Identification and Estimation of Nonseparable Triangular Equations with Mismeasured Instruments

Shaomin Wu

April 23, 2024

Abstract

In this paper, I study the nonparametric identification and estimation of the marginal effect of an endogenous variable X on the outcome variable Y, given a potentially mismeasured instrument variable W^*, without assuming linearity or separability of the functions governing the relationship between observables and unobservables. To address the challenges arising from the co-existence of measurement error and nonseparability, I first employ the deconvolution technique from the measurement error literature to identify the joint distribution of Y, X, W^* using two error-laden measurements of W^*. I then recover the structural derivative of the function of interest and the "Local Average Response" (LAR) from the joint distribution via the "unobserved instrument" approach in Matzkin (2016). I also propose nonparametric estimators for these parameters and derive their uniform rates of convergence. Monte Carlo exercises show evidence that the estimators I propose have good finite sample performance.

1 Introduction

This paper studies the nonparametric identification and estimation of the marginal effect of an endogenous variable X on the outcome variable Y, given a potentially mismeasured instrument variable W^*, without assuming linearity or separability of the functions governing the relationship between observables and unobservables. Without measurement error, this type of model is referred to as "nonseparable triangular equations" and its identification was studied by Chesher (2003), Imbens and Newey (2009), and Shaikh and Vytlacil (2011). Measurement error on the instrument variable poses additional challenges to the identification and estimation of the model primitives. Because of nonseparability, simply using an error-laden measurement as the instrument leads to inconsistent results. Because of measurement error, the true value of W^* is unobserved, making it impossible to proceed using existing methods in Imbens and Newey (2009). In this paper, I propose a way to deal with the two difficulties mentioned above and show the identification of average and individual-level marginal effects. I also propose estimators for these parameters. To illustrate ideas, I study the following model:

$$
\begin{aligned}
Y &= m(X, \epsilon) \\
X &= h(W^*, \eta),
\end{aligned}
$$

where X is endogenous in the sense that it's correlated with ϵ, and W^* is the instrument variable independent with both of the error terms ϵ, η but cannot be measured accurately. Following the measurement error literature, I assume that there are two error-laden measurements of W^*, denoted as W_1, W_2. Denote the corresponding measurement errors as $\Delta W_1, \Delta W_2$, i.e. $W^* = W_1 + \Delta W_1$, $W^* = W_2 + \Delta W_2$. I could allow a vector of observed exogenous control variables Z to enter both equations and all the arguments will carry over by conditioning on Z, so they are omitted here for simplicity.

To give an example of where this model can be used, consider the Engel curve estimation studied in Blundell, Chen and Kristensen (2007). They used Sieve Minimum Distance to estimate a semi-nonparametric model, but the curve can be estimated under less restrictive modeling assumptions as done in Imbens and Newey (2009). Y here is the share of expenditure on a commodity of a household. X is the log of the household's total expenditure. Z is a vector capturing the household's demographic composition and ϵ captures the household's unobserved heterogeneity. Researchers might be interested in the average response of households' expenditure on food Y, to changes in household's total expenditure X, holding the distribution of unobserved household heterogeneity $F_{\epsilon|X=x}$ fixed. This parameter is called

"Local Average Response" (LAR) by Altonji and Matzkin (2005) and if the function m is differentiable, it can be written as

$$\mathbf{LAR}(x) = \int \frac{\partial m(x,e)}{\partial x} f_{\epsilon|X=x}(e)de. \tag{1}$$

In addition, researchers might also be interested in the structural derivative of the function m. This is a more disaggregate level parameter. It stands for the marginal response of Y to changes in X for a specific household, say the household with log total expenditure equal $\bar{x}$, and share of food expenditure equal $\bar{y}$. Denote the structural derivative of this household as $\rho(\bar{y},\bar{x})$. If m is differentiable and strictly monotone in its second argument for all values of X, $\rho(\bar{y},\bar{x})$ can be written as

$$\rho(\bar{y},\bar{x}) = \left.\frac{\partial m(x,\epsilon)}{\partial x}\right|_{x=\bar{x},\epsilon=m^{-1}(\bar{y},\bar{x})}, \tag{2}$$

where m^{-1} is the inverse of m with respect to its second argument, and $m^{-1}(\bar{y},\bar{x})$ is the value of ϵ of the specific household one is interested in. To estimate these parameters, under the assumption that heterogeneity in earnings is not correlated with households' preferences over consumption, one can use the income of the head of the household as the IV W^*. The income variable is likely to suffer from measurement errors. Researchers could obtain multiple measurements of it from panel data, for example.

To see how nonseparability makes it harder to identify parameters like the LAR and the structural derivative when the IV W^* is mismeasured, consider the case when both equations are linear:

$$\begin{aligned} Y &= \alpha_0 + \alpha_1 X + \epsilon \\ X &= \beta_0 + \beta_1 W^* + \eta, \end{aligned}$$

where $E[\epsilon W^*] = 0$ (exclusion restriction) and $E[XW^*] \neq E[X]E[W^*]$ (relevance condition). Suppose I have $W_2 = W^* + \Delta W_2$ as an error-laden measurement of W^*, it's not hard to verify that $E[\epsilon W_2] = 0$ and $X \not\perp W_2$ still hold under mild assumptions on ΔW_2 (e.g. $E[\Delta W_2|Y,X] = 0$). This means in a linear model, even if the instrument variable suffers from measurement errors, one can still use the error-laden measurement W_2 as an IV and proceed as usual. However, this is not the case in nonseparable models. Plugging W_2 into the second equation yields $X = h(W_2 - \Delta W_2, \eta)$, where both ΔW_2 and η are unobervable and importantly, $W_2 \not\perp \Delta W_2$. This means if I were to use W_2 as an IV, both of the two equations in the triangular system would contain endogenous variables and nothing can be done without additional IVs outside of this system. This also means that if one simply uses W_2 as the instrument variable and proceeds with the standard techniques in Imbens and Newey (2009), they would get inconsistent results.

Given that nonseparability makes the problem much harder to solve, a natural question is why one wants to deal with nonseparable models instead of an additive separable or linear model. There are multiple reasons why researchers might prefer a nonseparable model. First, nonseparable models allow the observed variable X and the unobserved variable ϵ to interact in a flexible way. For example, a recent paper by Brancaccio, Kalouptsidi and Papageorgiou (2020) estimated the matching function $m(s, e)$ between ships (of number s) and exporters (of number e which is unobserved) at a seaport. Not imposing functional form assumptions (including separability) is important. It allows the authors to remain agnostic about the nature of the meeting process. In addition, the flexibility of functional forms can be key when deriving welfare and policy implications (see Brancaccio, Kalouptsidi and Papageorgiou (2020)). Second, nonseparability is also important when X and ϵ are correlated, since in many cases the source of endogeneity is the nonseparable nature of the model. For example, X could come from the optimization problem of maximizing the expected value of Y minus the production cost, given an exogenous variable W, and some noisy information about ϵ. Think of X as an individual's education level, or a firm's input level, and Y as the individual's lifetime earnings, or the firm's output. Then X could be the solution of $\max_x \{E[m(x, \epsilon)|\eta, W] - c(x, W)\}$, leading to $X = h(W, \eta)$, where η is some noisy proxy of ϵ, and $c(x, W)$ is the cost function. If the function m were additively separable in ϵ, the optimal choice of x would not even depend on η.

When there is no measurement error, Imbens and Newey (2009) proposed a way to identify the model primitives in a nonseparable triangular system, making use of the fact that $X \perp \epsilon | \eta$. They first estimate the control variable η (or a strictly monotone function of η, denoted as V in their paper) from the second equation, and then estimate the model primitives in the first equation by first conditioning on the estimated η, and then integrate it out. Their method cannot directly apply when the instrument variable W^* in the second equation is mismeasured, because the control variable (or control function) cannot be observed from error-laden measurements of W^*. A recent paper by Aradillas-Lopez (2022) studies inference in models where control functions are unobserved. The setup of his paper is different from this paper in many aspects. He requires the availability of observable or estimable bounds for the unobserved control functions, which is not required by the model in this paper. Instead, this paper requires the availability of error-laden measurements and builds upon the measurement error literature. Also, his focus is on constructing confidence sets for finite-dimensional parameters, while my focus is on point identification and estimation of infinite-dimensional parameters.

In this paper, I propose a method that makes use of the same intuition as in Imbens and Newey (2009), but can deal with mismeasured W^*. Same as Imbens and Newey (2009), I utilize the fact that the correlation between X and ϵ is merely coming from η, but instead of estimating and conditioning on η, I separate out this correlation by writing ϵ as a function of η and a uniformly distributed random variable which is independent with X and W^*. In this way, I am able to write the model primitives like the structural derivative and LAR as functionals of the joint distribution of Y, X, W^*, which can be recovered using the two error-laden measurements and the deconvolution technique developed in the measurement error literature (e.g. Fan (1991a), Fan and Truong (1993), Schennach (2004a), Schennach (2004b)). I can thus identify the model primitives in a constructive way and estimate them using plug-in estimators. I also derive uniform rates of convergence of the estimators.

This paper is most related to Schennach, White and Chalak (2012) (SWC, hereafter), where the authors also consider a triangular simultaneous equations model with a mismeasured exogenous instrument. This paper differs from their paper in the modeling assumptions, parameters that can be identified, and also theoretical methods. SWC shows that under separability assumption on the second equation, the instrument conditioned marginal response[1] $E[m_x(X,\epsilon)|W^*=w^*]$ can be written as a ratio of the derivative of the conditional mean of Y given W^* over the derivative of the conditional mean of X given W^*, both of which can be recovered from the data given two error-laden measurements W_1, W_2. Integrating out W^*, they can also recover the average response $E[m_x(X,\epsilon)]$. However, under the nonseparability of both of the equations, their method cannot recover either of the two parameters. Different from SWC, this paper shows that it's possible to identify not only the instrument-conditioned marginal response and the average marginal response but also the structural derivative and LAR, even when both of the equations are nonseparable. This conclusion, however, comes at the expense of more assumptions on the function m and unobservables compared with SWC. In particular, this paper assumes strict monotonicity of the function m on its second argument, and scalar unobservables, which are not required in SWC. Regarding estimation, both SWC and this paper employ plug-in estimators, but the asymptotic analysis in this paper is a bit more complex than in SWC, because of the observed variable components Y, X of the joint density f_{Y,X,W^*}. They have non-trivial implications on the asymptotic treatment (including the convergence rates) so that SWC's asymptotic analysis cannot directly apply here.

This paper is also closely related to Song, Schennach and White (2015). In their paper, the coauthors study a nonseparable model with mismeasured endogenous variable, assuming a correctly measured control variable is available. The setup of this paper is different from theirs. This paper also studies nonseparable models with endogeneity, but instead of mismeasured endogenous variable, this paper studies the case when the endogenous variable is correctly measured, and instead of assuming the existence of a correctly measured control variable, this paper assumes that a potentially mismeasured instrument variable is available. The two papers are complementary depending on the availability of data. Regarding the parameters of interest, in addition to the various parameters studied in Song, Schennach and White (2015), including the average marginal response conditional on the control variable, the average marginal response, the LAR, and the weighted average version of these variables, this paper additionally show identification of the structural derivative, which is more disaggregate and could be helpful when

[1] m_x denotes the partial derivative of the m function with respect to its first argument.

researchers are interested heterogeneous effects.

The rest of the paper consists of six parts. Section 2 introduces the model and assumptions. Section 3 talks about the identification of the model primitives. Section 4 proposes plug-in estimators of the model primitives identified in Section 3, and talks about their asymptotic properties. Section 5 conducts limited Monte Carlo studies to show the finite sample performance of the estimator. Section 6 concludes.

