Biometrika*, 78, 451–462.

——— (2000): "Measurement Error Bias Reduction," Working paper.

Chesher, A. and C. Schluter (2002): "Welfare measurement and measurement error," *The Review of Economic Studies*, 69, 357–378.

Evdokimov, K. S. and A. Zeleneev (2018): "Issues of Nonstandard Inference in Measurement Error Models," *Working Paper, Princeton University.*

——— (2022): "Simple Estimation of Semiparametric Models with Measurement Errors," *CeMMAP Working Paper.*

Hausman, J. A., H. Ichimura, W. K. Newey, and J. L. Powell (1991): "Identification and estimation of polynomial errors-in-variables models," *Journal of Econometrics*, 50, 273–295.

Hausman, J. A., W. K. Newey, and J. L. Powell (1995): "Nonlinear errors in variables Estimation of some Engel curves," *Journal of Econometrics*, 65, 205–233.

Hu, Y. and S. M. Schennach (2008): "Instrumental Variable Treatment of Nonclassical Measurement Error Models," *Econometrica*, 76, 195–216.

Li, T. and Q. Vuong (1998): "Nonparametric Estimation of the Measurement Error Model Using Multiple Indicators," *Journal of Multivariate Analysis*, 65, 139–165.

Newey, W. K. (2001): "Flexible Simulated Moment Estimation of Nonlinear Errors-in-Variables Models," *The Review of Economics and Statistics*, 83, 616–627.

Schennach, S. (2022): "Measurement Systems," *Journal of Economic Literature*, 60, 1223–1263.

Schennach, S. M. (2004): "Estimation of Nonlinear Models with Measurement Error," *Econometrica*, 72, 33–75.

——— (2007): "Instrumental Variable Estimation of Nonlinear Errors-in-Variables Models," *Econometrica*, 75, 201–239.

——— (2020): "Mismeasured and unobserved variables," in *Handbook of Econometrics*, Elsevier, 487–565.

Stone, C. J. (1980): "Optimal Rates of Convergence for Nonparametric Estimators," *Annals of Statistics*, 8, 1348–1360.

Wolter, K. M. and W. A. Fuller (1982): "Estimation of Nonlinear Errors-in-Variables Models," *The Annals of Statistics*, 10, 539–548.

A Proofs

A.1 Proof of Theorem 1

Before proving Theorem 1, we state and prove 3 auxiliary lemmas. Then we prove the main result.

A.1.1 Auxiliary lemmas

Lemma A.1. *Suppose that the hypotheses of Theorem 1 are satisfied. Then, for any x and z such that $x \in \mathcal{S}_{X^*}(z)$ (so, $f_{X^*|Z}(x|z) > 0$ and $f_{X^*}(x) > 0$), we have:*

(i)

$$q(x, z) = \rho(x) + v(x)\left(\rho'(x)\, s_{X^*|Z}(x|z) + \frac{1}{2}\rho''(x)\right) + \rho'(x)\, v'(x) + O(\tau^p), \tag{A.1}$$

(ii)

$$\left|s_{X|Z}(x|z) - s_{X^*|Z}(x|z)\right| + |q'(x) - \rho'(x)| = O(\tau^2). \tag{A.2}$$

Proof of Lemma A.1, Part (i). For concreteness, we establish the result for the case $p = 4$. The proof for the case $p = 3$ is analogous (but simpler).

We have

$$\begin{aligned} f_{X^*X|Z}(r, x|z) &= f_{\varepsilon|X^*}(x - r|r)\, f_{X^*|Z}(r|z), \\ f_{X^*|XZ}(r|x, z) &= \frac{f_{X^*X|Z}(r, x|z)}{f_{X|Z}(x|z)} = \frac{f_{\varepsilon|X^*}(x - r|r)\, f_{X^*|Z}(r|z)}{f_{X|Z}(x|z)}, \\ f_{X|Z}(x|z) &= \int f_{\varepsilon|X^*}(x - r|r)\, f_{X^*|Z}(r|z)\, dr. \end{aligned}$$

Also, notice that

$$\begin{aligned} q(x, z) &\equiv E[Y_i|X_i = x, Z_i = z] = E[E[Y_i|X_i^*, X_i, Z_i]\,|X_i = x, Z_i = z] \\ &= E[\rho(X_i^*)|X_i = x, Z_i = z], \end{aligned}$$

where the last equality follows from Assumptions 2.1 and 2.2.

We want to compute

$$E\left[\rho\left(X^{*}\right)|X=x,Z=z\right]=\frac{\int\rho\left(r\right)f_{X^{*}|XZ}\left(r|x,z\right)dr}{f_{X|Z}\left(x|z\right)}=\frac{\int\rho\left(r\right)f_{\varepsilon|X^{*}}\left(x-r|r\right)f_{X^{*}|Z}\left(r|z\right)dr}{f_{X|Z}\left(x|z\right)}.$$

Notice that on the RHS, the denominator is a special case of the numerator with $\rho\left(r\right)=1$ for all r. Thus, it is sufficient to focus on the numerator.

Next, we consider $\int\eta\left(r\right)f_{\varepsilon|X^{*}}\left(x-r|r\right)dr$, where $\eta(\cdot)$ is a bounded function with p bounded derivatives. Recall that we have $\varepsilon=\tau\xi$, so $f_{\varepsilon|X^{*}}\left(\varepsilon|r\right)=\frac{1}{\tau}f_{\xi|X^{*}}\left(\frac{\varepsilon}{\tau}|r\right)$. Then,

$$\begin{aligned}\int\eta\left(r\right)f_{\varepsilon|X^{*}}\left(x-r|r\right)dr&=\int\eta\left(x-\tau u\right)\tau f_{\varepsilon|X^{*}}\left(\tau u|x-\tau u\right)du\\&=\int\eta\left(x-\tau u\right)f_{\xi|X^{*}}\left(u|x-\tau u\right)du,\end{aligned}$$

where we used the substitution $r=x-\tau u$. Using Assumption 2.4, we have (for all u)

$$\begin{aligned}\eta\left(x-\tau u\right)f_{\xi|X^{*}}\left(u|x-\tau u\right)=&\eta(x)f_{\xi|X^{*}}(u|x)+\sum_{\ell=1}^{p-1}\frac{(-\tau)^{\ell}}{\ell!}u^{\ell}\nabla_{x}^{\ell}\{\eta(x)f_{\xi|X^{*}}(u|x)\}\\&+\frac{(-\tau)^{p}}{p!}u^{p}\nabla_{x}^{p}\left\{\eta(x)f_{\xi|X^{*}}(u|x)\right\}\Big|_{x=x-\tilde{\tau}(u)u}\end{aligned}$$

for some $\tilde{\tau}(u)\in(0,\tau)$. By boundedness of $\eta(\cdot)$ (and its derivatives) and Assumption 2.5, all the terms on the RHS of the equation above are integrable, and consequently we have

$$\begin{aligned}&\int\eta\left(x-\tau u\right)f_{\xi|X^{*}}\left(u|x-\tau u\right)du\\=&\int\eta\left(x\right)f_{\xi|X^{*}}\left(u|x\right)du-\tau\int u\left\{\eta'\left(x\right)f_{\xi|X^{*}}\left(u|x\right)+\eta\left(x\right)\nabla_{x}f_{\xi|X^{*}}\left(u|x\right)\right\}du\\&+\frac{\tau^{2}}{2}\int u^{2}\left\{\eta''\left(x\right)f_{\xi|X^{*}}\left(u|x\right)+2\eta'\left(x\right)\nabla_{x}f_{\xi|X^{*}}\left(u|x\right)+\eta\left(x\right)\nabla_{x}^{2}f_{\xi|X^{*}}\left(u|x\right)\right\}du+R_{\eta}(\tau)\\=&\eta\left(x\right)-\tau\left\{\eta'\left(x\right)E\left[\xi|X^{*}=x\right]+\eta\left(x\right)\nabla_{x}E\left[\xi|X^{*}=x\right]\right\}\\&+\frac{\tau^{2}}{2}\left\{\eta''\left(x\right)E\left[\xi^{2}|X^{*}=x\right]+2\eta'\left(x\right)\nabla_{x}E\left[\xi^{2}|X^{*}=x\right]+\eta\left(x\right)\nabla_{x}^{2}E\left[\xi^{2}|X^{*}=x\right]\right\}+R_{\eta}(\tau)\\=&\eta\left(x\right)+\frac{1}{2}\{\eta''\left(x\right)v\left(x\right)+2\eta'\left(x\right)\nabla_{x}v\left(x\right)+\eta\left(x\right)\nabla_{x}^{2}v\left(x\right)\}+R_{\eta}(\tau),\end{aligned}$$

where (for $p = 4$)

$$R_\eta(\tau) = -\frac{\tau^3}{6}\int u^3\nabla_x^3\{\eta(x)f_{\xi|X^*}(u|x)\}du + \frac{\tau^4}{24}\int u^4\nabla_x^4\left\{\eta(\tilde{x})f_{\xi|X^*}(u|\tilde{x})\right\}\big|_{x=x-\tilde{\tau}(u)u}du.$$

In the derivation above, the second equality uses $\int u^\ell \nabla_x^k f_{\xi|X^*}(u|x)du = \nabla_x^k \int u^\ell f_{\xi|X^*}(u|x)du = \nabla_x^k E[\xi^k|X^* = x]$ (differentiation under the integral sign is possible due to Assumptions 2.4 and 2.5 for non-negative integers $k, \ell \leqslant p$), and the last equality is due to $E[\xi|X^* = x] = 0$ (as a function of x) and uses the notation $v(x) = E[\varepsilon^2|X^* = x] = \tau^2 E[\xi^2|X^* = x]$.

Next, we establish $R_\eta(\tau) = O(\tau^p)$. Notice that

$$\begin{aligned}
&\int u^3\nabla_x^3\{\eta(x)f_{\xi|X^*}(u|x)\}du \\
=\ &\int u^3\{\eta'''(x)f_{\xi|X^*}(u|x) + 3\eta''(x)\nabla_x f_{\xi|X^*}(u|x) + 3\eta'(x)\nabla_x^2 f_{\xi|X^*}(u|x) + \eta(x)\nabla_x^3 f_{\xi|X^*}(u|x)\}du \\
=\ &\eta'''(x)E\left[\xi^3|X^* = x\right] + 3\eta''(x)\nabla_x E\left[\xi^3|X^* = x\right] + 3\eta'(x)\nabla_x^2 E\left[\xi^3|X^* = x\right] + \eta(x)\nabla_x^3 E\left[\xi^3|X^* = x\right] \\
=\ &0,
\end{aligned}$$

where the last equality uses $E[\xi^3|X^* = x] = 0$ (as a function of x). Finally,

$$\begin{aligned}
&\int u^4\nabla_x^4\left\{\eta(\tilde{x})f_{\xi|X^*}(u|\tilde{x})\right\}\big|_{\tilde{x}=x-\tilde{\tau}(u)u}du \\
=\ &\int u^4\Big\{\eta''''(x-\tilde{\tau}(u)u)f_{\xi|X^*}(u|x-\tilde{\tau}(u)u) + 4\eta'''(x-\tilde{\tau}(u)u)\nabla_x f_{\xi|X^*}(u|x-\tilde{\tau}(u)u) \\
&+6\eta''(x-\tilde{\tau}(u)u)\nabla_x^2 f_{\xi|X^*}(u|x-\tilde{\tau}(u)u) + 4\eta'(x-\tilde{\tau}(u)u)\nabla_x^3 f_{\xi|X^*}(u|x-\tilde{\tau}(u)u) \\
&+\eta(x-\tilde{\tau}(u)u)\nabla_x^4 f_{\xi|X^*}(u|x-\tilde{\tau}(u)u)\Big\}du.
\end{aligned}$$

Since η and its derivatives are uniformly bounded, for some $C > 0$, we have

$$|R_\eta(\tau)| \leqslant C\tau^4\int u^4\sup_{\tilde{x}\in\mathcal{S}_X}\left\{f_{\xi|X^*}(u|\tilde{x}) + \sum_{\ell=1}^4\left|\nabla_x^\ell f_{\xi|X^*}(u|\tilde{x})\right|\right\}du = O(\tau^4),$$

where the last equality uses Assumption 2.5. Hence, we conclude

$$\int\eta(r)f_{\varepsilon|X^*}(x-r|r)\,dr = \eta(x) + \frac{1}{2}\{\eta''(x)v(x) + 2\eta'(x)\nabla_x v(x) + \eta(x)\nabla_x^2 v(x)\} + O(\tau^p).$$

Note that this result also implies

$$\int \eta\left(r\right) f_{\varepsilon|X^*}\left(x-r|r\right) dr = \eta(x) + O(\tau^2). \tag{A.3}$$

Next, we apply the derived result to

$$q\left(x,z\right) = \frac{\int \rho\left(r\right) f_{X^*|Z}\left(r|z\right) f_{\varepsilon|X^*}\left(x-r|r\right) dr}{\int f_{X^*|Z}\left(r|z\right) f_{\varepsilon|X^*}\left(x-r|r\right) dr}.$$

For the numerator, we use $\eta\left(x\right) = \rho\left(x\right) f_{X^*|Z}\left(x|z\right)$, $\eta'\left(x\right) = \rho\left(x\right)' f_{X^*|Z}\left(x|z\right) + \rho\left(x\right) f'_{X^*|Z}\left(x|z\right)$, and $\eta''\left(x\right) = \rho''\left(x\right) f_{X^*|Z}\left(x|z\right) + 2\rho'\left(x\right) f'_{X^*|Z}\left(x|z\right) + \rho\left(x\right) f''_{X^*|Z}\left(x|z\right)$. Thus,

$$\begin{aligned}
&\int \rho\left(r\right) f_{X^*|Z}\left(r|z\right) f_{\varepsilon|X^*}\left(x-r|r\right) dr \\
=&\rho(x)\left(f_{X^*|Z}(x|z) + \frac{1}{2}\{f''_{X^*|Z}\left(x|z\right) v\left(x\right) + 2f'_{X^*|Z}\left(x|z\right)\nabla_x v\left(x\right) + f_{X^*|Z}\left(x|z\right)\nabla_x^2 v\left(x\right)\}\right) \\
&+ \frac{1}{2}\left(\rho''\left(x\right) f_{X^*|Z}\left(x|z\right) + 2\rho'\left(x\right) f'_{X^*|Z}\left(x|z\right)\right) v(x) + \rho\left(x\right)' f_{X^*|Z}\left(x|z\right)\nabla_x v\left(x\right) + O(\tau^p).
\end{aligned}$$

Similarly,

$$\begin{aligned}
f_{X|Z}(x|z) &= \int f_{X^*|Z}\left(r|z\right) f_{\varepsilon|X^*}\left(x-r|r\right) dr \\
&= f_{X^*|Z}(x|z) + \frac{1}{2}\{f''_{X^*|Z}\left(x|z\right) v\left(x\right) + 2f'_{X^*|Z}\left(x|z\right)\nabla_x v\left(x\right) + f_{X^*|Z}\left(x|z\right)\nabla_x^2 v\left(x\right)\} + O(\tau^p).
\end{aligned}$$

Hence, using $\nabla_x^\ell v(x) = O(\tau^2)$ for $l \in \{0,1,2\}$, we conclude

$$\begin{aligned}
q\left(x,z\right) &= \rho\left(x\right) + \frac{\frac{1}{2}\left(\rho''\left(x\right) f_{X^*|Z}\left(x|z\right) + 2\rho'\left(x\right) f'_{X^*|Z}\left(x|z\right)\right) v\left(x\right) + \rho\left(x\right)' f_{X^*|Z}\left(x|z\right)\nabla_x v\left(x\right)}{f_{X^*|Z}\left(x|z\right) + O(\tau^2)} + O\left(\tau^p\right) \\
&= \rho\left(x\right) + \left(\frac{1}{2}\rho''\left(x\right) + \rho'\left(x\right) s_{X^*|Z}\left(x|z\right)\right) v(x) + \rho'\left(x\right) v'(x) + O\left(\tau^p\right).
\end{aligned}$$

□

Proof of Lemma A.1, Part (ii). First, we consider

$$\begin{aligned}\nabla_x \int \eta(r) f_{\varepsilon|X^*}(x-r|r)dr &= \nabla_x \int \eta(x-\tau u) f_{\xi|X^*}(u|x-\tau u)du \\ &= \int \nabla_x \{\eta(x-\tau u) f_{\xi|X^*}(u|x-\tau u)\}du,\end{aligned}$$

where the second equality follows from Assumptions 2.4 and 2.5. Analogously to the expansion of $\eta(x-\tau u)f_{\xi|X^*}(u|x-\tau u)$ considered in the proof of Part (i), we have

$$\begin{aligned}\nabla_x\{\eta(x-\tau u)f_{\xi|X^*}(u|x-\tau u)\} =& \eta'(x)f_{\xi|X^*}(u|x)+\eta(x)\nabla_x f_{\xi|X^*}(u|x) \\ &-\tau u \nabla_x^2\{\eta(x)f_{\xi|X^*}(u|x)\} \\ &+\frac{\tau^2}{2}u^2\nabla_x^3\{\eta(x-\tilde{\tau}(u)u)f_{\xi|X^*}(u|x-\tilde{\tau}(u)u)\},\end{aligned}$$

for some $\tilde{\tau}(u)\in(0,\tau)$. Hence,

$$\begin{aligned}\nabla_x \int \eta(r) f_{\varepsilon|X^*}(x-r|r)dr &= \int \left(\eta'(x)f_{\xi|X^*}(u|x)+\eta(x)\nabla_x f_{\xi|X^*}(u|x)\right)du + O(\tau^2) \\ &= \eta'(x)+O(\tau^2).\end{aligned}$$

Plugging $\rho(x)f_{X^*}(x)$, $f_{X^*}(x)$, and $f_{X^*|Z}(x|z)$ as $\eta(x)$, we obtain

$$\begin{aligned}\nabla_x \int \rho(r)f_{X^*}(r)f_{\varepsilon|X^*}(x-r|r)dr &= \nabla_x\{\rho(x)f_{X^*}(x)\}+O(\tau^2), \\ f_X'(x) = \nabla_x \int f_{X^*}(r)f_{\varepsilon|X^*}(x-r|r)dr &= f_{X^*}'(x)+O(\tau^2), \\ f_{X|Z}'(x,z) = \nabla_x \int f_{X^*|Z}(r|z)f_{\varepsilon|X^*}(x-r|r)dr &= f_{X^*|Z}'(x|z)+O(\tau^2).\end{aligned}$$

Similarly, using equation (A.3) derived in the proof of Part (i),

$$\begin{aligned}\int \rho(r)f_{X^*}(r)f_{\varepsilon|X^*}(x-r|r)dr &= \rho(x)f_{X^*}(x)+O(\tau^2), \\ f_X(x) = \int f_{X^*}(r)f_{\varepsilon|X^*}(x-r|r)dr &= f_{X^*}(x)+O(\tau^2), \\ f_{X|Z}(x,z) = \int f_{X^*|Z}(r|z)f_{\varepsilon|X^*}(x-r|r)dr &= f_{X^*|Z}(x|z)+O(\tau^2).\end{aligned}$$

Using the results above, we establish

$$\begin{aligned} s_{X|Z}(x|z) - s_{X^*|Z}(x|z) &= \frac{f'_{X|Z}(x|z) f_{X^*|Z}(x|z) - f'_{X|Z}(x|z) f_{X^*|Z}(x|z)}{f_{X|Z}(x|z) f_{X^*|Z}(x|z)} \\ &= \frac{O(\tau^2)}{f_{X^*|Z}(x|z)^2 + O(\tau^2)} = O(\tau^2), \end{aligned} \tag{A.4}$$

Similarly,

$$\begin{aligned} q'(x) &= \nabla_x \left\{ \frac{\int \rho(r) f_{X^*}(r) f_{\varepsilon|X^*}(x-r|r)dr}{\int f_{X^*}(r) f_{\varepsilon|X^*}(x-r|r)dr} \right\} \\ &= \frac{(\rho'(x) f_{X^*}(x) + \rho(x) f'_{X^*}(x) + O(\tau^2))(f_{X^*}(x) + O(\tau^2)) - (f'_{X^*}(x) + O(\tau^2))(\rho(x) f_{X^*}(x) + O(\tau^2))}{(f_{X^*}(x) + O(\tau^2))^2} \\ &= \rho'(x) + O(\tau^2). \end{aligned} \tag{A.5}$$

which completes the proof. □

Before proving additional auxiliary results we introduce an additional notation.

Definition. *Consider a function $g : \mathcal{X} \times \mathbb{R}^+ \to \mathbb{R}$ such that $\nabla_x g(x,\tau)$ exists (for every $\tau \in \mathbb{R}^+$). We say that $g(x,\tau) = O_x(\tau^p)$ if $g(x,\tau) = O(\tau^p)$ and $\nabla_x g(x,\tau) = O(\tau^p)$ as $\tau \downarrow 0$.*

Lemma A.2. *Suppose that the hypotheses of Theorem 1 are satisfied. Then, we have $f_{X|Z}(x,z) = f_{X^*|Z}(x|z) + O_x(\tau^2)$, $s_{X|Z}(x|z) = s_{X^*|Z}(x|z) + O_x(\tau^{p-2})$, and $q'(x) = \rho'(x) + O_x(\tau^{p-2})$.*

Proof of Lemma A.2. For concreteness, we provide the proof for $p = 4$. The proof for $p = 3$ is analogous (but simpler).