2 The Model

I consider the following triangular model in this paper:

$$Y = m(X, Z, \epsilon)$$
$$X = h(W^*, Z, \eta)$$

where X is an observed endogenous variable that is correlated with ϵ. In the returns to education example, X stands for years of education, which is correlated with ϵ since it's chosen by the agent as an equilibrium outcome. Z is a vector of observed exogenous variables which are independent with ϵ and η. W^* is an instrument for X and satisfies $W^* \perp (\epsilon, \eta)$. Researchers cannot measure W^* exactly but have two error-laden measurements of it:

$$W_1 = W^* + \Delta W_1$$
$$W_2 = W^* + \Delta W_2.$$

The measurement errors satisfy $E[\Delta W_1 \mid W^*, \Delta W_2] = 0$ and $\Delta W_2 \perp W^*, Y, X, Z$. As in the literature studying the identification of nonseparable models (Chesher (2003), Matzkin (2003),Altonji and Matzkin (2005),Matzkin (2015)), I impose monotonicity assumptions on the structural functions. I assume that m is strictly increasing in ϵ and that h is strictly increasing in η. Since the vector of covariates Z is correctly measured and exogenous, all the analysis can be done conditional on Z. For brevity, from now on I omit the vector Z in my notations and work on the simplified model below, while other assumptions on the structural functions and distributions of variables remain unchanged.

$$Y = m(X, \epsilon) \tag{3}$$

$$X = h(W^*, \eta) \tag{4}$$

The assumptions mentioned above are stated formally below:

Assumption 1. *Function m and h are continuously differentiable with respect to both of their arguments and are strictly increasing in their respective second argument.*

Assumption 2. $W^* \perp (\epsilon, \eta)$.

Assumption 3. $W_1 = W^* + \Delta W_1$ *and* $W_2 = W^* + \Delta W_2$.

Assumption 4. $E[\Delta W_1 \mid W^*, \Delta W_2] = 0$*,* $\Delta W_2 \perp W^*, Y, X$.

Following the measurement error literature, I also impose:

Assumption 5. *For any finite* $t \in \mathbb{R}$*,* $|E[\exp(\mathbf{i}tW_2)]| > 0$.

Note that for the measurement error ΔW_1, only the mean independence assumption is imposed, which is weaker than the assumption on ΔW_2. This weaker assumption is sufficient to identify the distribution of W^* (Schennach (2004b)). On the other hand, although ΔW_2 satisfies strong independence assumptions, making W_2 a measurement with classical measurement error, one still cannot use W_2 as the instrument because of the nonseparability of the model. More specifically, plugging W_2 into the second equation in the triangular system yields

$$X = h(W_2 - \Delta W_2, \eta).$$

This is a nonseparable model with two unobservables ΔW_2 and η, and W_2 is not independent with ΔW_2. This means both of the two equations in the system contain endogenous variables and thus cannot be identified without further assumptions.

In addition, I also impose the following assumptions on the distribution of Y, X, W^*:

Assumption 6. *The distribution of* (Y, X, W^*) *has compact support, denoted as* $\mathbb{S}_{(Y,X,W^*)}$ *and has a joint density* f_{Y,X,W^*} *which is twice continuously differentiable on* R^3.

Assumption 7. *For any x, w^* belonging to the support of X, W^*, $\left|\frac{\partial F_{X|W^*=w^*}(x)}{\partial w^*}\right| > 0$.*

Assumption 7 is a sufficient condition to ensure $\left.\frac{\partial h(w^*,\eta)}{\partial w^*}\right|_{\eta=r(x,w^*)} \neq 0$ for all x, w^* on the support of X, W^*. This assumption is like the rank condition imposed in the linear instrument variable models. To illustrate this idea, suppose h is a linear function such that $X = \pi W^* + \eta$, then $F_{X|W^*=w^*}(x) = F_\eta(x - \pi w^*)$, and $\pi \neq 0$ is a neccessary condition for Assumption 7.

3 Identification

In this section, I show the identification of the model parameters in two steps. First, assuming that the joint distribution of Y, X, W^* is known, I show identification of the derivative of the structural function m with respect to x, when the value of ϵ is fixed, and the identification of other model parameters including the LAR and AR. Then I show how to identify the joint distribution of Y, X, W^* from the two measurements W_1 and W_2.

3.1 Identification of parameters assuming the joint distribution of Y, X, W^* is known

The independence between W^* and ϵ, η implies that conditional on η, W^* and ϵ are independent. The next lemma uses this fact to show that one can write ϵ as a function of η and another random variable which is independent with X, W^*. This lemma is similar to Proposition 5.1 in Matzkin (2016). Before stating the lemma, I first impose the following assumption on the conditional distribution of ϵ given η:

Assumption 8. *$F_{\epsilon|\eta=\bar{\eta}}(\bar{\epsilon})$ is strictly increasing in $\bar{\epsilon}$, given any values of $\bar{\eta}$.*

Lemma 1. *Under the model setup, suppose Assumption 8 is satisfied. There exists a function $s : \mathbb{R}^2 \to \mathbb{R}$ strictly increasing in its second argument and an unobservable random term δ such that*

$$\epsilon = s(\eta, \delta), \tag{5}$$

and δ is independent of (X, W^*) and is $U(0,1)$.

Proof. See Appendix B. ■

Next, I plug (5) into the first structural equation. The assumption that h is strictly increasing in η implies that one can write the inverse of $h(W^*, \eta)$ w.r.t. η as $r(X, W^*)$. Then I have

$$Y = m(X, \epsilon) = m(X, s(\eta, \delta)) = m(X, s(r(X, W^*), \delta)) \equiv v(X, W^*, \delta). \tag{6}$$

Equation (6) builds a bridge between the structural function m and the reduced form function v. Note that v is a function of observable variables X, W^* and an unobservable variable δ which is independent of the observables. The derivatives of v can be identified by applying identification techniques in standard nonseparable models (Matzkin (2003)). To identify the main parameter of interest, the derivative of the structural function m, one can utilize the last equality in 6: $m(X, s(r(X, W^*), \delta)) \equiv v(X, W^*, \delta)$. Taking derivative w.r.t X and W^* yields:

$$\left.\frac{\partial m(x,\epsilon)}{\partial x}\right|_{\epsilon=s(r(x,w^*),\delta)} + \left.\frac{\partial m(x,\epsilon)}{\partial \epsilon}\right|_{\epsilon=s(r(x,w^*),\delta)} \cdot \left.\frac{\partial s(\eta,\delta)}{\partial \eta}\right|_{\eta=r(x,w^*)} \cdot \frac{\partial r(x,w^*)}{\partial x} = \frac{\partial v(x,w^*,\delta)}{\partial x} \tag{7}$$

$$\left.\frac{\partial m(x,\epsilon)}{\partial \epsilon}\right|_{\epsilon=s(r(x,w^*),\delta)} \cdot \left.\frac{\partial s(\eta,\delta)}{\partial \eta}\right|_{\eta=r(x,w^*)} \cdot \frac{\partial r(x,w^*)}{\partial w^*} = \frac{\partial v(x,w^*,\delta)}{\partial w^*} \tag{8}$$

Plug (8) into (7) to cancel $\left.\frac{\partial m(x,\epsilon)}{\partial \epsilon}\right|_{\epsilon=s(r(x,w^*),\delta)} \cdot \left.\frac{\partial s(\eta,\delta)}{\partial \eta}\right|_{\eta=r(x,w^*)}$ yields

$$\left.\frac{\partial m(x,\epsilon)}{\partial x}\right|_{\epsilon=s(r(x,w^*),\delta)} = \frac{\partial v(x,w^*,\delta)}{\partial x} - \frac{\partial v(x,w^*,\delta)}{\partial w^*}\frac{\frac{\partial r(x,w^*)}{\partial x}}{\frac{\partial r(x,w^*)}{\partial w^*}}.$$

Taking derivative w.r.t. W^* on both sides of $X \equiv h(W^*, r(X, W^*))$ and cancelling out the unobserved $\left.\frac{\partial h(w^*,\eta)}{\partial \eta}\right|_{\eta=r(x,w^*)}$, one can get $\left.\frac{\partial h(w^*,\eta)}{\partial w^*}\right|_{\eta=r(x,w^*)} = -\frac{\frac{\partial r(x,w^*)}{\partial w^*}}{\frac{\partial r(x,w^*)}{\partial x}}$. Then one can write

$$\left.\frac{\partial m(x,\epsilon)}{\partial x}\right|_{\epsilon=s(r(x,w^*),\delta)} = \frac{\partial v(x,w^*,\delta)}{\partial x} + \frac{\partial v(x,w^*,\delta)}{\partial w^*}\frac{1}{\left.\frac{\partial h(w^*,\eta)}{\partial w^*}\right|_{\eta=r(x,w^*)}}. \tag{9}$$

The right-hand side of the equation (9) can be identified from the data. To show this, note that

$$F_{Y|X=x,W^*=w^*}(v(x,w^*,\delta)) = F_\delta(\delta) = \delta \tag{10}$$

$$
\begin{aligned}
&\Rightarrow \frac{\partial F_{Y|X=x,W^*=w^*}(y)}{\partial x} + \frac{\partial F_{Y|X=x,W^*=w^*}(y)}{\partial y}\frac{\partial v(x,w^*,\delta)}{\partial x} = 0 \\
&\Rightarrow \frac{\partial v(x,w^*,\delta)}{\partial x} = -\frac{\partial F_{Y|X=x,W^*=w^*}(y)}{\partial x} \Big/ \frac{\partial F_{Y|X=x,W^*=w^*}(y)}{\partial y},
\end{aligned}
$$

Similarly

$$
\begin{aligned}
&\frac{\partial v(x,w^*,\delta)}{\partial w^*} = -\frac{\partial F_{Y|X=x,W^*=w^*}(y)}{\partial w^*} \Big/ \frac{\partial F_{Y|X=x,W^*=w^*}(y)}{\partial y} \\
&\frac{\partial h(w^*,\eta)}{\partial w^*} = -\frac{\partial F_{X|W^*=w^*}(x)}{\partial w^*} \Big/ \frac{\partial F_{X|W^*=w^*}(x)}{\partial x},
\end{aligned}
$$

where δ is the unique value such that $y = v(x, w^*, \delta)$ and η is the unique value such that $x = h(w^*, \eta)$. Plug into (9), one can get

$$
\begin{aligned}
\rho(\bar{y},\bar{x}) &\equiv \left.\frac{\partial m(x,\epsilon)}{\partial x}\right|_{x=\bar{x},\epsilon=m^{-1}(\bar{x},\bar{y})} \\
&= -\frac{\frac{\partial F_{Y|X=\bar{x},W^*=w^*}(\bar{y})}{\partial x}}{\frac{\partial F_{Y|X=\bar{x},W^*=w^*}(\bar{y})}{\partial y}} + \frac{\frac{\partial F_{Y|X=\bar{x},W^*=w^*}(\bar{y})}{\partial w^*}}{\frac{\partial F_{Y|X=\bar{x},W^*=w^*}(\bar{y})}{\partial y}} \times \frac{\frac{\partial F_{X|W^*=w^*}(\bar{x})}{\partial x}}{\frac{\partial F_{X|W^*=w^*}(\bar{x})}{\partial w^*}}
\end{aligned} \tag{11}
$$

for all $\bar{y}, \bar{x}$ belongs to support, and values of $\bar{\epsilon}$ such that $\bar{y} = m(\bar{x}, \bar{\epsilon})$. Note that as long as $\bar{y}$ and $\bar{x}$ are fixed, $\bar{\epsilon}$ is fixed, no matter which value is picked for w^*. This means there's actually overidentification for the structural derivative. To identify the **LAR** and **AR**, one needs independent variations of ϵ given X. Rewrite (10) with slightly different notations yields:

$$
\begin{aligned}
&F_{Y|X=x,W^*=w^*}(v(x,w^*,\delta)) = F_\delta(\delta) = \delta \\
\Rightarrow &v(x,w^*,\delta) = F^{-1}_{Y|X=x,W^*=w^*}(\delta)
\end{aligned}
$$

so that one can write

$$
\left.\frac{\partial m(x,\epsilon)}{\partial x}\right|_{\epsilon=s(r(x,w^*),\delta)} = \frac{\partial F^{-1}_{Y|X=x,W^*=w^*}(\delta)}{\partial x} - \frac{\partial F^{-1}_{Y|X=x,W^*=w^*}(\delta)}{\partial w^*}\frac{\frac{\partial F_{X|W^*=w^*}(x)}{\partial x}}{\frac{\partial F_{X|W^*=w^*}(x)}{\partial w^*}}. \tag{12}
$$

To show the identification of the **LAR** and **AR**, one needs to show that the support of $s(r(X,W^*),\delta)$ given $X = x$ is the same as the support of ϵ given $X = x$. This is stated formally in the lemma below.