First, consider $\int \eta(r) f_{\varepsilon|X^*}(x-r|r)dr = \int \eta(x-\tau u) f_{\xi|X^*}(u|x-\tau u)du$, where $\eta(\cdot)$ is a bounded function with p bounded derivatives. Considering an expansion analogous to the one provided in the proof of Lemma A.1, Part (i), we obtain

$$\begin{aligned} \eta(x-\tau u) f_{\xi|X^*}(u|x-\tau u) =& \eta(x) f_{\xi|X^*}(u|x) - \tau u \nabla_x \{\eta(x) f_{\xi|X^*}(u|x)\} \\ &+ \int_0^\tau u^2 \nabla_x^2 \{\eta(x-tu) f_{\xi|X^*}(u|x-tu)\}(\tau - t)dt, \end{aligned}$$

where the remainder is given in the integral form. Then,

$$\begin{aligned}\int \eta(r) f_{\varepsilon|X^*}(x-r|r)dr &= \int \eta(x-\tau u) f_{\xi|X^*}(u|x-\tau u)du \\ &= \eta(x) + R_\eta(x;\tau),\end{aligned}$$

where

$$R_\eta(x;\tau) = \int \left(\int_0^\tau u^2 \nabla_x^2 \{\eta(x-tu) f_{\varepsilon|X^*}(u|x-tu)\}(\tau-t)dt \right) du.$$

Since η and its derivatives are bounded, for some $C>0$, we have

$$\left| \int_0^\tau u^2 \nabla_x^2 \{\eta(x-tu) f_{\varepsilon|X^*}(u|x-tu)\}(\tau-t)dt \right| \leqslant \frac{\tau^2}{2} |u|^2 \sup_{\tilde{x}\in\mathcal{S}_X} \sum_{\ell=0}^{2} \left|\nabla_x^\ell f_{\xi|X^*}(u|\tilde{x})\right|.$$

Combining this result with Assumption 2.5, we obtain $R_\eta(x;\tau) = O(\tau^2)$. Similarly, we conclude

$$\nabla_x R_\eta(x;\tau) = \int \left(\int_0^\tau u^2 \nabla_x^3 \{\eta(x-tu) f_{\varepsilon|X^*}(u|x-tu)\}(\tau-t)dt \right) du = O(\tau^2).$$

Hence, we conclude

$$\int \eta(r) f_{\varepsilon|X^*}(x-r|r)dr = \eta(x) + O_x(\tau^2). \tag{A.6}$$

Next, consider $\nabla_x \int \eta(r) f_{\varepsilon|X^*}(x-r|r)dr = \nabla_x \int \eta(x-\tau u) f_{\xi|X^*}(u|x-\tau u)du$, where $\eta(\cdot)$ is a bounded function with p bounded derivatives. Considering an expansion analogous to the one derived in the proof of Lemma A.1, Part (ii), we obtain

$$\begin{aligned}\nabla_x \int \eta(r) f_{\varepsilon|X^*}(x-r|r)dr &= \nabla_x \int \eta(x-\tau u) f_{\xi|X^*}(u|x-\tau u)du \\ &= \eta'(x) + r_\eta(x;\tau),\end{aligned}$$

where

$$r_\eta(x,\tau) = \int \left(\int_0^\tau u^2 \nabla_x^3 \{\eta(x-tu) f_{\xi|X^*}(u|x-tu)\}(\tau-t)dt \right) du = \nabla_x R_\eta(x;\tau) = O(\tau^2).$$

Similarly,

$$\nabla_x r_\eta(x,\tau) = \int \left(\int_0^\tau u^2 \nabla_x^4 \{\eta(x-tu) f_{\xi|X^*}(u|x-tu)\}(\tau-t)dt \right) du = O(\tau^2),$$

which demonstrates

$$\nabla_x \int \eta(r) f_{\varepsilon|X^*}(x-r|r)dr = \eta'(x) + O_x(\tau^2). \tag{A.7}$$

Applying (A.6) and (A.7) with $\rho(x)f_{X^*}(x)$, $f_{X^*}(x)$, and $f_{X^*|Z}(x|z)$ as $\eta(x)$ we obtain

$$\begin{aligned}
\int \rho(r) f_{X^*}(r) f_{\varepsilon|X^*}(x-r|r)dr &= \rho(x)f_{X^*}(x) + O_x(\tau^2),\\
f_X(x) = \int f_{X^*}(r) f_{\varepsilon|X^*}(x-r|r)dr &= f_{X^*}(x) + O_x(\tau^2),\\
f_{X|Z}(x,z) = \int f_{X^*|Z}(r|z) f_{\varepsilon|X^*}(x-r|r)dr &= f_{X^*|Z}(x|z) + O_x(\tau^2),
\end{aligned}$$

and

$$\begin{aligned}
\nabla_x \int \rho(r) f_{X^*}(r) f_{\varepsilon|X^*}(x-r|r)dr &= \nabla_x\{\rho(x)f_{X^*}(x)\} + O_x(\tau^2),\\
f'_X(x) = \nabla_x \int f_{X^*}(r) f_{\varepsilon|X^*}(x-r|r)dr &= f'_{X^*}(x) + O_x(\tau^2),\\
f'_{X|Z}(x,z) = \nabla_x \int f_{X^*|Z}(r|z) f_{\varepsilon|X^*}(x-r|r)dr &= f'_{X^*|Z}(x|z) + O_x(\tau^2).
\end{aligned}$$

Combining the results above with derivations as in equations (A.4) and (A.5) completes the proof. □

Lemma A.3. *Suppose that the hypotheses of Theorem 1 are satisfied. Then,*

$$q(x,z_1) - q(x,z_2) = \rho'(x)\, v(x) \left(s'_{X^*|Z}(x|z_1) - s'_{X^*|Z}(x|z_2) \right) + O_x(\tau^p).$$

Proof of Lemma A.3. For concreteness, we provide the proof for $p=4$. The proof for $p=3$ is analogous (but simpler).

1. Note that by bringing the difference of ratios to the common denominator we can write

$$q(x,z_1) - q(x,z_2) = \frac{N(x,z_1,z_2,\tau)}{D(x,z_1,z_2,\tau)},$$

where $D(x, z_1, z_2, \tau) = f_{X|Z}(x|z_1) f_{X|Z}(x|z_2)$. The numerator is

$$\begin{aligned}
& N(x, z_1, z_2, \tau) \\
\equiv\ & \int \rho(r) f_{X^*|Z}(r|z_1) f_{\varepsilon|X^*}(x-r|r)\, dr \times \int f_{X^*|Z}(r|z_2) f_{\varepsilon|X^*}(x-r|r)\, dr \\
& \quad - \int \rho(r) f_{X^*|Z}(r|z_2) f_{\varepsilon|X^*}(x-r|r)\, dr \times \int f_{X^*|Z}(r|z_1) f_{\varepsilon|X^*}(x-r|r)\, dr \\
=\ & \iint \int \{\rho(x-\tau u_1) - \rho(x-\tau u_2)\} f_{X^*|Z}(x-\tau u_1|z_1) f_{X^*|Z}(x-\tau u_2|z_2) \\
& \quad \times f_{\xi|X^*}(u_1|x-\tau u_1) f_{\xi|X^*}(u_2|x-\tau u_2)\, du_1 du_2 \\
=\ & \iint \{\rho(x-\tau u_1) - \rho(x-\tau u_2)\} f_{\xi,X^*|Z}(u_1, x-\tau u_1|z_1) f_{\xi,X^*|Z}(u_2, x-\tau u_2|z_2)\, du_1 du_2.
\end{aligned}$$

Note that for any q_1, q_2, Q_1, Q_2 we have

$$q_1 q_2 = (q_1 - Q_1) Q_2 + Q_1 (q_2 - Q_2) + (q_1 - Q_1)(q_2 - Q_2) + Q_1 Q_2,$$

and taking $q_j \equiv f_{\xi,X^*|Z}(u_j, x-\tau u_j|z_j)$, $Q_j = f_{\xi,X^*|Z}(u_j, x|z_j) - f'_{\xi,X^*|Z}(u_j, x|z_j)\tau u_j$ for $j \in [2]$ we have

$$\begin{aligned}
& f_{\xi,X^*|Z}(u_1, x-\tau u_1|z_1) f_{\xi X^*|Z}(u_2, x-\tau u_2|z_2) \\
=\ & T_{ff,1} + T_{ff,2} + T_{ff,3} + T_{ff,4} \\
T_{ff,1} \equiv\ & \left(f_{\xi,X^*|Z}(u_1, x-\tau u_1|z_1) + f'_{\xi,X^*|Z}(u_1, x|z_1)\tau u_1 - f_{\xi,X^*|Z}(u_1, x|z_1)\right) \\
& \quad \times \left(-f'_{\xi,X^*|Z}(u_2, x|z_2)\tau u_2 + f_{\xi,X^*|Z}(u_2, x|z_2)\right) \\
T_{ff,2} \equiv\ & \left(f_{\xi,X^*|Z}(u_2, x-\tau u_2|z_2) + f'_{\xi,X^*|Z}(u_2, x|z_2)\tau u_2 - f_{\xi,X^*|Z}(u_2, x|z_2)\right) \\
& \quad \times \left(-f'_{\xi,X^*|Z}(u_1, x|z_1)\tau u_1 + f_{\xi,X^*|Z}(u_1, x|z_1)\right) \\
T_{ff,3} \equiv\ & \left(f_{\xi,X^*|Z}(u_1, x-\tau u_1|z_1) + f'_{\xi,X^*|Z}(u_1, x|z_1)\tau u_1 - f_{\xi,X^*|Z}(u_1, x|z_1)\right) \\
& \quad \times \left(f_{\xi X^*|Z}(u_2, x-\tau u_2|z_2) + f'_{\xi,X^*|Z}(u_2, x|z_2)\tau u_2 - f_{\xi,X^*|Z}(u_2, x|z_2)\right) \\
T_{ff,4} \equiv\ & \left(-f'_{\xi,X^*|Z}(u_1, x|z_1)\tau u_1 + f_{\xi,X^*|Z}(u_1, x|z_1)\right)\left(-f'_{\xi,X^*|Z}(u_2, x|z_2)\tau u_2 + f_{\xi,X^*|Z}(u_2, x|z_2)\right)
\end{aligned}$$

and let

$$I_{ff,j}(x) \equiv \iint \{\rho(x-\tau u_1) - \rho(x-\tau u_2)\} T_{ff,j}(\ldots)\, du_1 du_2, \quad j \in [4].$$

2. Note that

$$
\begin{aligned}
& f_{\xi,X^*|Z}\left(u_j, x-\tau u_j|z_j\right) + f'_{\xi,X^*|Z}\left(u_j, x|z_j\right)\tau u_j - f_{\xi,X^*|Z}\left(u_j, x|z_j\right) \\
= & \int_0^\tau f''_{\xi,X^*|Z}\left(u_j, x-tu_j|z_j\right) u_j^2 \left(\tau - t\right) dt \qquad \text{(A.8a)} \\
= & \frac{\tau^2}{2} f''_{\xi,X^*|Z}\left(u_j, x|z_j\right) u_j^2 - \int_0^\tau \frac{1}{2} f'''_{\xi,X^*|Z}\left(u_j, x-tu_j|z_j\right) u_j^3 \left(\tau - t\right)^2 dt \qquad \text{(A.8b)}
\end{aligned}
$$

and

$$
\begin{aligned}
& \rho\left(x-\tau u_1\right) - \rho\left(x-\tau u_2\right) \\
= & \sum_{j=1}^{2} (-1)^j \int_0^\tau \rho'(x-tu_j)u_j dt \qquad \text{(A.9a)} \\
= & \tau\rho'\left(x\right)\left(u_2-u_1\right) + \sum_{j=1}^{2} (-1)^{j-1} \int_0^\tau \rho''\left(x-tu_j\right) u_j^2 \left(\tau-t\right) dt \qquad \text{(A.9b)} \\
= & \tau\rho'\left(x\right)\left(u_2-u_1\right) + \frac{\tau^2}{2}\rho''\left(x\right)\left(u_1^2-u_2^2\right) \\
& + \sum_{j=1}^{2} (-1)^j \int_0^\tau \frac{1}{2}\rho'''\left(x-tu_j\right) u_j^3 \left(\tau-t\right)^2 dt. \qquad \text{(A.9c)}
\end{aligned}
$$

We will use the notation

$$
\begin{aligned}
\mathbb{M}_{|\zeta_1\zeta_2 z_1 z_2}\left[a\left(x, U_1, U_2\right)\right] & \equiv \mathbb{E}_{|\zeta_1\zeta_2 z_1 z_2}\left[a\left(x, U_1, U_2\right)\right] f_{X^*|Z}\left(\zeta_1|z_1\right) f_{X^*|Z}\left(\zeta_2|z_2\right), \\
\mathbb{E}_{|\zeta_1\zeta_2 z_1 z_2}\left[a\left(x, U_1, U_2\right)\right] & \equiv \iint a\left(x, u_1, u_2\right) f_{\xi|X^*,Z}\left(u_1|\zeta_1, z_1\right) f_{\xi|X^*,Z}\left(u_2|\zeta_2, z_2\right) du_1 du_2,
\end{aligned}
$$

i.e., $\mathbb{E}_{|\zeta_1\zeta_2 z_1 z_2}[\cdot]$ is the expectation w.r.t. U_1 and U_2, where U_1 and U_2 are independent and $U_j \sim f_{\xi|X^*,Z}\left(\cdot|\zeta_j, z_j\right) = f_{\xi|X^*}(\cdot|\zeta_j)$ for $j \in [2]$.

Let

$$
\begin{aligned}
\Delta\left(x,q,s\right) & \equiv \left\{\rho\left(x-\tau u_1\right) - \rho\left(x-\tau u_2\right)\right\} \times \nabla_x^q f_{\xi,X^*|Z}\left(u_1, x|z_1\right) u_1^q \times \nabla_x^s f_{\xi,X^*|Z}\left(u_2, x|z_2\right) u_2^s, \\
I_{\Delta(x,q,s)} & \equiv \iint \Delta\left(x,q,s\right) du_1 du_2 = \nabla_{\zeta_1}^q \nabla_{\zeta_2}^s \mathbb{M}_{|\zeta_1\zeta_2 z_1 z_2}\left[\left(\rho\left(x-\tau U_1\right) - \rho\left(x-\tau U_2\right)\right) U_1^q U_2^s\right]\Big|_{\zeta_1=\zeta_2=x}.
\end{aligned}
$$

Also, let

$$
\begin{aligned}
h(x,a,b,q,s) &\equiv u_1^a u_2^b \times \nabla_x^q f_{\xi,X^*|Z}(u_1,x|z_1)\, u_1^q \times \nabla_x^s f_{\xi,X^*|Z}(u_2,x|z_2)\, u_2^s \\
&= u_1^{a+q} u_2^{b+s} \times \nabla_x^q f_{\xi,X^*|Z}(u_1,x|z_1)\, \nabla_x^s f_{\xi,X^*|Z}(u_2,x|z_2)
\end{aligned}
$$

and

$$
I_{h(x,a,b,q,s)} \equiv \iint h(x,a,b,q,s)\, du_1 du_2 = \nabla_{\zeta_1}^q \nabla_{\zeta_2}^s \mathbb{M}_{|\zeta_1\zeta_2 z_1 z_2}\left[U_1^{a+q} U_2^{b+s}\right]\Big|_{\zeta_1=\zeta_2=x}. \tag{A.10}
$$

From equation (A.9c) we have

$$
\begin{aligned}
&\Delta(x,q,s) \\
= &\left\{\tau\rho'(x)(u_2-u_1) + \frac{\tau^2}{2}\rho''(x)\left(u_1^2-u_2^2\right) + \sum_{j=1}^{2}(-1)^j \int_0^\tau \frac{1}{2!}\rho'''(x-tu_j)\, u_j^3 (\tau-t)^2 dt\right\} \\
&\quad \times h(x,0,0,q,s) \\
= &\ \tau\rho'(x)\{h(x,0,1,q,s)-h(x,1,0,q,s)\} + \frac{\tau^2}{2}\rho''(x)\{h(x,2,0,q,s)-h(x,0,2,q,s)\} \\
&\quad + \sum_{j=1}^{2}(-1)^j \int_0^\tau \frac{1}{2}\rho'''(x-tu_j)\, u_j^3 (\tau-t)^2 dt \times h(x,0,0,q,s).
\end{aligned}
$$

Thus, for $q,s \in \{0,1\}$, we have

$$
\begin{aligned}
I_{\Delta(x,q,s)} &= \tau\rho'(x)\left\{I_{h(x,0,1,q,s)} - I_{h(x,1,0,q,s)}\right\} \\
&\quad + \frac{\tau^2}{2}\rho''(x)\left\{I_{h(x,2,0,q,s)} - I_{h(x,0,2,q,s)}\right\} + O_x\left(\tau^3\right).
\end{aligned} \tag{A.11}
$$

Notice that if $q=2$ or $s=2$, we use equation (A.9b) instead to obtain

$$
\begin{aligned}
&\Delta(x,q,s) \\
= &\left\{\tau\rho'(x)(u_2-u_1) + \sum_{j=1}^{2}(-1)^{j-1}\int_0^\tau \rho''(x-tu_j)\, u_j^2(\tau-t)\, dt\right\} h(x,0,0,q,s) \\
= &\ \tau\rho'(x)\{h(x,0,1,q,s)-h(x,1,0,q,s)\} \\
&\quad + \sum_{j=1}^{2}(-1)^{j-1}\int_0^\tau \rho''(x-tu_j)\, u_j^2(\tau-t)\, dt \times h(x,0,0,q,s).
\end{aligned}
$$

Thus, for both $q, s \in \{0,1,2\}$, we have

$$I_{\Delta(x,q,s)} = \tau\rho'(x)\left\{I_{h(x,0,1,q,s)} - I_{h(x,1,0,q,s)}\right\} + O_x\left(\tau^2\right). \tag{A.12}$$

3.1. Using equation (A.8a) we have

$$T_{ff,3}(x) = \int_0^\tau f''_{\xi,X^*|Z}(u_1, x - t_1 u_1|z_1)\, u_1^2 (\tau - t_1)\, dt_1 \times \int_0^\tau f''_{\xi,X^*|Z}(u_2, x - t_2 u_2|z_2)\, u_2^2 (\tau - t_2)\, dt_2.$$

First,

$$|T_{ff,3}(x)| \leqslant u_1^2 \sup_{\tilde{x}\in\mathcal{S}_X} \left|f''_{\xi,X^*|Z}(u_1,\tilde{x}|z_1)\right| u_2^2 \sup_{\tilde{x}\in\mathcal{S}_X} \left|f''_{\xi,X^*|Z}(u_2,\tilde{x}|z_2)\right| \frac{\tau^4}{4}.$$

Next,

$$\begin{aligned} f''_{\xi,X^*|Z}(u,x|z) &= \nabla_x^2\{f_{\xi|X^*}(u|x) f_{X^*|Z}(x|z)\} \\ &= f''_{\xi|X^*}(u|x) f_{X^*|Z}(x|z) + 2 f'_{\xi|X^*}(u|x) f'_{X^*|Z}(x|z) + f_{\xi|X^*}(u|x) f''_{X^*|Z}(x|z). \end{aligned}$$

Hence, using Assumption 2.5 and boundedness of ρ and $f_{\xi|X^*}$ and its derivatives, we conclude

$$I_{ff,3}(x) = O(\tau^4).$$

Similarly,

$$\left|T'_{ff,3}(x)\right| \leqslant \sum_{j=0}^{1} \sup_{\tilde{x}\in\mathcal{S}_X} \left|\nabla_x^{3-j} f_{\xi,X^*|Z}(u_1,\tilde{x}|z_1)\right| \sup_{\tilde{x}\in\mathcal{S}_X} \left|\nabla_x^{2+j} f_{\xi,X^*|Z}(u_2,\tilde{x}|z_2)\right| u_1^2 u_2^2 \frac{\tau^4}{4},$$

and

$$I'_{ff,3}(x) = \iint \sum_{j=0}^{1} \left\{\nabla_x^{1-j}\rho(x - \tau u_1) - \nabla_x^{1-j}\rho(x - \tau u_2)\right\} \nabla_x^j T_{ff,3}(x)\, du_1 du_2 = O(\tau^4),$$

where the second inequality follows from Assumption 2.5, and boundedness of ρ and $f_{X^*|Z}$ (and their derivatives). Hence we conclude $I_{ff,3}(x) = O_x(\tau^4)$.

3.2.

$$T_{ff,1} \equiv \left(f_{\xi,X^*|Z}\left(u_1, x-\tau u_1|z_1\right) + f'_{\xi,X^*|Z}\left(u_1, x|z_1\right)\tau u_1 - f_{\xi,X^*|Z}\left(u_1, x|z_1\right)\right) \times \left(-f'_{\xi,X^*|Z}\left(u_2, x|z_2\right)\tau u_2 + f_{\xi,X^*|Z}\left(u_2, x|z_2\right)\right).$$

Thus,

$$\begin{aligned} I_{ff,1}(x) &= \iint \left\{\rho\left(x-\tau u_1\right) - \rho\left(x-\tau u_2\right)\right\} \\ &\quad \times \left(\frac{\tau^2}{2} f''_{\xi,X^*|Z}\left(u_1, x|z_1\right) u_1^2 - \int_0^\tau \frac{1}{2} f'''_{\xi,X^*|Z}\left(u_1, x-tu_1|z_1\right) u_1^3 \left(\tau-t\right)^2 dt\right) \\ &\quad \times \left(-f'_{\xi,X^*|Z}\left(u_2, x|z_2\right)\tau u_2 + f_{\xi,X^*|Z}\left(u_2, x|z_2\right)\right) du_1 du_2 \\ &= \frac{\tau^2}{2}\left(-\tau I_{\Delta(x,2,1)} + I_{\Delta(x,2,0)}\right) + I_{ff,1,2}(x), \end{aligned}$$

where the remainder is

$$\begin{aligned} I_{ff,1,2}(x) &= -\iint \left\{\rho\left(x-\tau u_1\right) - \rho\left(x-\tau u_2\right)\right\} \left(\int_0^\tau \frac{1}{2} f'''_{\xi,X^*|Z}\left(u_1, x-tu_1|z_1\right) u_1^3 \left(\tau-t\right)^2 dt\right) \\ &\quad \times \left(-f'_{\xi,X^*|Z}\left(u_2, x|z_2\right)\tau u_2 + f_{\xi,X^*|Z}\left(u_2, x|z_2\right)\right) du_1 du_2. \end{aligned}$$