Lemma 2. *Define random variable* $\mathbf{s} \equiv s(r(X, W^*), \delta)$. *Denote the support of* $\mathbf{s}$ *and* ϵ *conditional on* $X = x$ *as* $\mathbb{S}_{\mathbf{s}|X=x}$ *and* $\mathbb{S}_{\epsilon|X=x}$, *respectively. If Assumption 8 holds, then* $\mathbb{S}_{\mathbf{s}|X=x} = \mathbb{S}_{\epsilon|X=x}$, *for all* x *belonging to its support.*

Proof. See Appendix B. ■

Lemma 1 and 2 ensures that $F_{\mathbf{s}|X}$ is the same as $F_{\epsilon|X}$, so that integrating $\frac{\partial m(x,\epsilon)}{\partial x}$ over ϵ given $X = x$ will be equivalent to integrating $\frac{\partial m(x,\mathbf{s})}{\partial x}$ over $\mathbf{s}$ given $X = x$. Then I have the following identification results:

$$\mathbf{LAR}: \; E\left[\frac{\partial m(X,\epsilon)}{\partial x} \mid X = x\right] = \int\int_0^1 \left.\frac{\partial m(x,\epsilon)}{\partial x}\right|_{\epsilon=s(r(x,w^*),\delta)} f_{W^*|X=x}(w^*)d\delta dw^* \quad (13)$$

$$\mathbf{AR}: \; E\left[\frac{\partial m(X,\epsilon)}{\partial x}\right] = \int\int\int_0^1 \left.\frac{\partial m(x,\epsilon)}{\partial x}\right|_{\epsilon=s(r(x,w^*),\delta)} f_{X,W^*}(x,w^*)d\delta dw^* dx. \quad (14)$$

3.2 Identification of the joint density of Y, X, W^*

First note that by Theorem 1 in Schennach (2004a)

$$\phi_{W^*}(t) = \exp\left(\int_0^t \frac{E\left[iW_1 e^{i\xi W_2}\right]}{E\left[e^{i\xi W_2}\right]} d\xi\right).$$

Then I can identify the density $f_{W^*}(w^*)$ by taking the Inverse Fourier Transform:

$$f_{W^*}(w^*) = \frac{1}{2\pi}\int e^{-itw^*}\phi_{W^*}(t)dt.$$

For the joint density $f_{Y,X,W^*}(y,x,w^*)$, by Assumption 4, I have the following convolution:

$$f_{Y,X,W_2}(x,w_2) = \int f_{Y,X,W^*}(x,\nu) f_{\Delta W_2}(w_2 - \nu)d\nu$$

Applying Fourier transformation on both sides (holding x constant) yields

$$\phi_{f_{Y,X,W_2}(x,\cdot)}(t) = \phi_{f_{Y,X,W^*}(y,x,\cdot)}(t)\phi_{\Delta W_2}(t),$$

which, by Assumption 5 implies

$$\phi_{f_{Y,X,W^*}(y,x,\cdot)}(t) = \frac{\phi_{f_{Y,X,W_2}(y,x,\cdot)}(t)}{\phi_{\Delta W_2}(t)} = \frac{\phi_{f_{Y,X,W_2}(y,x,\cdot)}(t)\phi_{W^*}(t)}{\phi_{W_2}(t)}. \quad (15)$$

Then applying the inverse Fourier transform, one can get

$$f_{Y,X,W^*}(x, w^*) = \frac{1}{2\pi}\int e^{-itw^*}\frac{\phi_{f_{X,W_2}(y,x,\cdot)}(t)\phi_{W^*}(t)}{\phi_{W_2}(t)}dt,, \tag{16}$$

and similarly,

$$f_{Y,X,W^*}(y, x, w^*) = \frac{1}{2\pi}\int e^{-itw^*}\frac{\phi_{f_{Y,X,W_2}(y,x,\cdot)}(t)\phi_{W^*}(t)}{\phi_{W_2}(t)}dt. \tag{17}$$

Then $F_{X|W^*=w^*}(x)$ follows from

$$F_{X|W^*=w^*}(x) = \frac{\int_{-\infty}^{x} f_{X,W^*}(y, x, w^*)dx}{\int_{-\infty}^{\infty} f_{X,W^*}(y, x, w^*)dx} \tag{18}$$

and $F_{Y|X=x,W^*=w^*}(y)$ follows from

$$F_{Y|X=x,W^*=w^*}(y) = \frac{\int_{-\infty}^{y} f_{Y,X,W^*}(y, x, w^*)dy}{\int_{-\infty}^{\infty} f_{Y,X,W^*}(y, x, w^*)dy}.$$

From the equations above, one can write the derivatives of the conditional CDFs as functionals of the joint densities:

$$\begin{aligned}
&\frac{\partial F_{Y|X=x,W^*=w^*}(y)}{\partial w^*} \\
&= \frac{\left(\int_{-\infty}^{y}\frac{\partial f_{Y,X,W^*}(s,x,w^*)}{\partial w^*}ds\right) f_{X,W^*}(x, w^*) - \left(\int_{-\infty}^{y} f_{Y,X,W^*}(s, x, w^*)ds\right)\frac{\partial f_{X,W^*}(x,w^*)}{\partial w^*}}{f^2_{X,W^*}(x, w^*)} \\
&\frac{\partial F_{Y|X=x,W^*=w^*}(y)}{\partial x} \\
&= \frac{\left(\int_{-\infty}^{y}\frac{\partial f_{Y,X,W^*}(s,x,w^*)}{\partial x}ds\right) f_{X,W^*}(x, w^*) - \left(\int_{-\infty}^{y} f_{Y,X,W^*}(s, x, w^*)ds\right)\frac{\partial f_{X,W^*}(x,w^*)}{\partial x}}{f^2_{X,W^*}(x, w^*)} \\
&\frac{\partial F_{Y|X=x,W^*=w^*}(y)}{\partial y} = \frac{f_{Y,X,W^*}(y, x, w^*)}{f_{X,W^*}(x, w^*)} \\
&\frac{\partial F_{X|W^*=w^*}(x)}{\partial w^*} = \frac{\left(\int_{-\infty}^{x}\frac{\partial f_{X,W^*}(s,w^*)}{\partial w^*}ds\right) f_{W^*}(w^*) - \left(\int_{-\infty}^{x} f_{X,W^*}(s, w^*)ds\right)\frac{\partial f_{W^*}(w^*)}{\partial w^*}}{f^2_{W^*}(w^*)} \\
&\frac{\partial F_{X|W^*=w^*}(x)}{\partial x} = \frac{f_{X,W^*}(x, w^*)}{f_{W^*}(w^*)}.
\end{aligned}$$

This means that the right-hand side of the equation (11) can be written as a functional of the joint density f_{Y,X,W^*}, which can be identified from the data. To show identification of the left-hand side of the equation (12), one needs to build the relationship between the derivatives of conditional quantiles and the joint densities, which is shown in the following lemma:

Lemma 3. *Let $\delta \in [0,1]$ be a constant, then*

$$\frac{\partial F^{-1}_{Y|X=x,W^*=w^*}(\delta)}{\partial w^*} = \frac{\delta \frac{\partial f_{X,W^*}(x,w^*)}{\partial w^*} - \int_{-\infty}^{F^{-1}_{Y|X=x,W^*=w^*}(\delta)} \frac{\partial f_{Y,X,W^*}(y,x,w^*)}{\partial w^*} dy}{f_{Y,X,W^*}\left(F^{-1}_{Y|X=x,W^*=w^*}(\delta), x, w^*\right)}$$

$$\frac{\partial F^{-1}_{Y|X=x,W^*=w^*}(\delta)}{\partial x} = \frac{\delta \frac{\partial f_{X,W^*}(x,w^*)}{\partial x} - \int_{-\infty}^{F^{-1}_{Y|X=x,W^*=w^*}(\delta)} \frac{\partial f_{Y,X,W^*}(y,x,w^*)}{\partial x} dy}{f_{Y,X,W^*}\left(F^{-1}_{Y|X=x,W^*=w^*}(\delta), x, w^*\right)}.$$

Proof. see Appendix B. ■

The identification of the left-hand side of (13) and (14) is a straightforward implication of Lemma 3 and equation (12).

4 Estimation

4.1 The Estimator

From now on, I use the following function to denote the joint density of Y, X, W^*, or its derivative with respect to Y, X, W^*:

$$g_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*) = \frac{\partial^{\lambda} f_{Y,X,W^*}(y,x,w^*)}{\partial w^{*\lambda_1} \partial y^{\lambda_{2,1}} \partial x^{\lambda_{2,2}}},$$

where $\lambda_1, \lambda_{2,1}, \lambda_{2,2} \in \{0,1\}$ and $\lambda \equiv \max\{\lambda_1, \lambda_{2,1}, \lambda_{2,2}\} \leq 1$. The $0-th$ order derivative of a function is defined as the function itself. For convenience of notation, I also define $\lambda_2 \equiv \max\{\lambda_{2,1}, \lambda_{2,2}\}$. In this section, I first propose an estimator for $g_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*)$, and then propose plug-in estimators for the structural derivative and a weighted average version of the **LAR** since they can be written as known functionals of $g_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*)$.

To make the expressions more transparent and clear, I show the explicit forms of the functionals by taking the structural derivative as an example. I write each component on the right-hand-side of equation 11 as

$$\frac{\partial F_{Y|X=x,W^*=w^*}(y)}{\partial w^*}$$

$$= \frac{\left(\int_{-\infty}^{y} g_{1,0,0}(s,x,w^*)ds\right)}{\left(\int_{-\infty}^{\infty} g_{0,0,0}(y,x,w^*)dy\right)} - \frac{\left(\int_{-\infty}^{y} g_{0,0,0}(s,x,w^*)ds\right)\left(\int_{-\infty}^{\infty} g_{1,0,0}(y,x,w^*)dy\right)}{\left(\int_{-\infty}^{\infty} g_{0,0,0}(y,x,w^*)dy\right)^2}$$

and

$$\frac{\partial F_{Y|X=x,W^*=w^*}(y)}{\partial x}$$
$$= \frac{\left(\int_{-\infty}^{y} g_{0,0,1}(s,x,w^*)ds\right)}{\left(\int_{-\infty}^{\infty} g_{0,0,0}(y,x,w^*)dy\right)} - \frac{\left(\int_{-\infty}^{y} g_{0,0,0}(s,x,w^*)ds\right)\left(\int_{-\infty}^{\infty} g_{0,0,1}(y,x,w^*)dy\right)}{\left(\int_{-\infty}^{\infty} g_{0,0,0}(y,x,w^*)dy\right)^2}$$

and

$$\frac{\partial F_{Y|X=x,W^*=w^*}(y)}{\partial y} = \frac{g_{0,0,0}(y,x,w^*)}{\left(\int_{-\infty}^{\infty} g_{0,0,0}(y,x,w^*)dy\right)}$$

and

$$\frac{\partial F_{X|W^*=w^*}(x)}{\partial w^*}$$
$$= \frac{\left(\int_{-\infty}^{x}\int_{-\infty}^{\infty} g_{1,0,0}(y,s,w^*)dyds\right)}{\left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g_{0,0,0}(y,x,w^*)dydx\right)}$$
$$- \frac{\left(\int_{-\infty}^{x}\int_{-\infty}^{\infty} g_{0,0,0}(y,s,w^*)dyds\right)\left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g_{1,0,0}(y,x,w^*)dydx\right)}{\left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g_{0,0,0}(y,x,w^*)dydx\right)^2}$$

and

$$\frac{\partial F_{X|W^*=w^*}(x)}{\partial x} = \frac{\left(\int_{-\infty}^{\infty} g_{0,0,0}(y,x,w^*)dy\right)}{\left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g_{0,0,0}(y,x,w^*)dydx\right)}.$$

For the **LAR**, I will introduce a weighted average version of it, and propose an estimator for this weighted average version of the **LAR**. The aim of introducing the weight is to ensure that the integration is taken on a set where the joint density f_{Y,X,W^*} is bounded away from zero. This has the benefit of guaranteeing that the relevant functional is Fréchet differentiable when f_{Y,X,W^*} appears in the denominator of it. The weighted average version of the **LAR** is defined as follows:

$$\mathbf{WLAR}(x) \equiv E\left[\omega(X,W^*,\epsilon)\frac{\partial m(X,\epsilon)}{\partial x} \mid X = x\right]$$

$$= \int \int_0^1 \omega(x, w^*, \epsilon) \left. \frac{\partial m(x, \epsilon)}{\partial x} \right|_{\epsilon = s(r(x,w^*),\delta)} f_{W^*|X=x}(w^*) d\delta dw^*, \tag{19}$$

where $\omega(x, w^*, \epsilon)$ is a known or estimable weighting function taking value 0 outside a compact set $\mathbb{M}$. As a weighting function, $\omega(x, w^*, \epsilon)$ also satisfies that given any values of x such that $\omega(x, \cdot, \cdot)$ could obtain non-zero values, $\int_{w^*} \int_{\epsilon} \omega(x, w^*, \epsilon) f_{\epsilon, W^*|X=x}(\epsilon, w^*) d\epsilon dw^* = 1$. The specific form of ω is up to the researcher's choice, however, to ensure easy calculation of $\omega(x, w^*, s(r(x, w^*), \delta))$ on the right-hand side of (19), I require that function $\omega(x, w^*, \epsilon)$ satisfy the following restriction: $\omega(x, w^*, \epsilon) = \tilde{\omega}(x, w^*)$ if $\epsilon \in [q_{x,w^*}(\tau_l), q_{x,w^*}(\tau_u)]$ and $\omega(x, w^*, \epsilon) = 0$ otherwise, where $\tilde{\omega}(x, w^*)$ is a known function with compact support, and $q_{x,w^*}(\tau)$ denotes the conditional-τ quantile of ϵ given $X = x, W^* = w^*$. The specific form of function $\tilde{w}$ and the specific values of τ_l and τ_u are up to the researcher's choice. Under this restriction, $\omega(x, w^*, s(r(x, w^*), \delta))$ is equal to $\tilde{w}(x, w^*)$ if δ is between τ_l and τ_u, and is equal to 0 otherwise. An example of function $\omega(x, w^*, \epsilon)$ is when $\tilde{\omega}(x, w^*)$ is a constant equal to $\left[(\tau_u - \tau_l) \int_{\underline{w}^*}^{\bar{w}^*} f_{W^*|X=x}(w^*) dw^*\right]^{-1}$. The **WLAR** in this case can be interpreted as the Local Average Response of a subgroup of individuals whose unobserved ϵ is between its τ_l and τ_u conditional quantile given X, W^* and whose W^* is between $\underline{w}^*$ and $\bar{w}^*$. Note that even though X is correlated with ϵ, neither the distribution of $\epsilon|X$ nor the subgroup of individuals changes as I take the derivative with respect to x. This means that my objective of interest is the same group of people when studying the effect of the counterfactual changes of the endogenous X. The **WLAR** can also be written as a functional of $g_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y, x, w^*)$. The explicit form can be derived from Lemma 3 and equation (12).

So far I have written the parameters of interest as functionals of $g_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y, x, w^*)$ in explicit forms. Next I focus on the estimation of $g_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y, x, w^*)$. To deal with the well-known ill-posed inverse problem when inverting a convolution operator, I base my estimator on a smoothed version of $g_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y, x, w^*)$:

$$g_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y, x, w^*, h_1) \equiv \int \frac{1}{h_1} K\left(\frac{\tilde{w}^* - w^*}{h_1}\right) g_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y, x, \tilde{w}^*) d\tilde{w}^*,$$

and use kernel functions satisfying the following assumption:

Assumption 9. *The kernel functions $K : \mathbb{R} \to \mathbb{R}$, $G_Y : \mathbb{R} \to \mathbb{R}$ and $G_X : \mathbb{R} \to \mathbb{R}$ are* *measurable, symmetric.* $\int K(x)dx = 1$, $\int G_Y(y)dy = 1$, $\int G_X(x)dx = 1$. G_Y, G_X *are differentiable.* G_Y, G_X, *and their derivatives are bounded, and denote the maximum of these bounds as* $\bar{G}$. *Their Fourier transforms* $\xi \to \phi_K(\xi)$, $\xi \to \phi_{G_Y}(\xi)$ *and* $\xi \to \phi_{G_X}(\xi)$ *obey: (i)* ϕ_F *is compactly supported (without loss of generality, the support is* $[-1, 1]$*) for* $F \in \{K, G_X, G_Y\}$*; and (ii) there exists* $\bar{\xi}_F$ *such that* $\phi_F(\xi) = 1$ *for* $|\xi| \leq \bar{\xi}_F$*, where* $F \in \{K, G_X, G_Y\}$.

By assuming (ii), I use flat-top kernels proposed by Politis and Romano (1999). The benefit of using flat-top kernels is that the rate of decrease of the bias term will not be affected by the order of the kernel, and will only be affected by the smoothness of the function to be estimated. The restriction of compact support of the Fourier transform is without loss of generality. As discussed by Schennach (2004b), given any kernel K, one can always create a modified kernel $\tilde{K}$ that satisfies the assumption by using a "windowing" function. For example,

$$\phi_{\tilde{K}}(t) = W(t)\phi_K(t).$$

where

$$W(t) = \begin{cases} 1 & \text{if } |t| \leq \bar{t} \\ \left(1 + \exp\left((1-\bar{t})\left((1-|t|)^{-1} - (|t|-\bar{t})^{-1}\right)\right)\right)^{-1} & \text{if } 1 \geq |t| > \bar{t} \\ 0 & \text{if } |t| > 1 \end{cases}$$

for some $\bar{t} \in (0, 1)$.

Assumption 10. $E[|\Delta W_1|] < \infty$.

The following lemma shows that the smoothed version $g_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y, x, w^*, h_1)$ can be written as an expression involving the Fourier transform of the kernel function and the characteristic functions of different variables. This expression makes it more straightforward to introduce the estimator that will appear soon below.

Lemma 4. *For* $(y, x, w^*) \in \mathbb{S}_{(Y,X,W^*)}$ *and* $h > 0$*, let*

$$g_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y, x, w^*, h_1) \equiv \int \frac{1}{h_1} K\left(\frac{\tilde{w}^* - w^*}{h_1}\right) g_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y, x, \tilde{w}^*) d\tilde{w}^*.$$

where K *satisfies Assumption 9. Then under Assumptions 5, 6, and 10,*

$$g_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y, x, w^*, h_1) = \frac{1}{2\pi} \int (-\mathbf{i}t)^{\lambda_1} e^{-\mathbf{i}tw^*} \phi_K(h_1 t) \frac{\phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(y,x,\cdot)}(t)\phi_{W^*}(t)}{\phi_{W_2}(t)} dt$$

where $f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(y, x, w_2) \equiv \frac{\partial^{\lambda_2} f_{Y,X,W_2}(y,x,w_2)}{\partial y^{\lambda_{2,1}} \partial x^{\lambda_{2,2}}}$.

Proof. See Appendix B. ■

I now define my estimator for $g_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y, x, w^*, h_1)$ by replacing $\phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(y,x,\cdot)}(t)$, $\phi_{W^*}(t)$, and $\phi_{W_2}(t)$ in Lemma 4 with their sample analogues. Formally,

Definition 1. *Let* $h_n \equiv (h_{1n}, h_{2n,1}, h_{2n,2}) \to 0$ *as* $n \to 0$*. The estimator for* $g_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y, x, w^*)$ *is defined as*

$$\hat{g}_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y, x, w^*, h_n) \equiv \frac{1}{2\pi} \int (-\mathbf{i}t)^{\lambda_1} e^{-\mathbf{i}tw^*} \phi_K(h_{1n} t) \frac{\hat{\phi}_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(y,x,\cdot)}(t)\hat{\phi}_{W^*}(t)}{\hat{\phi}_{W_2}(t)} dt,$$

where

$$\hat{\phi}_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(y,x,\cdot)}(t) \equiv \hat{E}\left[e^{\mathbf{i}tW_2} \frac{1}{h^{1+\lambda_{2,1}}_{2n,1} h^{1+\lambda_{2,2}}_{2n,2}} G^{(\lambda_{2,1})}_Y\left(\frac{y - Y}{h_{2n,1}}\right) G^{(\lambda_{2,2})}_X\left(\frac{x - X}{h_{2n,2}}\right)\right],$$

$$\hat{\phi}_{W_2}(t) \equiv \hat{E}\left[e^{\mathbf{i}tW_2}\right],$$

$$\hat{\phi}_{W^*}(t) \equiv \exp\left(\int_0^t \frac{\hat{E}\left[iW_1 e^{i\xi W_2}\right]}{\hat{E}\left[e^{i\xi W_2}\right]} d\xi\right),$$

where $\hat{E}$ *denotes the sample average,* G_Y, G_X *denotes the kernel functions for* Y, X*, respectively, and* $G^{(\lambda_{2,2})}_X(x) \equiv \frac{\partial^{\lambda_{2,2}} G_X(x)}{\partial x^{\lambda_{2,2}}}$*.* G_Y, G_X *are not necessarily the same as* K*.*

Having defined $\hat{g}_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y, x, w^*, h_n)$, I can now define the estimator for $\rho(\bar{y}, \bar{x})$ by replacing $g_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y, x, w^*)$ with its estimator on the right-hand-side of Equation (11). Note that by Fubini's Theorem

$$\int_{-\infty}^{y} \hat{g}_{\lambda_1,0,0,\lambda_{2,2}}(s, x, w^*, h_n) ds = \frac{1}{2\pi} \int (-\mathbf{i}t)^{\lambda_1} e^{-\mathbf{i}tw^*} \phi_K(h_{1n} t) \frac{\int_{-\infty}^{y} \hat{\phi}_{f^{(0,\lambda_{2,2})}_{Y,X,W_2}(s,x,\cdot)}(t) ds \hat{\phi}_{W^*}(t)}{\hat{\phi}_{W_2}(t)} dt$$

$$\int_{-\infty}^{\infty} \hat{g}_{\lambda_1,0,0,\lambda_{2,2}}(y,x,w^*,h_n)dy = \frac{1}{2\pi}\int (-\mathbf{i}t)^{\lambda_1} e^{-\mathbf{i}tw^*}\phi_K(h_{1n}t)\frac{\hat{\phi}_{f^{(0,\lambda_{2,2})}_{X,W_2}(x,\cdot)}(t)\hat{\phi}_{W^*}(t)}{\hat{\phi}_{W_2}(t)}dt$$
$$\int_{-\infty}^{x}\int_{-\infty}^{\infty} \hat{g}_{\lambda_1,0,0}(y,s,w^*,h_n)dyds = \frac{1}{2\pi}\int (-\mathbf{i}t)^{\lambda_1} e^{-\mathbf{i}tw^*}\phi_K(h_{1n}t)\frac{\int_{-\infty}^{x}\hat{\phi}_{f_{X,W_2}(s,\cdot)}(t)ds\hat{\phi}_{W^*}(t)}{\hat{\phi}_{W_2}(t)}dt$$
$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \hat{g}_{\lambda_1,0,0}(y,x,w^*,h_n)dydx = \frac{1}{2\pi}\int (-\mathbf{i}t)^{\lambda_1} e^{-\mathbf{i}tw^*}\phi_K(h_{1n}t)\hat{\phi}_{W^*}(t)dt,$$

where I've let $\tilde{G}_Y(y) \equiv \int_{-\infty}^{y} G_Y(u)\,du$ and $\tilde{G}_X(x) \equiv \int_{-\infty}^{x} G_X(u)\,du$ denote the kernel CDFs, and I've defined:

$$\int_{-\infty}^{y}\hat{\phi}_{f^{(0,\lambda_{2,2})}_{Y,X,W_2}(s,x,\cdot)}(t)ds = \hat{E}\left[e^{\mathbf{i}tW_2}\frac{1}{h_{2n,2}^{1+\lambda_{2,2}}}\tilde{G}_Y\left(\frac{y-Y}{h_{2n,1}}\right)G_X^{(\lambda_{2,2})}\left(\frac{x-X}{h_{2n,2}}\right)\right]$$
$$\hat{\phi}_{f^{(\lambda_{2,2})}_{X,W_2}(x,\cdot)}(t)ds = \hat{E}\left[e^{\mathbf{i}tW_2}\frac{1}{h_{2n,2}^{1+\lambda_{2,2}}}G_X^{(\lambda_{2,2})}\left(\frac{x-X}{h_{2n,2}}\right)\right]$$
$$\int_{-\infty}^{x}\hat{\phi}_{f_{X,W_2}(s,\cdot)}(t)ds = \hat{E}\left[e^{\mathbf{i}tW_2}\tilde{G}_X\left(\frac{x-X}{h_{2n,2}}\right)\right].$$

4.2 The Asymptotic Properties of the Estimator

In this subsection, I analyze the asymptotic properties for $\hat{g}_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h_n)$. I begin by decomposing the difference between the estimator and the true value of the parameter as the sum of a bias term, a variance term, and a remainder term.