By an argument similar to the one provided in part 3.1 above,

$$\begin{aligned} I_{ff,1,2,1}(x) &= \iint \left\{\rho\left(x-\tau u_1\right) - \rho\left(x-\tau u_2\right)\right\} \left(\int_0^\tau \frac{1}{2} f'''_{\xi,X^*|Z}\left(u_1, x-tu_1|z_1\right) u_1^3 \left(\tau-t\right)^2 dt\right) \\ &\quad \times f'_{\xi,X^*|Z}\left(u_2, x|z_2\right)\tau u_2 du_1 du_2 \\ &= O_x\left(\tau^4\right). \end{aligned}$$

Next,

$$\begin{aligned} I_{ff,1,2,2}(x) &= -\iint \left\{\rho\left(x-\tau u_1\right) - \rho\left(x-\tau u_2\right)\right\} \left(\int_0^\tau \frac{1}{2} f'''_{\xi,X^*|Z}\left(u_1, x-tu_1|z_1\right) u_1^3 \left(\tau-t\right)^2 dt\right) \\ &\quad \times f_{\xi,X^*|Z}\left(u_2, x|z_2\right) du_1 du_2 \\ &= -\iint \left(\sum_{j=1}^{2} (-1)^j \int_0^\tau \rho'(x-tu_j) u_j dt\right) \left(\int_0^\tau \frac{1}{2} f'''_{\xi,X^*|Z}\left(u_1, x-tu_1|z_1\right) u_1^3 \left(\tau-t\right)^2 dt\right) \\ &\quad \times f_{\xi,X^*|Z}\left(u_2, x|z_2\right) du_1 du_2 \\ &= O_x\left(\tau^4\right), \end{aligned}$$

where the second equality follows from equation (A.9a), and the last follows from an argument similar to the one provided in part 3.1. Then, we conclude

$$I_{ff,1,2}(x) = I_{ff,1,2,1}(x) + I_{ff,1,2,2}(x) = O_x(\tau^4).$$

Next, note that $I_{h(x,0,1,2,0)} = 0$ and $I_{h(x,1,0,2,0)} = 0$ since $\mathbb{M}_{|\zeta_1\zeta_2 z_1 z_2}[U_1^2 U_2] = 0$ and $\mathbb{M}_{|\zeta_1\zeta_2 z_1 z_2}[U_1^3] = 0$, respectively. Thus, using equation (A.12)

$$\begin{aligned} I_{ff,1}(x) &= \frac{\tau^2}{2}\left(-\tau I_{\Delta(x,2,1)} + I_{\Delta(x,2,0)}\right) + O_x\left(\tau^4\right) = \tau^3 O_x\left(\tau\right) + \frac{\tau^2}{2} I_{\Delta(x,2,0)} + O_x\left(\tau^4\right) \\ &= \frac{\tau^2}{2}\left(\tau\rho'\left(x\right)\left\{I_{h(x,0,1,2,0)} - I_{h(x,1,0,2,0)}\right\} + O_x\left(\tau^2\right)\right) + O_x\left(\tau^4\right) \\ &= O_x\left(\tau^4\right). \end{aligned}$$

Similarly,

$$I_{ff,2}\left(x\right) = O_x\left(\tau^4\right).$$

3.3 The main term is

$$\begin{aligned} I_{ff,4}\left(x\right) &\equiv \iint \left\{\rho\left(x - \tau u_1\right) - \rho\left(x - \tau u_2\right)\right\}\left(-f'_{\xi,X^*|Z}\left(u_1, x|z_1\right)\tau u_1 + f_{\xi,X^*|Z}\left(u_1, x|z_1\right)\right) \\ &\qquad \times \left(-f'_{\xi,X^*|Z}\left(u_2, x|z_2\right)\tau u_2 + f_{\xi,X^*|Z}\left(u_2, x|z_2\right)\right) du_1 du_2 \\ &= \tau^2 I_{\Delta(x,1,1)} - \tau I_{\Delta(x,1,0)} - \tau I_{\Delta(x,0,1)} + I_{\Delta(x,0,0)} \\ &= -\tau\left(I_{\Delta(x,1,0)} + I_{\Delta(x,0,1)}\right), \end{aligned}$$

where the last equality holds because $I_{\Delta(x,0,0)} = I_{\Delta(x,1,1)} = 0$ due to the symmetry. Using equation (A.11) we have

$$I_{\Delta(x,1,0)} = \tau\rho'\left(x\right)\left\{I_{h(x,0,1,1,0)} - I_{h(x,1,0,1,0)}\right\} + \frac{\tau^2}{2}\rho''\left(x\right)\left\{I_{h(x,2,0,1,0)} - I_{h(x,0,2,1,0)}\right\} + O_x\left(\tau^3\right).$$

Using equation (A.10),

$$
\begin{aligned}
I_{h(x,0,1,1,0)} &= \nabla_{\zeta_1}\mathbb{M}_{|\zeta_1\zeta_2z_1z_2}\left[U_1U_2\right]\Big|_{\zeta_1=\zeta_2=x}=0,\\
I_{h(x,1,0,1,0)} &= \nabla_{\zeta_1}\mathbb{M}_{|\zeta_1\zeta_2z_1z_2}\left[U_1^2\right]\Big|_{\zeta_1=\zeta_2=x}=\nabla_{\zeta_1}\left(E[\xi_i^2|X_i^*=\zeta_1]f_{X^*|Z}\left(\zeta_1|z_1\right)f_{X^*|Z}\left(\zeta_2|z_2\right)\right)\Big|_{\zeta_1=\zeta_2=x}\\
&= \left[E[\xi_i^2|X_i^*=x]f'_{X^*|Z}\left(x|z_1\right)+\nabla_xE[\xi_i^2|X_i^*=x]f_{X^*|Z}\left(x|z_1\right)\right]f_{X^*|Z}\left(x|z_2\right),\\
I_{h(x,2,0,1,0)} &= \nabla_{\zeta_1}\mathbb{M}_{|\zeta_1\zeta_2z_1z_2}\left[U_1^3\right]\Big|_{\zeta_1=\zeta_2=x}=\nabla_{\zeta_1}0\Big|_{\zeta_1=\zeta_2=x}=0,\\
I_{h(x,0,2,1,0)} &= \nabla_{\zeta_1}\mathbb{M}_{|\zeta_1\zeta_2z_1z_2}\left[U_1U_2^2\right]\Big|_{\zeta_1=\zeta_2=x}=0,
\end{aligned}
$$

and hence

$$
I_{\Delta(x,1,0)}=-\tau\rho'\left(x\right)\left[E[\xi_i^2|X_i^*=x]f'_{X^*|Z}\left(x|z_1\right)+\nabla_xE[\xi_i^2|X_i^*=x]f_{X^*|Z}\left(x|z_1\right)\right]f_{X^*|Z}\left(x|z_2\right)+O_x\left(\tau^3\right).
$$

Similarly,

$$
I_{\Delta(x,0,1)}=\tau\rho'\left(x\right)\left[E[\xi_i^2|X_i^*=x]f'_{X^*|Z}\left(x|z_2\right)+\nabla_xE[\xi_i^2|X_i^*=x]f_{X^*|Z}\left(x|z_2\right)\right]f_{X^*|Z}\left(x|z_1\right)+O_x\left(\tau^3\right),
$$

and hence

$$
\begin{aligned}
I_{ff,4}\left(x\right) &= -\tau\left(I_{\Delta(x,1,0)}+I_{\Delta(x,0,1)}\right)\\
&= \rho'\left(x\right)v\left(x\right)\left(f'_{X^*|Z}\left(x|z_1\right)f_{X^*|Z}\left(x|z_2\right)-f_{X^*|Z}\left(x|z_1\right)f'_{X^*|Z}\left(x|z_2\right)\right)+O_x\left(\tau^4\right),
\end{aligned}
$$

where we used $v(x)=\tau^2E[\xi_i^2|X_i^*=x]$.
4. Putting all $I_{ff,j}\left(x\right)$ together we have

$$
\begin{aligned}
&N\left(x,z_1,z_2,\tau\right)\\
=\;& \rho'\left(x\right)v\left(x\right)\left(f'_{X^*|Z}\left(x|z_1\right)f_{X^*|Z}\left(x|z_2\right)-f_{X^*|Z}\left(x|z_1\right)f'_{X^*|Z}\left(x|z_2\right)\right)+O_x\left(\tau^4\right)\\
=\;& \rho'\left(x\right)v\left(x\right)\left(s_{X^*|Z}\left(x|z_1\right)-s_{X^*|Z}\left(x|z_2\right)\right)f_{X^*|Z}\left(x|z_1\right)f_{X^*|Z}\left(x|z_2\right)+O_x\left(\tau^4\right).
\end{aligned}
$$

Moreover, Lemma A.2 establishes $f_{X|Z}(x,z_j)=f_{X^*|Z}(x|z_j)+O_x(\tau^2)$ for $j\in\{1,2\}$. Hence, the denominator $D\left(x,z_1,z_2,\tau\right)=f_{X|Z}\left(x|z_1\right)f_{X|Z}\left(x|z_2\right)$ satisfies

$$
D\left(x,z_1,z_2,\tau\right)=f_{X^*|Z}\left(x|z_1\right)f_{X^*|Z}\left(x|z_2\right)+O_x(\tau^2).
$$

Thus,

$$\begin{aligned}
&\frac{N(x,z_1,z_2,\tau)}{D(x,z_1,z_2,\tau)}\\
= &\frac{\rho'(x)\,v(x)\,f_{X^*|Z}(x|z_1)\,f_{X^*|Z}(x|z_2)\left(s_{X^*|Z}(x|z_1)-s_{X^*|Z}(x|z_2)\right)+O_x\left(\tau^4\right)}{f_{X^*|Z}(x|z_1)\,f_{X^*|Z}(x|z_2)+O_x\left(\tau^2\right)}\\
= &\rho'(x)\,v(x)\left(s_{X^*|Z}(x|z_1)-s_{X^*|Z}(x|z_2)\right)+O_x\left(\tau^4\right),
\end{aligned}$$

which completes the proof. □

A.1.2 Proof of the main result

Equipped with Lemmas A.1-A.3, we are ready to prove Theorem 1.