Lemma 5. *Suppose $\{Y_i, X_i, W_i^*, \Delta W_{1,i}, W_{2,i}\}$ is an IID sequence satisfying Assumptions 3, 4, 5, 6, and 10, and that Assumption 9 holds. Then for $(y,x,w^*) \in \mathbb{S}_{(Y,X,W^*)}$, and $h > 0$,*

$$\begin{aligned}&\hat{g}_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h) - g_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*)\\ &= B_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h) + L_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h) + R_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h)\end{aligned}$$

where $B_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h)$ is the bias term admitting the linear representation:

$$\begin{aligned}&B_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h)\\ &= E\left[\bar{g}_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h)\right] - g_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*)\\ &= g_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h_1) - g_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*)\\ &\quad + \frac{1}{2\pi}\int (-\mathbf{i}t)^{\lambda_1} e^{-\mathbf{i}tw^*}\phi_K(h_{1n}t)\frac{E\left[\hat{\phi}_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(y,x,\cdot)}(t) - \phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(y,x,\cdot)}(t)\right]\phi_{W^*}(t)}{\phi_{W_2}(t)}dt;\end{aligned}$$

$L_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h)$ is the variance term admitting the linear representation:

$$\begin{aligned}
&L_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h)\\
&=\bar{g}_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h)-E\left[\bar{g}_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h)\right]\\
&=\hat{E}\Bigg[\int\Psi_{1,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,y,x,w^*,h_1)\left(W_1e^{\mathbf{i}\xi W_2}-E[W_1e^{\mathbf{i}\xi W_2}]\right)d\xi\\
&\quad+\int\Psi_{2,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,y,x,w^*,h_1)\left(e^{\mathbf{i}\xi W_2}-E[e^{\mathbf{i}\xi W_2}]\right)d\xi\\
&\quad+\int\Psi_{3,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,y,x,w^*,h_1)\times\\
&\qquad\Bigg(e^{\mathbf{i}\xi W_2}\frac{1}{h_{2,1}^{1+\lambda_{2,1}}h_{2,2}^{1+\lambda_{2,2}}}G_Y^{(\lambda_{2,1})}\left(\frac{y-Y}{h_{2,1}}\right)G_X^{(\lambda_{2,2})}\left(\frac{x-X}{h_{2,2}}\right)\\
&\qquad\quad-E\left[e^{\mathbf{i}\xi W_2}\frac{1}{h_{2,1}^{1+\lambda_{2,1}}h_{2,2}^{1+\lambda_{2,2}}}G_Y^{(\lambda_{2,1})}\left(\frac{y-Y}{h_{2,1}}\right)G_X^{(\lambda_{2,2})}\left(\frac{x-X}{h_{2,2}}\right)\right]\Bigg)d\xi\Bigg]\\
&=\hat{E}[l_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h;Y,X,W_1,W_2)],
\end{aligned}$$

where I've let $\theta(\xi)\equiv E\left[W_1e^{\mathbf{i}\xi W_2}\right]$ and defined

$$\begin{aligned}
&\Psi_{1,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,y,x,w^*,h_1)=\frac{1}{2\pi}\frac{\mathbf{i}}{\phi_{W_2}(\xi)}\int_\xi^{\pm\infty}(-\mathbf{i}t)^{\lambda_1}e^{-\mathbf{i}tw^*}\phi_K(h_1t)\phi_{f_{Y,X,W^*}^{(\lambda_{2,1},\lambda_{2,2})}(y,x,\cdot)}(t)dt\\
&\Psi_{2,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,y,x,w^*,h_1)\\
&=-\frac{1}{2\pi}\frac{\mathbf{i}\theta(\xi)}{(\phi_{W_2}(\xi))^2}\int_\xi^{\pm\infty}(-\mathbf{i}t)^{\lambda_1}e^{-\mathbf{i}tw^*}\phi_K(h_1t)\phi_{f_{Y,X,W^*}^{(\lambda_{2,1},\lambda_{2,2})}(y,x,\cdot)}(t)dt\\
&\quad-\frac{1}{2\pi}(-\mathbf{i}t)^{\lambda_1}e^{-\mathbf{i}\xi w^*}\phi_K(h_1\xi)\frac{\phi_{f_{Y,X,W^*}^{(\lambda_{2,1},\lambda_{2,2})}(y,x,\cdot)}(\xi)}{\phi_{W_2}(\xi)}\\
&\Psi_{3,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,y,x,w^*,h_1)=\frac{1}{2\pi}(-\mathbf{i}t)^{\lambda_1}e^{-\mathbf{i}\xi w^*}\phi_K(h_1\xi)\frac{\phi_{W^*}(\xi)}{\phi_{W_2}(\xi)}.
\end{aligned}$$

where for a given function $\zeta\to f(\zeta)$, I've written $\int_\xi^{\pm\infty}f(\zeta)d\zeta\equiv\lim_{c\to+\infty}\int_\xi^{c\xi}f(\zeta)d\zeta$; and $R_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h)$ is the remainder term:

$$R_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h)=\hat{g}_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h)-\bar{g}_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h).$$

Proof. See Appendix B. ■

To derive the uniform rate of convergence, I impose some assumptions on the smoothness of $g_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*)$, stated in terms of the rate of decay of the tail of its Fourier transform:

Assumption 11. *(i) There exist constants $C_\phi > 0$, $\alpha_\phi \leq 0$, $\beta_\phi \geq 0$ and $\gamma_\phi \in \mathbb{R}$ such that $\beta_\phi \gamma_\phi \geq 0$ and for $j = 1, 2$*

$$\max_{\lambda_2 \in \{0,1\}} \sup_{(y,x) \in \mathbb{S}_{(Y,X)}} \left| \phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W^*}(y,x,\cdot)}(t) \right| \leq C_\phi (1 + |t|)^{\gamma_\phi} \exp\left(\alpha_\phi |t|^{\beta_\phi}\right)$$
$$|\phi_{W^*}(t)| \leq C_\phi (1 + |t|)^{\gamma_\phi} \exp\left(\alpha_\phi |t|^{\beta_\phi}\right),$$

where $f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(y, x, w_2) \equiv \frac{\partial^{\lambda_2} f_{Y,X,W_2}(y,x,w_2)}{\partial y^{\lambda_{2,1}} \partial x^{\lambda_{2,2}}}$. Moreover, if $\beta_\phi = 0$, then for given $\lambda \in \{0, 1\}$, $\gamma_\phi < -\lambda_1 - 1$.
(ii) Denote the Fourier transform of $f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W^*}(y, x, w^*)$ as $\phi_{f^{(\lambda_{2,1},\lambda_{2,2})}}(t_1, t_2, \xi)$. There exists constants $C_{f_1} > 0$, $C_{f_2} > 0$, $\alpha_{f_1} \leq 0$, $\alpha_{f_2} \leq 0$, $\beta_{f_1} \geq 0$, $\beta_{f_2} \geq 0$ and $\gamma_{f_1} \in \mathbb{R}$, $\gamma_{f_2} \in \mathbb{R}$ such that $\beta_{f_1}\gamma_{f_1} \geq 0$, $\beta_{f_2}\gamma_{f_2} \geq 0$, and

$$\left| \phi_{f^{(\lambda_{2,1},\lambda_{2,2})}}(t_1, t_2, \xi) \right| \leq C_{f_1} C_{f_2} (1 + |t_1|)^{\gamma_{f_1}} (1 + |t_2|)^{\gamma_{f_2}} \exp\left(\alpha_{f_1} |t_1|^{\beta_{f_1}} + \alpha_{f_2} |t_2|^{\beta_{f_2}}\right) |\phi_{W^*}(t)|.$$

Moreover, if $\beta_{f_1} = 0$, $\gamma_{f_1} < -1$, and if $\beta_{f_2} = 0$, $\gamma_{f_2} < -1$.

In Assumption 11(i) I impose the same bound for the tail behavior of $\phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W^*}(y,x,\cdot)}$ and $\phi_{W^*}(t)$. This is without loss of generality since they have the same effect on the convergence rate. Assumption 11(ii) is similar to Assumption A7 in Li (2002), and can be viewed as a generalization of smoothness condition from the univariate case to the multivariate case.

I next state the first main result of this paper, the uniform asymptotic rate of the bias term:

Theorem 1. *Let the conditions of Lemma 5 hold with $\{Y_i, X_i, W_i^*, \Delta W_{1,i}, \Delta W_{2,i}\}$ IID, and suppose in addition that Assumption 11 holds. Then for $h > 0$,*

$$\sup_{(y,x,w^*) \in \mathbb{S}_{(Y,X,W^*)}} \left| B_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y, x, w^*, h) \right| = O\left(\left(h_1^{-1}\right)^{1+\gamma_\phi+\lambda_1} \exp\left(\alpha_\phi \bar{\zeta}_K^{\beta_\phi} \left(h_1^{-1}\right)^{\beta_\phi}\right)\right).$$

Proof. See Appendix B. ■

I impose the following assumption to ensure finite variance:

Assumption 12. *For some $\delta > 0$, $E[|W_1|^{2+\delta}] < \infty$, $\sup_{w_2 \in \mathbb{S}_{W_2}} E\left[|W_1|^{2+\delta} \mid W_2 = w_2\right] < \infty$*
.

To derive the rate for $L_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h)$, I impose the following assumption on the tail behavior of the Fourier transforms involved. These are common in the deconvolution literature (e.g. Fan (1991b), Fan and Truong (1993), and Schennach (2004b)).

Assumption 13. *(i) There exist constants $C_2>0$, $\alpha_2\leq 0$, $\beta_2\geq\beta_\phi\geq 0$ and $\gamma_2\in\mathbb{R}$ such that $\beta_2\gamma_2\geq 0$ and*

$$|\phi_{W_2}(t)|\geq C_2(1+|t|)^{\gamma_2}\exp\left(\alpha_2|t|^{\beta_2}\right).$$

Moreover, if $\beta_2=0$, $\lambda_1-\gamma_2+\gamma_\phi>0$.

(ii) There exist constants $C_*>0$ and $\gamma_*\geq 0$ such that

$$\left|\frac{\phi'_{W^*}(t)}{\phi_{W^*}(t)}\right|\leq C_*\left(1+|t|\right)^{\gamma_*}.$$

I explicitly impose $\beta_2\geq\beta_\phi$ because

$$\begin{aligned}C_\phi\left(1+|t|\right)^{\gamma_\phi}\exp\left(\alpha_{W^*}|t|^{\beta_\phi}\right)&\geq|\phi_{W^*}(t)|=\left|E[e^{\mathbf{i}tW^*}]\right|\\&\geq\left|E[e^{\mathbf{i}tW^*}]\right|\left|E[e^{\mathbf{i}t\Delta W_2}]\right|\geq\left|E[e^{\mathbf{i}tW_2}]\right|\\&=|\phi_{W_2}(t)|\geq C_2\left(1+|t|\right)^{\gamma_2}\exp\left(\alpha_2|t|^{\beta_2}\right).\end{aligned}$$

The following theorem states the asymptotic properties of the linear term $L_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h)$ and facilitates the analysis of various quantities of interest later.

Theorem 2. *Suppose the conditions of Lemma 5 hold. (i) Then for each $(y,x,w^*)\in\mathbb{S}_{(y,x,w^*)}$, $E[L_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h)]=0$, and if Assumption 12 also holds, then $E[L^2_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h)]=n^{-1}\Omega_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h)$, where $\Omega_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h)\equiv E[\left(l_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h;Y,X,W_1,W_2)\right)^2]$.*

Further, if Assumption 11 and 13 also holds then

$$\begin{aligned}&\sqrt{\sup_{(y,x,w^*)\in\mathbb{S}_{(Y,X,W^*)}}\Omega_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h)}\\&=O\Bigg(\max\left\{(h_1^{-1})^{1+\gamma_*},(h_{2,1}^{-1})^{1+\lambda_{2,1}}(h_{2,2}^{-1})^{1+\lambda_{2,2}}\right\}\\&\qquad(h_1^{-1})^{1-\gamma_2+\gamma_\phi+\lambda_1}\exp\left(\left(\alpha_\phi\mathbf{1}\{\beta_\phi=\beta_2\}-\alpha_2\right)\left(h_1^{-1}\right)^{\beta_2}\right)\Bigg).\end{aligned}\tag{20}$$

I also have

$$\sup_{(y,x,w^*)\in\mathbb{S}_{(Y,X,W^*)}}\left|L_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h)\right|$$

$$= O\Big(n^{-\frac{1}{2}} \max\left\{(h_1^{-1})^{1+\gamma_*}, (h_{2,1}^{-1})^{1+\lambda_{2,1}}(h_{2,2}^{-1})^{1+\lambda_{2,2}}\right\}$$
$$(h_1^{-1})^{1-\gamma_2+\gamma_\phi+\lambda_1} \exp\left(\left(\alpha_\phi \mathbf{1}\{\beta_\phi = \beta_2\} - \alpha_2\right)\left(h_1^{-1}\right)^{\beta_2}\right)\Big). \tag{21}$$

(ii) If Assumption 12 also holds, and if for each $(y, x, w^*) \in \mathbb{S}_{(y,x,w^*)}$, $\Omega_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y, x, w^*, h_n) > 0$ *for all n sufficiently large, then for each* $(y, x, w^*) \in \mathbb{S}_{(y,x,w^*)}$

$$n^{1/2}\left(\Omega_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y, x, w^*, h_n)\right)^{-1/2} L_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y, x, w^*, h_n) \xrightarrow{d} N(0, 1).$$

Proof. See Appendix B. ■

Finally, I bound the remainder term. To do that, I first impose some restrictions on the moments of W_2:

Assumption 14. $E[|W_2|] < \infty, E[|W_1W_2|] < \infty$.

I then impose the following bounds for the bandwidths:

Assumption 15. *If* $\beta_2 = 0$ *in Assumption 13, then* $h_{1n}^{-1} = O\left(n^{(4+4\gamma_*-4\gamma_2)^{-1}-\eta}\right)$ *and* $h_{2n,j}^{-1} = O\left(n^{(16+8\lambda_2)^{-1}-\eta}\right)$*, for some* $\eta > 0$*,* $j = 1, 2$*; otherwise* $h_{1n}^{-1} = O\left((\ln n)^{\beta_2^{-1}-\eta}\right)$ *and* $h_{2n,j}^{-1} = O\left(n^{(8+4\lambda_2)^{-1}-\eta}\right)$*, for some* $\eta > 0$*,* $j = 1, 2$*.*

Theorem 3. *(i) Suppose the conditions of Lemma 5 and Assumption 11, 12, 13, and 14 hold. Then*

$$\sup_{(y,x,w^*)\in\mathbb{S}_{(Y,X,W^*)}} \left|R_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y, x, w^*, h_n)\right|$$
$$= o_p\left(h_{2,1}^{-2}(h_{2,2}^{-1})^{2+2\lambda_2} n^{-1+2\epsilon}(h_{1n}^{-1})^{1+\gamma_*-\gamma_2} \exp\left(-\alpha_2\left(h_{1n}^{-1}\right)^{\beta_2}\right)\right)$$
$$\times O\Big(\max\left\{(h_1^{-1})^{1+\gamma_*}, (h_{2,1}^{-1})^{1+\lambda_{2,1}}(h_{2,2}^{-1})^{1+\lambda_{2,2}}\right\}$$
$$(h_1^{-1})^{1-\gamma_2+\gamma_\phi+\lambda_1} \exp\left(\left(\alpha_\phi \mathbf{1}\{\beta_\phi = \beta_2\} - \alpha_2\right)\left(h_1^{-1}\right)^{\beta_2}\right)\Big)$$

for arbitrarily small $\epsilon > 0$*. (ii) If Assumption 15 also holds, then*

$$\sup_{(y,x,w^*)\in\mathbb{S}_{(Y,X,W^*)}} \left|R_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y, x, w^*, h_n)\right| =$$
$$o_p\Big(n^{-\frac{1}{2}} \max\left\{(h_1^{-1})^{1+\gamma_*}, (h_{2,1}^{-1})^{1+\lambda_{2,1}}(h_{2,2}^{-1})^{1+\lambda_{2,2}}\right\}$$
$$(h_1^{-1})^{1-\gamma_2+\gamma_\phi+\lambda_1} \exp\left(\left(\alpha_\phi \mathbf{1}\{\beta_\phi = \beta_2\} - \alpha_2\right)\left(h_1^{-1}\right)^{\beta_2}\right)\Big).$$

Proof. See Appendix B. ■

Collecting the results from Theorem 1,2 and 3 yields a straightforward corollary of the uniform rate of convergence of $\hat{g}_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h_n)$:

Corollary 1. *If the conditions of Theorem 3 (ii) hold, then*

$$\begin{aligned}
&\sup_{(y,x,w^*)\in\mathbb{S}_{(Y,X,W^*)}} \left|\hat{g}_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h_n) - g_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*)\right| \\
&= O_p\left(\epsilon_{n,\lambda_1}\right) + O_p\left(\tilde{\epsilon}_{n,\lambda_1,\lambda_{2,1},\lambda_{2,2}}\right).
\end{aligned}$$

where

$$\begin{aligned}
&\epsilon_{n,\lambda_1} \equiv \left(h_1^{-1}\right)^{1+\gamma_\phi+\lambda_1} \exp\left(\alpha_\phi \bar{\zeta}_K^{\beta_\phi} \left(h_1^{-1}\right)^{\beta_\phi}\right) \\
&\tilde{\epsilon}_{n,\lambda_1,\lambda_{2,1},\lambda_{2,2}} \\
&\equiv n^{-1/2} \max\left\{(h_1^{-1})^{1+\gamma_*}, (h_{2,1}^{-1})^{1+\lambda_{2,1}}(h_{2,2}^{-1})^{1+\lambda_{2,2}}\right\} (h_1^{-1})^{1-\gamma_2+\gamma_\phi+\lambda_1} \\
&\qquad\qquad\qquad\qquad \times \exp\left(\left(\alpha_\phi \mathbf{1}\{\beta_\phi=\beta_2\} - \alpha_2\right)\left(h_1^{-1}\right)^{\beta_2}\right).
\end{aligned}$$

The uniform (over the whole support) rate of convergence of the kernel estimators of the density or its derivatives has been considered in Andrews (1995), Hansen (2008) and Schennach, White and Chalak (2012)[2]. Hansen (2008) obtained a faster rate than Andrews (1995) and Schennach, White and Chalak (2012) (Theorem 3.2, for the no-measurement error case), but requires more assumptions on the kernel functions (their Assumption 1 and 3), including bounded support or an integrable tail, which, unfortunately, are not satisfied by infinite order kernels. Andrews (1995)'s conclusion is stated assuming a kernel with finite order. Infinite order kernels are not essential when there is no measurement error (like in Andrews (1995) and Hansen (2008)), but are especially advantageous when there are measurement errors. If infinite order kernels are used, only the smoothness of the functions, but not the order of the kernels will affect the rate of convergence. Corollary 1 is similar to Corollary 4.7 in Schennach, White and Chalak (2012).

With the uniform rate of convergence of $\hat{g}$, I can then derive the uniform rate of convergence of the plug-in estimators of the structural derivative $\rho(y,x)$ and the **WLAR**, denoted as $\widehat{\rho(y,x)}$ and $\widehat{\mathbf{WLAR}}(x)$, respectively.

[2]Schennach, White and Chalak (2012) studies the case both with and without measurement errors. Their Theorem 3.2 corresponds to the no-measurement error case, and Corollary 4.7 corresponds to the case with measurement errors.

Theorem 4. *Suppose that* $\left\{Y_j, X_j, W_j^*, \Delta W_{1,j}, W_{2,j}\right\}$ *is an IID sequence satisfying the conditions of Corollary 1 with* $\lambda_1, \lambda_{2,1}, \lambda_{2,2} \in \{0,1\}$ *and* $\max\{\lambda_1, \lambda_{2,1}, \lambda_{2,2}\} \leq 1$. *Suppose in addition, Assumption 7 and 8 hold.*

(i) Define $\mathbb{S}_{\tau_n} \equiv \left\{(y, x, w^*) \in \mathbb{S}_{(Y,X,W^*)} : \left|\frac{\partial F_{X|W^*=w^*}(x)}{\partial w^*}\right| > \tau_n \text{ and } f_{Y,X,W^*}(y, x, w^*) > \tau_n\right\}$. *Then*

$$\sup_{(y,x,w^*)\in\mathbb{S}_{\tau_n}} \left|\widehat{\rho(y,x)} - \rho(y,x)\right| \leq O_p\left(\tilde{\epsilon}_{n,0,0,1}\right) + \frac{O_p\left(\epsilon_{n,1}\right) + O_p\left(\tilde{\epsilon}_{n,1,0,0}\right)}{\tau_n^2}.$$

and there exists $\{\tau_n\}$ *such that* $\tau_n > 0, \tau_n \to 0$ *as* $n \to \infty$, *and*

$$\sup_{(y,x,w^*)\in\mathbb{S}_{\tau_n}} \left|\widehat{\rho(y,x)} - \rho(y,x)\right| = o_p(1).$$

(ii)

$$\sup_{x\in\mathbb{S}_X} \left|\widehat{\mathbf{WLAR}}(x) - \mathbf{WLAR}(x)\right| \leq O_p\left(\epsilon_{n,1}\right) + O_p\left(\tilde{\epsilon}_{n,0,0,1}\right) + O_p\left(\tilde{\epsilon}_{n,1,0,0}\right).$$

Proof. See Appendix B. ■

In Appendix C, I state an asymptotic normality result for $\hat{g}_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y, x, w^*, h_n)$ and $\widehat{\rho(y,x)}$. The proof of them requires a lower bound on $\Omega_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y, x, w^*.h_n)$ relative to $B_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y, x, w^*.h_n)$ and $R_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y, x, w^*.h_n)$. The assumptions to ensure the lower bound are stated at a high level, and more primitive sufficient conditions need to be derived.

5 Monte Carlo Simulations

In this section, I conduct some Monte Carlo simulation exercises to study the finite sample performance of my proposed estimators. I consider two simulation designs. Design 1 is when both of the equations are linear, and Design 2 is when the equations are nonlinear.

Design 1

$$Y = 0.25X + 0.25\epsilon$$

$$X = W^* + \eta$$

Design 2

$$Y = \ln(\exp(X + \epsilon) + 1)$$
$$X = \frac{3^3}{4^4}\frac{W^{*4}}{(-\eta)^3}$$

In both cases, ϵ and η are correlated through a common component θ: $\epsilon = \theta + \epsilon_1$, and $\eta = \theta + \eta_1$, where θ, ϵ_1 and η_1 are mutually independent, and are independent with W^*. In both designs, I have two error-laden measurements $W_1 = W^* + \Delta W_1, W_2 = W^* + \Delta W_2$, where ΔW_1 and ΔW_2 are independent and they are independent with all other variables. The distributions of variables in both designs are listed in Table 1 below:

Table 1: Distributions of Variables

Variables	Design 1 (linear)	Design 2 (nonlinear)
W^*	$N(0, \sqrt{0.5})$	$N(6, 1)$
θ	$N(0, 0.5)$	$N(-3, \sqrt{0.5})$
ϵ_1	$N(0, \sqrt{0.75})$	$N(3, \sqrt{0.5})$
η_1	$N(0, 0.5)$	$N(-3, \sqrt{0.5})$
ϵ	$N(0, 1)$	$N(0, 1)$
η	$N(0, \sqrt{0.5})$	$N(-6, 1)$
ΔW_1	$N(0, \sqrt{0.5})$	$N(0, \sqrt{0.5})$
ΔW_2	$\chi^2(2) - 2$	$\chi^2(2) - 2$

In the estimation, I use the following flat-top kernel[3] proposed by Politis and Romano (1999):

$$K(x) = \frac{\sin^2(2\pi u) - \sin^2(\pi u)}{\pi^2 u^2}$$

I use the sample size of 500 and replicate the estimation 500 times.