Proof of Theorem 1. Using equation (A.1) we have

$$q(x,z_1)-q(x,z_2)=v(x)\,\rho'(x)\left[s_{X^*|Z}(x|z_1)-s_{X^*|Z}(x|z_2)\right]+O\left(\tau^p\right). \tag{A.13}$$

Combined with (A.2), the above implies that

$$q(x,z_1)-q(x,z_2)=v(x)\,q'(x)\left[s_{X|Z}(x|z_1)-s_{X|Z}(x|z_2)\right]+O\left(\tau^p\right).$$

Hence, for any x with $\rho'(x)\left[s_{X^*|Z}(x|z_1)-s_{X^*|Z}(x|z_2)\right]\neq 0$ (or x such that $\left|q'(x)\left[s_{X|Z}(x|z_1)-s_{X|Z}(x|z_2)\right]\right|>C>0$),

$$\widetilde{v}(x)=v(x)+O\left(\tau^p\right),\qquad \text{where } \widetilde{v}(x)\equiv\frac{q(x,z_1)-q(x,z_2)}{q'(x)\left[s_{X|Z}(x|z_1)-s_{X|Z}(x|z_2)\right]}. \tag{A.14}$$

Next by Lemma A.3, we have

$$q(x,z_1)-q(x,z_2)=\tau^2\rho'(x)\,v(x)\left(s_{X^*|Z}(x|z_1)-s_{X^*|Z}(x|z_2)\right)+O_x\left(\tau^p\right).$$

By Lemma A.2, we also have $q'(x)=\rho'(x)+O_x(\tau^{p-2})$ and $s_{X|Z}(x|z)=s_{X^*|Z}(x|z)+O_x(\tau^{p-2})$. Hence,

$$q(x,z_1)-q(x,z_2)=q'(x)\,v(x)\left(s_{X|Z}(x|z_1)-s_{X|Z}(x|z_2)\right)+O_x\left(\tau^p\right).$$

This implies that

$$q'(x,z_1) - q'(x,z_2) = \nabla_x \left[q'(x)\, v(x) \left(s_{X|Z}(x|z_1) - s_{X|Z}(x|z_2)\right)\right] + O(\tau^p).$$

Hence, $v'(x)$ satisfies

$$\begin{aligned} v'(x) &= \frac{q'(x,z_1) - q'(x,z_2)}{q'(x)\left[s_{X|Z}(x|z_1) - s_{X|Z}(x|z_2)\right]} - v(x)\frac{\nabla_x\left(q'(x)\left[s_{X|Z}(x|z_1) - s_{X|Z}(x|z_2)\right]\right)}{q'(x)\left[s_{X|Z}(x|z_1) - s_{X|Z}(x|z_2)\right]} + O(\tau^p) \\ &= \underbrace{\frac{q'(x,z_1) - q'(x,z_2)}{q'(x)\left[s_{X|Z}(x|z_1) - s_{X|Z}(x|z_2)\right]} - \widetilde{v}(x)\frac{\nabla_x\left(q'(x)\left[s_{X|Z}(x|z_1) - s_{X|Z}(x|z_2)\right]\right)}{q'(x)\left[s_{X|Z}(x|z_1) - s_{X|Z}(x|z_2)\right]}}_{=\widetilde{v}'(x)} + O(\tau^p), \end{aligned}$$

where the second equality uses equation (A.14). Hence, we conclude

$$\widetilde{v}'(x) = v'(x) + O(\tau^p). \tag{A.15}$$

Next, notice that since $q'(x) = \rho'(x) + O_x(\tau^{p-2})$ (Lemma A.2), we also have

$$|q''(x) - \rho''(x)| = O(\tau^{p-2}). \tag{A.16}$$

Finally, recall that (A.1) implies

$$\begin{aligned} \rho(x) &= q(x,z_1) - v(x)\left[\rho'(x)\, s_{X^*|Z}(x|z) + \tfrac{1}{2}\rho''(x)\right] - v'(x)\,\rho'(x) + O(\tau^p), \\ &= q(x,z_1) - v(x)\left[q'(x)\, s_{X|Z}(x|z) + \tfrac{1}{2}q''(x)\right] - v'(x)\, q'(x) + O(\tau^p) \\ &= q(x,z_1) - \widetilde{v}(x)\left[q'(x)\, s_{X|Z}(x|z) + \tfrac{1}{2}q''(x)\right] - \widetilde{v}'(x)\, q'(x) + O(\tau^p) \\ &= \widetilde{\rho}(x,z_1) + O(\tau^p), \end{aligned}$$

where the second equality uses $v(x) = O(\tau^2)$, $v'(x) = O(\tau^2)$, and equations (A.2) and (A.16), and the third equality uses (A.14) and (A.15). □

A.2 Proofs of Lemma 3

Proof of Lemma 3. First, consider $\int \eta(r) f_{\varepsilon|X^*}(x-r|r)dr = \int \eta(x-\tau u) f_{\xi|X^*}(u|x-\tau u)du$ where $\eta(\cdot)$ is a bounded function with $m \geq p$ bounded derivatives with respect to x. Then, notice that Assumptions 2.4 and 2.5 ensure that $\int \eta(x-\tau u) f_{\xi|X^*}(u|x-\tau u)du$ has m bounded derivatives with respect to x. Hence,

$f_{X|Z}(x|z) = \int f_{X^*|Z}(r|z) f_{\varepsilon|X^*}(x-r|r)dr$, $f_X(x) = \int f_{X^*}(r) f_{\varepsilon|X^*}(x-r|r)dr$, $q(x,z) = \int \rho(r) f_{X^*|Z}(r|z) f_{\varepsilon|X^*}(x-r|r)dr / f_{X|Z}(x|z)$, and $q(x) = \int \rho(r) f_{X^*}(r) f_{\varepsilon|X^*}(x-r|r)dr/f_X(x)$ have 4 bounded derivatives with respect to x (for $q(x,z)$ and $q(x)$, at least, in a neighborhood of x where $f_{X^*|Z}(x|z)$ and $f_{X^*}(z)$ are bounded away from zero).

If the tuning parameters are chosen optimally, under the usual regularity conditions, the rates of convergence of these estimators are $\widehat{q}(x,z) - q(x,z) = O_p\left(n^{-\frac{m}{2m+1}}\right)$, $\widehat{q}(x) - q(x) = O_p\left(n^{-\frac{m}{2m+1}}\right)$, $\widehat{q}'(x,z) - q'(x,z) = O_p\left(n^{-\frac{m-1}{2m+1}}\right)$, $\widehat{q}'(x) - q'(x) = O_p\left(n^{-\frac{m-1}{2m+1}}\right)$, $\widehat{q}''(x) - q''(x) = O_p\left(n^{-\frac{m-2}{2m+1}}\right)$, $\widehat{s}_{X|Z}(x|z) - s_{X|Z}(x|z) = O_p\left(n^{-\frac{m-1}{2m+1}}\right)$, and $\widehat{s}'_{X|Z}(x|z) - s'_{X|Z}(x|z) = O_p\left(n^{-\frac{m-2}{2m+1}}\right)$ for $x \in S_{X^*}(z)$ where $\widehat{s}_{X|Z}(x|z) \equiv \widehat{f}'_{X|Z}(x|z) \big/ \widehat{f}_{X|Z}(x|z)$ (see, e.g., Stone, 1980).

First, note that

$$\begin{aligned}
\widehat{v}(x) &= \frac{\widehat{q}(x,z_1) - \widehat{q}(x,z_2)}{\widehat{q}'(x)\left[\widehat{s}_{X|Z}(x|z_1) - \widehat{s}_{X|Z}(x|z_2)\right]} \\
&= \underbrace{\frac{q(x,z_1) - q(x,z_2)}{q'(x)\left[s_{X|Z}(x|z_1) - s_{X|Z}(x|z_2)\right]}}_{=\widetilde{v}(x)} + O_p\left(n^{-\frac{m}{2m+1}} + \tau_n^2 n^{-\frac{m-1}{2m+1}}\right) \qquad \text{(A.17)}
\end{aligned}$$

Here the second equality follows from $\widehat{a}/\widehat{b} - a/b = (\widehat{a} - a)/\widehat{b} + a(1/\widehat{b} - 1/b)$, $q(x,z_1) - q(x,z_2) = O(\tau_n^2)$ (implied by (A.13)), and the rates of convergence listed above.

Next,

$$\widehat{v}'(x) = \frac{\widehat{q}'(x,z_1) - \widehat{q}'(x,z_2)}{\widehat{q}'(x)\left[\widehat{s}_{X|Z}(x|z_1) - \widehat{s}_{X|Z}(x|z_2)\right]} - \widehat{v}(x)\frac{\nabla_x\left(\widehat{q}'(x)\left[\widehat{s}_{X|Z}(x|z_1) - \widehat{s}_{X|Z}(x|z_2)\right]\right)}{\widehat{q}'(x)\left[\widehat{s}_{X|Z}(x|z_1) - \widehat{s}_{X|Z}(x|z_2)\right]}.$$

Analogously to the derivation provided above, we inspect that

$$\frac{\widehat{q}'(x,z_1) - \widehat{q}'(x,z_2)}{\widehat{q}'(x)\left[\widehat{s}_{X|Z}(x|z_1) - \widehat{s}_{X|Z}(x|z_2)\right]} = \frac{q'(x,z_1) - q'(x,z_2)}{q'(x)\left[s_{X|Z}(x|z_1) - s_{X|Z}(x|z_2)\right]} + O_p\left(n^{-\frac{m-1}{2m+1}}\right),$$

and

$$\frac{\nabla_x\left(\widehat{q}'(x)\left[\widehat{s}_{X|Z}(x|z_1) - \widehat{s}_{X|Z}(x|z_2)\right]\right)}{\widehat{q}'(x)\left[\widehat{s}_{X|Z}(x|z_1) - \widehat{s}_{X|Z}(x|z_2)\right]} = \frac{\nabla_x\left(q'(x)\left[s_{X|Z}(x|z_1) - s_{X|Z}(x|z_2)\right]\right)}{q'(x)\left[s_{X|Z}(x|z_1) - s_{X|Z}(x|z_2)\right]} + O_p\left(n^{-\frac{m-2}{2m+1}}\right).$$

Combining the latter with (A.17), we obtain

$$\widehat{v}\left(x\right)\frac{\nabla_x\left(\widehat{q}'\left(x\right)\left[\widehat{s}_{X|Z}\left(x|z_1\right)-\widehat{s}_{X|Z}\left(x|z_2\right)\right]\right)}{\widehat{q}'\left(x\right)\left[\widehat{s}_{X|Z}\left(x|z_1\right)-\widehat{s}_{X|Z}\left(x|z_2\right)\right]} = \widetilde{v}\left(x\right)\frac{\nabla_x\left(q'\left(x\right)\left[s_{X|Z}\left(x|z_1\right)-s_{X|Z}\left(x|z_2\right)\right]\right)}{q'\left(x\right)\left[s_{X|Z}\left(x|z_1\right)-s_{X|Z}\left(x|z_2\right)\right]} + O_p\left(n^{-\frac{m}{2m+1}}+\tau_n^2 n^{-\frac{m-2}{2m+1}}\right),$$

where we also used $\widetilde{v}(x) = O(\tau_n^2)$. Hence, we conclude

$$\widehat{v}'(x) = \widetilde{v}'(x) + O_p\left(n^{-\frac{m-1}{2m+1}}+\tau_n^2 n^{-\frac{m-2}{2m+1}}\right). \tag{A.18}$$

Finally,

$$\begin{aligned}\widehat{\rho}(x) &= \widehat{q}(x,z_1) - \widehat{v}\left(x\right)\left[\widehat{q}'\left(x\right)\widehat{s}_{X|Z}\left(x|z_1\right)+\tfrac{1}{2}\widehat{q}''\left(x\right)\right] - \widehat{v}'\left(x\right)\widehat{q}'\left(x\right)\\ &= \widetilde{\rho}(x,z_1) + O_p\left(n^{-\frac{m-1}{2m+1}}+\tau_n^2 n^{-\frac{m-2}{2m+1}}\right)\\ &= \rho(x) + O_p\left(n^{-\frac{m-1}{2m+1}}+\tau_n^2 n^{-\frac{m-2}{2m+1}}+\tau_n^p\right),\\ &= \rho(x) + O_p\left(n^{-\frac{m-1}{2m+1}}+\tau_n^p\right).\end{aligned}$$

Here, the second equality uses the definition of $\widetilde{\rho}(x,z_1)$, and equations (A.17) and (A.18). The third equality follow from Theorem 1. For the last equality, note that $\tau_n^2 n^{-\frac{m-2}{2m+1}} \lesssim n^{-\frac{m-1}{2m+1}}$ for $\tau_n \lesssim n^{-\frac{1}{2(2m+1)}}$ and $\tau_n^2 n^{-\frac{m-2}{2m+1}} \lesssim \tau_n^p$ for $\tau_n \gtrsim n^{-\frac{m-2}{(p-2)(2m+1)}}$, and $n^{-\frac{1}{2(2m+1)}} > n^{-\frac{m-2}{(p-2)(2m+1)}}$ since $2m-p-2>0$.