First I show the performance of my proposed method versus two other methods for estimating the structural derivative: (1) 2SLS using the error-laden measurement W_2 as IV; (2) a plug-in estimator replacing W^* with W_2 in my identification equation (11). Method

[3] I may use different flat-top kernels for G_X and G_Y. For simplicity I use the same flat-top kernel.

(1) will be valid under the linear design (Design 1), since W_2, although error-laden, still satisfies the exclusion restriction and the rank condition for linear IV estimation. However, because of its misspecification of the model, it won't be valid under the nonlinear design (Design 2). Method (2) won't be valid under either the linear or nonlinear design, as the estimator will converge to a population value that is not equal to the right-hand side of equation (11) in general. I estimate $\rho(y, x)$ evaluated at $y = 0, x = 0, w^* = 0.7$ for Design 1, and $y = 0.6, x = 0.6, w^* = 7$ for Design 2, yielding a true value of 0.25 and 0.4512, respectively. There are three bandwidths used in the estimation: h_1, h_{21}, h_{22}. The latter two correspond to estimating the unknown distribution of observed variables X and Y. I thus choose the values of h_{21} and h_{22} by cross-validation (i.e. minimizing the estimated MISE of $\hat{f}_{Y,X}(y, x)$). For the bandwidth h_1, I scan a set of values ranging from 0.5 to 3 for my proposed estimator and a set of values ranging from 0.5 to 6 for the method that uses the error-laden measurement. Table 2, 3, and 4 show the MSE, VAR, and the absolute value of BIAS of my proposed estimator, Method (2), and Method (1), respectively.

From Table 3, one can see that the bias from the error-contaminated estimator does not shrink toward zero as bandwidth decreases. Comparing results from Table 2 and Table 3, one can see that my estimator also gives smaller variances than the error-contaminated method. As a result, my estimator performs better than the error-contaminated method in terms of MSE under both the linear and the nonlinear designs. Comparing Table 2 with Table 4, one can see that under the linear design, both my estimator and 2SLS work well except that my estimator has a slightly larger variance. Under the nonlinear design, the bias and MSE of 2SLS are much larger than my estimator, which aligns with the theory.

Table 5 below shows the Monte Carlo simulation results of my proposed estimator as a function of the sample size. The value of h_{21} and h_{22} are selected by cross-validation for the respective sample sizes and the optimal values of h_1 are selected by minimizing the MSE in the corresponding set of values the same as when $N = 500$. The MSE, VAR, and Bias decrease as the sample size increases, which is in accordance with the theory.

Table 2: Monte Carlo simulation results of my proposed estimator for $\rho(y,x)$

Design 1 (linear)						Design 2 (nonlinear)					
h_1	h_{21}	h_{22}	MSE	VAR	abs(BIAS)	h_1	h_{21}	h_{22}	MSE	VAR	abs(BIAS)
0.5	1.05	2.92	0.01072	0.01072	0.00009	0.5	2.09	0.88	0.31292	0.30489	0.08961
0.75	1.05	2.92	0.00766	0.00766	0.00050	0.75	2.09	0.88	0.17724	0.16339	0.11769
1	1.05	2.92	0.00728	0.00728	0.00138	1	2.09	0.88	0.15425	0.13973	0.12050
1.25	1.05	2.92	0.00738	0.00738	0.00253	1.25	2.09	0.88	0.14824	0.13360	0.12098
1.5	1.05	2.92	0.00755	0.00754	0.00328	1.5	2.09	0.88	0.14611	0.13129	0.12175
1.75	1.05	2.92	0.00772	0.00771	0.00376	1.75	2.09	0.88	0.14540	0.13034	0.12273
2	1.05	2.92	0.00789	0.00787	0.00408	2	2.09	0.88	0.14540	0.13008	0.12380
2.25	1.05	2.92	0.00806	0.00804	0.00431	2.25	2.09	0.88	0.14581	0.13023	0.12482
2.5	1.05	2.92	0.00822	0.00820	0.00447	2.5	2.09	0.88	0.14640	0.13060	0.12569
2.75	1.05	2.92	0.00836	0.00834	0.00461	2.75	2.09	0.88	0.14707	0.13109	0.12640
3	1.05	2.92	0.00851	0.00849	0.00474	3	2.09	0.88	0.14777	0.13163	0.12703
Optimal						Optimal					
h_1	h_{21}	h_{22}	MSE	VAR	BIAS	h_1	h_{21}	h_{22}	MSE	VAR	BIAS
1	1.05	2.92	0.00728	0.00728	0.00138	1.75	2.09	0.88	0.14540	0.13034	0.12273

Table 3: Monte Carlo simulation results of plugging in error-laden W_2 to (11) to estimate $\rho(y,x)$ (Method (2) above)

Design 1 (linear)						Design 2 (nonlinear)					
h_1	h_{21}	h_{22}	MSE	VAR	abs(BIAS)	h_1	h_{21}	h_{22}	MSE	VAR	abs(BIAS)
0.5	1.05	2.92	831.49239	827.49615	1.99906	0.5	2.09	0.88	342390.30868	342380.03811	3.20477
1	1.05	2.92	15.36480	15.18324	0.42611	1	2.09	0.88	3268.04002	3264.78862	1.80316
1.5	1.05	2.92	1609.24991	1605.98439	1.80708	1.5	2.09	0.88	795.12130	795.09105	0.17391
2	1.05	2.92	981.89810	980.63434	1.12417	2	2.09	0.88	1113.49474	1113.43014	0.25417
2.5	1.05	2.92	2739.03247	2729.13653	3.14578	2.5	2.09	0.88	7439.31131	7399.31171	6.32452
3	1.05	2.92	48.28533	48.27034	0.12243	3	2.09	0.88	2589.00861	2583.05235	2.44054
3.5	1.05	2.92	227.69069	227.32253	0.60677	3.5	2.09	0.88	7367.37049	7363.00243	2.08999
4	1.05	2.92	326.60891	325.63719	0.98576	4	2.09	0.88	4810.95117	4805.88266	2.25133
4.5	1.05	2.92	29.37044	29.30256	0.26055	4.5	2.09	0.88	12005.60726	11951.43663	7.36007
5	1.05	2.92	7904.23502	7887.80380	4.05354	5	2.09	0.88	863.16882	862.91399	0.50480
5.5	1.05	2.92	16.49962	16.48066	0.13769	5.5	2.09	0.88	6884.05497	6883.55213	0.70911
6	1.05	2.92	54.17185	54.16807	0.06147	6	2.09	0.88	844.96901	843.85308	1.05637
Optimal						Optimal					
h_1	h_{21}	h_{22}	MSE	VAR	BIAS	h_1	h_{21}	h_{22}	MSE	VAR	BIAS
1	1.05	2.92	15.36480	15.18324	0.42611	1.5	2.09	0.88	795.12130	795.09105	0.17391

Table 4: Monte Carlo simulation results of using 2SLS and error-laden W_2 to estimate $\rho(y, x)$ (Method (1) above)

Design 1 (linear)			Design 2 (nonlinear)		
MSE	VAR	abs(BIAS)	MSE	VAR	abs(BIAS)
0.00267	0.00267	0.00190	1.11718	0.00147	1.05627

Table 5: Monte Carlo simulation results of my $\hat{\rho}(y, x)$ as a function of the sample size

Design 1 (linear)							Design 2 (nonlinear)						
N	h_1	h_{21}	h_{22}	MSE	VAR	abs(BIAS)	N	h_1	h_{21}	h_{22}	MSE	VAR	abs(BIAS)
500	1	1.05	2.92	0.00728	0.00728	0.00138	500	1.75	2.09	0.88	0.14540	0.13034	0.12273
1000	1	0.97	2.61	0.00354	0.00353	0.00359	1000	1.5	1.58	0.87	0.08341	0.06822	0.12323
5000	1.25	0.82	2.05	0.00102	0.00102	0.00026	5000	1.75	0.93	0.65	0.03053	0.03048	0.00708

Next, I study the performance of my estimator of the weighted LAR. For Design 1 (linear), I evaluate the WLAR at $x = 0$ and set the weighting function to be $\omega(0, w^*, \epsilon) = \omega_1$ if $\epsilon \in [q_{0,w^*}(0.25), q_{0,w^*}(0.35)]$ and $w^* \in [0.70, 0.90]$. The constant ω_1 is selected such that $\int_{0.7}^{0.9} \int_{q_{0,w^*}(0.25)}^{q_{0,w^*}(0.35)} \omega_1 f_{\epsilon,W^*|X=0}(\epsilon, w^*) d\epsilon dw^* = 1$. For Design 2 (nonlinear), I evaluate the WLAR at $x = 0.6$ and set the weighting function to be $\omega(0.6, w^*, \epsilon) = \omega_2$ if $\epsilon \in [q_{0.6,w^*}(0.25), q_{0.6,w^*}(0.35)]$ and $w^* \in [6, 6.23]$. The constant ω_2 is selected to ensure that $\int_{6}^{6.23} \int_{q_{0.6,w^*}(0.25)}^{q_{0.6,w^*}(0.35)} \omega_2 f_{\epsilon,W^*|X=0.6}(\epsilon, w^*) d\epsilon dw^* = 1$. These choices of the weighting functions yield a true value of the WLAR of 0.25 and 0.5032 for Design 1 and 2, respectively. In the estimation, I use the optimal values of h_1 coming from Table 2, and the same values of h_{21} and h_{22} as in Table 2. Table 6 below shows the results of my estimators of the WLAR under both designs. The performance of my estimator of the WLAR is comparable to the performance of my estimator for the point-wise structural derivative $\rho(y, x)$ (shown in Table 2). This shows evidence that my estimator could not only perform well at single points but also could maintain good performance in a global sense. The supports of the weighting functions I used here are not large. This is mainly due to computation constraints. One direction of future work is to look at the performance of my estimator for WLAR when the support of the weighting function is larger. This will give us a better understanding of the global performance of my estimator.

Table 6: Monte Carlo simulation results of my proposed estimator for the WLAR

	h_1	h_{21}	h_{22}	true value	MSE	VAR	abs(BIAS)
Design 1 (linear)	1	1.05	2.92	0.25	0.00779	0.00777	0.00405
Design 2 (nonlinear)	1.75	1.75	2.09	0.5032	0.15082	0.13829	0.11193

The Monte Carlo simulation exercise shown in this section is very limited. To have a comprehensive examination of the performance of my estimators, there are several future directions that I can work on. First, in my current exercise, the distribution of the measurement error ΔW_2 is set to be χ^2, which is ordinarily smooth. I can further explore the performance of my estimators under different distributions of the measurement error ΔW_2 and other variables. Second, when estimating the WLAR, I set the support for the weighting function to be small because of the computation constraint. With more time and more computation power, I can look at the performance of my estimator when the support of the weighting function in the WLAR is larger, i.e. the WLAR is taking the average marginal effect over a larger population. Third, the way I select h_1 is by minimizing the actual MSE, i.e., the mean squared difference between my estimator and the true value of the parameter. This is infeasible in practice when one deals with real-world data. It will be helpful if I could propose a feasible way to select the bandwidth.

6 Conclusion

I study the nonparametric identification and estimation of the nonseparable triangular equations model when the instrument variable W^* is mismeasured. I don't assume linearity or separability of the functions governing the relationship between observables and unobservables. To deal with the challenges caused by the co-existence of the measurement error and nonseparability, I first employ the deconvolution technique developed in the measurement error literature (Schennach (2004a),Schennach (2004b)) to identify the joint distribution of Y, X, W^* using two error-laden measurements of W^*. I then recover the structural derivative of the function of interest and the "Local Average Response" (LAR) via the "unobserved instrument" approach in Matzkin (2016). Based on the constructive identification results, I propose plug-in nonparametric estimators for these parameters and derive their uniform rates of convergence. I also conducted limited Monte Carlo exercises to show the finite sample performance of my estimators.