For the naive estimator, we have

$$\widehat{\rho}^{\text{Naive}}\left(x\right) = \widehat{q}\left(x\right) = \rho\left(x\right) + O_p\left(n^{-\frac{m}{2m+1}}+\tau_n^2\right),$$

where the second equality uses $\widehat{q}(x) = q(x) + O\left(n^{-\frac{m}{2m+1}}\right)$ and $q(x) = \rho(x) + O(\tau_n^2)$ (e.g., (A.1)). □

A.3 Proof of Theorem 4

Before proving Theorem 4, in Section A.3.1, we first demonstrate that the CDF $F_{X^*}(\cdot)$ and the quantile function $Q_{X^*}(\cdot)$ of X_i^* are identified up to an error of order $O(\tau^p)$. The proof of Theorem 4 is then provided in Section A.3.2.

A.3.1 Identification of $F_{X^*}(\cdot)$ and $Q_{X^*}(\cdot)$

Lemma A.4. *Suppose that Assumptions 2.3-2.5 and 3.1 are satisfied. Suppose either (i) $p=3$, or (ii) $E\left[\xi_i^3|X_i^*\right]=0$ and $p=4$. Then, for any $x \in supp\left(X_i^*\right)$ and for any $s \in (0,1)$ such that $f_{X^*}(Q_{X^*}(s))>0$, as $\tau \to 0$,*

$$\begin{aligned} F_{X^*}(x) &= F_X(x) - \frac{1}{2}\nabla_x\left(f_X(x)v(x)\right) + O\left(\tau^p\right), \\ Q_{X^*}(s) &= Q_X(s) + \frac{1}{2}\left\{s_X\left(Q_X(s)\right)v\left(Q_X(s)\right) + \nabla_x v\left(Q_X(s)\right)\right\} + O\left(\tau^p\right). \end{aligned}$$

Remark A.1. *Assumptions 2.3 and 2.4 are stronger than needed for the result of Lemma A.4 to hold. For Assumption 2.3, the requirement $\xi \perp Z_i|X_i^*$ can be dropped. Assumption 2.4 can be replaced by requiring the densities $f_{X^*}(x)$ and $f_{\xi|X^*}(\xi|x)$ to be bounded functions that have at least $p \geqslant 3$ bounded derivatives (all with respect to x).*

Lemma A.4 extends the results obtained earlier for CME (e.g., Chesher, 1991) to the WCME model. It also provides a way of recovering $F_{X^*}(x)$ and $Q_{X^*}(s)$ up to an error of order $O(\tau^p)$ using

$$\begin{aligned} \tilde{F}_{X^*}(x) &\equiv F_X(x) - \frac{1}{2}\nabla_x\left(f_X(x)\tilde{v}(x)\right), \\ \tilde{Q}_{X^*}(s) &\equiv Q_X(s) + \frac{1}{2}\left\{s_X\left(Q_X(s)\right)\tilde{v}\left(Q_X(s)\right) + \nabla_x\tilde{v}\left(Q_X(s)\right)\right\}, \end{aligned}$$

where $\tilde{v}(\cdot)$ is given by (3).

Proof of Lemma A.4. In the proof of Lemma A.1, we demonstrated that, for a bounded function $\eta(\cdot)$ with p bounded derivatives, we have

$$\begin{aligned} \int \eta(r) f_{\varepsilon|X^*}(x-r|r)\,dr &= \int \eta(x-\tau u) f_{\xi|X^*}(u|x-\tau u)\,du \\ &= \eta(x) + \frac{1}{2}\nabla_x^2\left\{\eta(x)v(x)\right\} + R_\eta(x;\tau). \end{aligned}$$

Here the remainder $R_\eta(x;\tau)$ can be represented in the integral form as

$$R_\eta(x;\tau) = \int\left(\int_0^\tau \frac{(-1)^p}{(p-1)!}u^p\nabla_x^p\{\eta(x-tu)f_{\xi|X^*}(u|x-tu)\}(\tau-t)^{p-1}dt\right)du.$$

Then, for $f_X(x) = \int f_{X^*}(r) f_{\varepsilon|X^*}(x-r|r)\,dr$ with $\eta(\cdot) = f_{X^*}(\cdot)$, we have

$$f_X(x) = f_{X^*}(x) + \frac{1}{2}\nabla_x^2 (f_{X^*}(x)\,v(x)) + R_{f_{X^*}}(x;\tau).$$

Then,

$$\begin{aligned}
F_X(\bar{x}) &= \int_{-\infty}^{\bar{x}} f_X(x)dx \\
&= \int_{-\infty}^{\bar{x}} \left(f_{X^*}(x) + \frac{1}{2}\nabla_x^2 (f_{X^*}(x)\,v(x)) + R_{f_{X^*}}(x;\tau) \right) dx \\
&= F_{X^*}(x) + \frac{1}{2}\nabla_x (f_{X^*}(\bar{x})\,v(\bar{x})) + \int_{-\infty}^{\bar{x}} R_{f_{X^*}}(x;\tau)dx.
\end{aligned}$$

Here we used

$$\int_{-\infty}^{\bar{x}} \nabla_x^2 (f_{X^*}(x)\,v(x)) = \nabla_x (f_{X^*}(\bar{x})\,v(\bar{x}))$$

since $\lim_{x\to-\infty} \nabla_x (f_{X^*}(x)\,v(x)) = 0$ because $v(x) = \tau^2 E[\xi^2|X_i^* = x]$, and $E[\xi^2|X_i^* = x]$ and $\nabla_x E[\xi^2|X_i^* = x]$ are bounded under Assumption 2.5. The remainder takes the form

$$\int_{-\infty}^{\bar{x}} R_{f_{X^*}}(x;\tau)dx = \int_{-\infty}^{\bar{x}} \left\{ \int \left(\int_0^{\tau} \varphi(x,u,t)dt \right) du \right\} dx,$$

where

$$\varphi(x,u,t) \equiv \frac{(-1)^p}{(p-1)!} u^p \nabla_x^p \{ f_{X^*}(x-tu) f_{\xi|X^*}(u|x-tu) \} (\tau-t)^{p-1}.$$

Next, note that, using Assumption 3.1, for all u and $t \in [0,\tau]$

$$\int |\varphi(x,u,t)|\,dx \le C\,|u|^p\,(\tau-t)^{p-1} \sup_{\tilde{x}\in\mathcal{S}_X} \sum_{\ell=0}^{p} \left|\nabla_x^p f_{\xi|X^*}(u|\tilde{x})\right|,$$

where $C > 0$ is a generic constant. Next, using Assumption 2.5,

$$\int \left(\int_0^{\tau} |u|^p\,(\tau-t)^{p-1} \sup_{\tilde{x}\in\mathcal{S}_X} \sum_{\ell=0}^{p} \left|\nabla_x^p f_{\xi|X^*}(u|\tilde{x})\right| dt \right) du \le C\tau^p.$$

Hence, using Fubini–Tonelli's theorem, we conclude that

$$\iiint\limits_{(-\infty,\bar{x}]\times\mathbb{R}\times[0,\tau]} |\phi(x,u,t)|\, dxdudt \le C\tau^p,$$

and the order of integration can be interchanged. Specifically, it also implies that

$$\int_{-\infty}^{\bar{x}} R_{f_{X^*}}(x;\tau)dx \le C\tau^p.$$

Hence, we conclude

$$F_X(\bar{x}) = F_{X^*}(\bar{x}) + \frac{1}{2}\nabla_x \left(f_{X^*}\left(\bar{x}\right) v\left(\bar{x}\right)\right) + O(\tau^p),$$

so

$$\begin{aligned} F_{X^*}(\bar{x}) &= F_X(\bar{x}) - \frac{1}{2}\nabla_x \left(f_{X^*}\left(\bar{x}\right) v\left(\bar{x}\right)\right) + O(\tau^p) \\ &= F_X(\bar{x}) - \frac{1}{2}\left(f'_{X^*}\left(\bar{x}\right) v\left(\bar{x}\right) + f_{X^*}\left(\bar{x}\right) v'\left(\bar{x}\right)\right) + O(\tau^p) \\ &= F_X(\bar{x}) - \frac{1}{2}\left(f'_{X}\left(\bar{x}\right) v\left(\bar{x}\right) + f_{X}\left(\bar{x}\right) v'\left(\bar{x}\right)\right) + O(\tau^p) \\ &= F_X(\bar{x}) - \frac{1}{2}\nabla_x \left(f_{X}\left(\bar{x}\right) v\left(\bar{x}\right)\right) + O(\tau^p), \end{aligned} \tag{A.19}$$

where we used $f_X(\bar{x}) = f_{X^*}(\bar{x}) + O(\tau^2)$ and $f'_X(\bar{x}) = f'_{X^*}(\bar{x}) + O(\tau^2)$ (established the the proof of Lemma A.2). This completes the proof of the first part.

Next, note that

$$\begin{aligned} Q_{X^*}(s) - Q_X(s) &= Q_{X^*}(s) - Q_{X^*}\left(F_{X^*}\left(Q_X(s)\right)\right) \\ &= Q'_{X^*}(s)\left(F_X\left(Q_X(s)\right) - F_{X^*}\left(Q_X(s)\right)\right) + \frac{1}{2}Q''_{X^*}(\tilde{s})\left(F_X\left(Q_X(s)\right) - F_{X^*}\left(Q_X(s)\right)\right)^2, \end{aligned}$$

where $Q'_{X^*}(s) = \frac{1}{f'_{X^*}(Q_{X^*}(s))}$ and $Q''_{X^*}(s) = -\frac{f''_{X^*}(Q_{X^*}(s))}{\left(f'_{X^*}(Q_{X^*}(s))\right)^3}$, and $\tilde{s}$ lies between $s = F_X(Q_X(s))$ and $F_{X^*}\left(Q_X(s)\right)$. Recall that (A.19) implies

$$\begin{aligned} F_X(Q_x(s)) - F_{X^*}(Q_X(s)) &= \frac{1}{2}\nabla_x \left\{f_X(x)v(x)\right\}\big|_{x=Q_x(s)} + O(\tau^p) \\ &= O(\tau^2). \end{aligned}$$

Hence, we conclude $Q_{X^*}(s) - Q_X(s) = O(\tau^2)$. This implies

$$\begin{aligned}\frac{1}{f'_{X^*}(Q_{X^*}(s))} &= \frac{1}{f'_{X^*}(Q_X(s))} + O(\tau^2)\\ &= \frac{1}{f'_X(Q_X(s))} + O(\tau^2).\end{aligned}$$

Thus, we obtain

$$\begin{aligned}Q_{X^*}(s) - Q_X(s) &= \frac{1}{f'_X\left(Q_X(s)\right)}\,\frac{1}{2}\nabla_x\left\{f_X(x)v(x)\right\}\big|_{x=Q_x(s)} + O(\tau^p)\\ &= \frac{1}{2}\left\{s_X\left(Q_X\left(s\right)\right)v\left(Q_X\left(s\right)\right) + \nabla_x v\left(Q_X\left(s\right)\right)\right\} + O(\tau^p).\end{aligned}$$

□

A.3.2 Proof of Theorem 4

Proof of Theorem 4. First, in the proof of Theorem 1, we show that

$$\begin{aligned}\tilde{v}\left(Q_{X^*}\left(F_{\mathcal{X}^*}(\varkappa)\right)\right) &= v\left(Q_{X^*}\left(F_{\mathcal{X}^*}(\varkappa)\right)\right) + O(\tau^p),\\ \tilde{v}'\left(Q_{X^*}\left(F_{\mathcal{X}^*}(\varkappa)\right)\right) &= v'\left(Q_{X^*}\left(F_{\mathcal{X}^*}(\varkappa)\right)\right) + O(\tau^p).\end{aligned}$$

Combining the above with the result of Lemma A.4, we establish

$$\tilde{Q}_{X^*}(F_{\mathcal{X}^*}(\varkappa)) - Q_{X^*}\left(F_{\mathcal{X}^*}(\varkappa)\right) = O(\tau^p). \tag{A.20}$$

Since $\rho'(\cdot)$ is bounded, we also have

$$\left|\rho(\tilde{Q}_{X^*}(F_{\mathcal{X}^*}(\varkappa))) - \rho\left(Q_{X^*}\left(F_{\mathcal{X}^*}(\varkappa)\right)\right)\right| = O(\tau^p). \tag{A.21}$$

Second, there exists $\delta > 0$ such that $f_{X^*|Z}(x|z_1)$ and $f_{X^*}(x|z_2)$ are bounded away from zero for $x \in B_\delta\left(Q_{X^*}\left(F_{\mathcal{X}^*}(\varkappa)\right)\right)$. Since all the remainders such as $O(\cdot)$ and $O_x(\cdot)$ are explicitly bounded in the proof of Theorem 1 and those bounds are uniform in $x \in \mathcal{B}_\delta(Q_{X^*}(F_{\mathcal{X}^*}(\varkappa)))$, we have

$$\sup_{x\in B_\delta(Q_{X^*}(F_{\mathcal{X}^*}(\varkappa)))} |\tilde{\rho}_{X^*}(x, z_1) - \rho_{X^*}(x)| = O(\tau^p). \tag{A.22}$$

Finally,

$$
\begin{aligned}
|\tilde{\rho}_{\mathcal{X}^*}(\varkappa, z_1) - \rho_{\mathcal{X}^*}(\varkappa)| = & \left|\tilde{\rho}_{X^*}(\tilde{Q}_{X^*}\left(F_{\mathcal{X}^*}(\varkappa)\right), z_1) - \rho_{X^*}\left(Q_{X^*}\left(F_{\mathcal{X}^*}(\varkappa)\right)\right)\right| \\
\leq & \sup_{x \in B_\delta(Q_{X^*}(F_{\mathcal{X}^*}(\varkappa)))} |\tilde{\rho}_{X^*}(x, z_1) - \rho_{X^*}(x)| \\
& + \left|\rho(\tilde{Q}_{X^*}(F_{\mathcal{X}^*}(\varkappa))) - \rho\left(Q_{X^*}\left(F_{\mathcal{X}^*}(\varkappa)\right)\right)\right|,
\end{aligned}
$$

where the inequality holds for sufficiently small values of τ due to (A.20). Combining (A.21) and (A.22) together completes the proof. □

B Verification of Assumption 2.5

Lemma B.1. *Suppose $\xi = \sigma(X^*)\zeta$, where ζ is independent of X^*. Also, suppose that the following conditions are satisfied:*

(i) $\sigma(\cdot)$ and $f_\zeta(\cdot)$ are bounded (non-negative) functions with m bounded derivatives, and $\sigma(\cdot)$ is bounded away from 0;

(ii) for sufficiently large values of t, $t^{m+k+1}|f_\zeta^{(k)}(t)|$ and $t^{m+k+1}|f_\zeta^{(k)}(-t)|$, for $k \in \{0, \dots, m\}$, are decreasing functions;

(iii) $\int |\zeta|^{m+k} \left|f_\zeta^{(k)}(\zeta)\right| d\zeta < C$ for $k \in \{0, \dots, m\}$.

Then, Assumption 2.5 is satisfied.

Proof of Lemma B.1. First, notice that $f_{\xi|X^*}(u|x) = \frac{1}{\sigma(x)} f_\zeta\left(\frac{u}{\sigma(x)}\right)$.

We want to show that

$$
\int |u|^m \sup_{\tilde{x} \in \mathcal{S}_X} \left|\nabla_x^\ell f_{\xi|X^*}(u|\tilde{x})\right| du < C
$$

for $\ell \in \{0, \dots, m\}$, where $C > 0$ is a universal constant. Since the derivatives of $f_{\xi|X^*}(u|x)$ are uniformly bounded (by condition (i)), we just focus on showing

$$
\int_{\underline{u}}^\infty |u|^m \sup_{\tilde{x} \in \mathcal{S}_X} \left|\nabla_x^\ell f_{\xi|X^*}(u|\tilde{x})\right| du < C
$$

for some (generic) $\underline{u} > 0$ ($\int_{-\infty}^{-\underline{u}} |u|^m \sup_{\tilde{x} \in \mathcal{S}_X} \left|\nabla_x^\ell f_{\xi|X^*}(u|\tilde{x})\right| du$ can be bounded using a similar argument).

Starting with $\ell = 0$, we verify

$$\int_{\underline{u}}^{\infty} |u|^m \sup_{\tilde{x} \in \mathcal{S}_X} f_{\xi|X^*}(u|\tilde{x}) du < C.$$

Notice that

$$\sup_{x \in \mathcal{S}_X} f_{\xi|X^*}(u|x) = \sup_{s \in \Sigma} \left\{ \frac{1}{s} f_\zeta \left(\frac{u}{s} \right) \right\},$$

where $\Sigma = \{\sigma(x) : x \in \mathcal{S}_X\}$ is bounded. Let $s^* = \sup_{x \in \mathcal{S}_X} \sigma(x)$. Condition (ii) implies that there exists $\underline{u} > 0$ such that for all $|u| > \underline{u}$ we have

$$\sup_{s \in \Sigma} \left\{ \frac{1}{s} f_\zeta \left(\frac{u}{s} \right) \right\} = \frac{1}{s^*} f_\zeta \left(\frac{u}{s^*} \right).$$

Then,

$$\int_{\underline{u}}^{\infty} |u|^m \sup_{\tilde{x} \in \mathcal{S}_X} f_{\xi|X^*}(u|\tilde{x}) du = \int_{\underline{u}}^{\infty} |u|^m \frac{1}{s^*} f_\zeta \left(\frac{u}{s^*} \right) du < C,$$

where the last inequality is due to condition (iii).

Next, we verify that

$$\int_{\underline{u}}^{\infty} |u|^m \sup_{\tilde{x} \in \mathcal{S}_X} \left| \nabla_x^\ell f_{\xi|X^*}(u|\tilde{x}) \right| du < C$$

for $\ell \in \{1, \ldots, m\}$. Since the derivatives of $\sigma(\cdot)$ are bounded, we have

$$\sup_{\tilde{x} \in \mathcal{S}_X} \left| \nabla_x^\ell f_{\xi|X^*}(u|\tilde{x}) \right| \leqslant C \sum_{k=1}^{\ell} \sup_{s \in \Sigma} \left| \nabla_s^k \left\{ \frac{1}{s} f_\zeta \left(\frac{u}{s} \right) \right\} \right|,$$

so it is sufficient to verify

$$\int_{\underline{u}}^{\infty} |u|^m \sup_{s \in \Sigma} \left| \nabla_s^\ell \left\{ \frac{1}{s} f_\zeta \left(\frac{u}{s} \right) \right\} \right| du < C,$$

for $\ell \in \{1, \ldots, m\}$. Next,

$$\nabla_s^\ell \left\{ \frac{1}{s} f_\zeta \left(\frac{u}{s} \right) \right\} = \sum_{k=0}^{\ell} a_{\ell k} \frac{u^k}{s^{\ell+k+1}} f^{(k)} \left(\frac{u}{s} \right)$$

for some constants $a_{\ell k}$. Then, condition (ii) implies that there exists $\underline{u} > 0$ such that for all $|u| > \underline{u}$ we have

$$\sup_{s \in \Sigma} \left| \nabla_s^\ell \left\{ \frac{1}{s} f_\zeta \left(\frac{u}{s} \right) \right\} \right| = \left| \nabla_s^\ell \left\{ \frac{1}{s} f_\zeta \left(\frac{u}{s} \right) \right\} \right|_{s=s^*} = \left| \sum_{k=0}^{\ell} a_{\ell k} \frac{u^k}{(s^*)^{\ell+k+1}} f^{(k)} \left(\frac{u}{s^*} \right) \right|.$$

Then,

$$\begin{aligned} \int_{\underline{u}}^{\infty} |u|^m \sup_{\tilde{x} \in \mathcal{S}_X} \left| \nabla_x^\ell f_{\xi|X^*}(u|\tilde{x}) \right| du &< C \int_{\underline{u}}^{\infty} \sum_{k=0}^{\ell} \frac{|u|^{m+k}}{(s^*)^{\ell+k+1}} \left| f_\zeta^{(k)} \left(\frac{u}{s^*} \right) \right| du \\ &< C \sum_{k=0}^{\ell} \int |\zeta|^{m+k} \left| f_\zeta^{(k)} (\zeta) \right| d\zeta \\ &< C, \end{aligned}$$

where the last inequality is due to condition (iii). □