I recognize that there are some important future directions for this paper. First, in the main text, I only demonstrated the uniform rate of convergence of the estimators. The appendix contains proofs of their asymptotic normality, which rely on high-level assumptions. To further enhance my results, it would be beneficial to derive more primitive sufficient conditions for these assumptions. Second, as discussed in Section 5, to conduct a comprehensive examination of the performance of my estimators, additional Monte Carlo studies are necessary. It will also be extremely useful if I could propose a feasible way to select the bandwidth h_1. Last, it will be an interesting exercise to apply my proposed method to real-world data.

7 Appendix

7.1 Appendix A: Useful Lemmas

Before stating the lemmas, I define some notations.

Let Y be a random variable, Z be a random vector of dimension 2. Denote $Z_1 \equiv X, Z_2 \equiv W$. Let $\mathbb{S}_{(Y,Z)}$ be the support of (Y, Z), and define

$$\mathbb{S}_\tau \equiv \left\{ (y, z) \in \mathbb{S}_{(Y,Z)} : \left| \frac{\partial F_{X|W=w}(x)}{\partial w} \right| \geq \tau, \text{ and } f_{Y,X,W}(y, x, w) > \tau \right\}$$

and $\mathcal{S}_\tau \equiv \left\{ (\delta, z) \in [0, 1] \times \mathbb{S}_Z : \left| \frac{\partial F_{X|W=w}(x)}{\partial w} \right| \geq \tau \text{ and } f_{Y,X,W}(v(x, w, \delta), x, w) > \tau \right\}$. Let $\bar{\mathbb{M}}$ be a compact subset of the support of (Y, X, W). Define the mapping **invm** as $\mathbf{invm}(Y, X, W) \equiv (m^{-1}(Y, X), X, W)$, where m^{-1} is as defined in (2). Define the mapping **invs** as $\mathbf{invs}(Y, X, W) \equiv (s^{-1}(r(X, W), m^{-1}(X, Y)), X, W)$, where s^{-1} is the inverse of the s function defined in Lemma 1 with respect to its second argument. Let $\mathbb{M} \equiv \mathbf{invm}(\bar{\mathbb{M}})$, and $\mathcal{M} = \mathbf{invs}(\bar{\mathbb{M}})$. By Assumption 1 and 6, $\mathbb{M}$ and $\mathcal{M}$ are also compact, and there exist $\tau > 0$ such that $\bar{\mathbb{M}} \subset \mathbb{S}_\tau$ and $\mathcal{M} \subset \mathcal{S}_\tau$.

For any function $g : \mathbb{R}^3 \to \mathbb{R}$, define $\tilde{g}(z) \equiv \int g(y, z)dy$, $\tilde{g}_{Z_2}(z_2) \equiv \int \tilde{g}(z)dx$, $G(y, z) \equiv \int_{-\infty}^{y} \int_{-\infty}^{z} g(s, t)dtds$, and if $\tilde{g}(z) \neq 0$, define $G_{Y|Z=z}(y) = \left(\int_{-\infty}^{y} g(s, z)ds\right) / \tilde{g}(z)$; if $\tilde{g}_{Z_2}(z_2) \neq 0$, define $G_{Z_1|Z_2=z_2}(z_1) = \left(\int_{-\infty}^{z_1} \tilde{g}(s, z_2)ds\right) / \tilde{g}_{Z_2}(z_2)$. Let F denote a set of twice continuously differentiable functions $g : \mathbb{R}^3 \to \mathbb{R}$ such that g vanishes outside of $\mathbb{S}_{(Y,Z)}$. Let f denote the joint density of (Y, Z). Assume that f belongs to F. For any value $(y, z) \in \mathbb{S}_{Y,Z}$, and any value $\delta \in [0, 1]$, define the functionals $\Lambda(\cdot)$, $\alpha(\cdot)$, $\Phi_{(j)}(\cdot)$, $\Psi_{(j)}(\cdot)$ and $\tilde{\Lambda}(\cdot)$, $\tilde{\Psi}(\cdot)$ on F by $\Lambda(g) \equiv G_{Y|Z=z}(y)$, $\alpha(g) \equiv G^{-1}_{Y|Z=z}(\delta)$, $\Phi_{(j)}(g) \equiv \frac{\partial G^{-1}_{Y|Z=z}(\delta)}{\partial z_j}$, $\Psi_{(j)}(g) \equiv -\frac{\partial G_{Y|Z=z}(y)}{\partial z_j} \Big/ \frac{\partial G_{Y|Z=z}(y)}{\partial y}$ and $\tilde{\Lambda}(g) \equiv G_{X|W=w}(x)$, $\tilde{\Psi}(g) \equiv -\frac{\partial G_{X|W=w}(x)}{\partial w} \Big/ \frac{\partial G_{X|W=w}(x)}{\partial x}$. For simplicity, I leave the argument (z, y, δ) implicit.

Lemma 6.

$$\Phi_{(j)}(f) = \frac{\delta \frac{\partial \tilde{f}(z)}{\partial z_j} - \int_{-\infty}^{\alpha(f)} \frac{\partial f(y,z)}{\partial z_j} dy}{f(\alpha(f), z)}$$

$$\Psi_{(j)}(f) = \frac{\Lambda(f)\frac{\partial \tilde{f}(z)}{\partial z_j} - \int_{-\infty}^{y} \frac{\partial f(s,z)}{\partial z_j} ds}{f(y,z)}$$

$$\tilde{\Psi}(f) = \frac{\tilde{\Lambda}(f)\frac{\partial \tilde{f}_W(w)}{\partial w} - \int_{-\infty}^{x} \frac{\partial \tilde{f}(s,w)}{\partial w} ds}{\tilde{f}(x,w)}.$$

Proof. By definition,

$$\begin{aligned}
\delta &= F_{Y|Z=z}(\alpha(f)) \\
&= \frac{\int_{-\infty}^{\alpha(f)} f(y,z)dy}{\int_{-\infty}^{\infty} f(y,z)dy} \\
\implies \delta \tilde{f}(z) &= \int_{-\infty}^{\alpha(f)} f(y,z)dy
\end{aligned}$$

Taking derivatives on both sides with respect to z_j yields

$$\begin{aligned}
f(\alpha(f), z)\Phi_{(j)}(f) + \int_{-\infty}^{\alpha(f)} \frac{\partial f(y,z)}{\partial z_j} dy &= \delta \frac{\partial \tilde{f}(z)}{\partial z_j} \\
\implies \Phi_{(j)}(f) &= \frac{\delta \frac{\partial \tilde{f}(z)}{\partial z_j} - \int_{-\infty}^{\alpha(f)} \frac{\partial f(y,z)}{\partial z_j} dy}{f(\alpha(f), z)}.
\end{aligned}$$

Then note that

$$\frac{\partial G_{Y|Z=z}(x)}{\partial z_j} = \frac{\left(\int_{-\infty}^{y} \frac{\partial f(s,z)}{\partial z_j} ds\right) \tilde{f}(z) - \left(\int_{-\infty}^{y} f(s,z)ds\right) \frac{\partial \tilde{f}(z)}{\partial z_j}}{\tilde{f}^2(z)}$$

$$\frac{\partial G_{Y|Z=z}(x)}{\partial y} = \frac{f(y,z)}{\tilde{f}(z)}.$$

Then

$$\begin{aligned}
\Psi_{(j)}(f) &\equiv - \frac{\partial G_{Y|Z=z}(x)}{\partial z_j} \Bigg/ \frac{\partial G_{Y|Z=z}(x)}{\partial y} \\
&= \frac{F_{Y|Z=z}(y)\frac{\partial \tilde{f}(z)}{\partial z_j} - \int_{-\infty}^{y} \frac{\partial f(s,z)}{\partial z_j} ds}{f(y,z)}.
\end{aligned}$$

The conclusion for $\tilde{\Psi}(f)$ follows by the same argument. ■

Lemma 7. *For any h in F such that $\sup_{(y,z)\in \mathbb{S}_{(Y,Z)}} |h|$ is small enough, I have that*

$$\Lambda(f+h) - \Lambda(f) = D\Lambda(f;h) + R\Lambda(f;h)$$

$$\tilde{\Lambda}(f+h) - \tilde{\Lambda}(f) = D\tilde{\Lambda}(f;h) + R\tilde{\Lambda}(f;h),$$

where

$$\begin{aligned}
D\Lambda(f;h) &= \frac{\int_{-\infty}^{y} h(s,z)ds - \tilde{h}(z)F_{Y|Z=z}(y)}{\tilde{f}(z)} \\
R\Lambda(f;h) &= \left[\frac{\int_{-\infty}^{y} h(s,z)ds - \tilde{h}(z)F_{Y|Z=z}(y)}{\tilde{f}(z)}\right]\left[\frac{\tilde{h}(z)}{\tilde{f}(z)+\tilde{h}(z)}\right] \\
D\tilde{\Lambda}(f;h) &= \frac{\int_{-\infty}^{x} \tilde{h}(s,w)ds - \tilde{h}_W(w)F_{X_1|W=w}(x)}{\tilde{f}_W(w)} \\
R\tilde{\Lambda}(f;h) &= \left[\frac{\int_{-\infty}^{x} \tilde{h}(s,w)ds - \tilde{h}_W(w)F_{X_1|W=w}(x)}{\tilde{f}_W(w)}\right]\left[\frac{\tilde{h}_W(w)}{\tilde{f}_W(w)+\tilde{h}_W(w)}\right]
\end{aligned}$$

and for some $a_1 < \infty$,

$$\begin{aligned}
\sup_{(y,z)\in\mathbb{S}_\tau} |D\Lambda(f;h)| &\le a_1 \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}} |h| \\
\sup_{(y,z)\in\mathbb{S}_\tau} |R\Lambda(f;h)| &\le a_1 \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}} |h|^2;
\end{aligned}$$

for some $\tilde{a}_1 < \infty$,

$$\begin{aligned}
\sup_{(y,z)\in\mathbb{S}_\tau} \left|D\tilde{\Lambda}(f;h)\right| &\le \tilde{a}_1 \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}} |h| \\
\sup_{(y,z)\in\mathbb{S}_\tau} \left|R\tilde{\Lambda}(f;h)\right| &\le \tilde{a}_1 \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}} |h|^2.
\end{aligned}$$

Proof.

$$\begin{aligned}
\Lambda(f+h) - \Lambda(f) &= \frac{\int_{-\infty}^{y}(f(s,z)+h(s,z))ds}{\tilde{f}(z)+\tilde{h}(z)} - \frac{\int_{-\infty}^{y} f(s,z)ds}{\tilde{f}(z)} \\
&= \frac{\int_{-\infty}^{y} hds}{\tilde{f}+\tilde{h}} - \frac{\tilde{h}\int_{-\infty}^{y} fds}{\tilde{f}(\tilde{f}+\tilde{h})} \\
&= \frac{\int_{-\infty}^{y} hds - \tilde{h}F_{Y|Z=z}(y)}{\tilde{f}} + \left[\frac{\int_{-\infty}^{y} hds - \tilde{h}F_{Y|Z=z}(y)}{\tilde{f}}\right]\left[\frac{\tilde{h}}{\tilde{f}+\tilde{h}}\right]; \\
\tilde{\Lambda}(f+h) - \tilde{\Lambda}(f) &= \frac{\int_{-\infty}^{x} \tilde{h}ds - \tilde{h}_W F_{X_1|W=w}(x)}{\tilde{f}_W} + \left[\frac{\int_{-\infty}^{x} \tilde{h}ds - \tilde{h}_W F_{X_1|W=w}(x)}{\tilde{f}_W}\right]\left[\frac{\tilde{h}_W}{\tilde{f}_W+\tilde{h}_W}\right].
\end{aligned}$$

Define

$$
\begin{aligned}
D\Lambda(f;h) &= \frac{\int_{-\infty}^{y} hds - \tilde{h}F_{Y|Z=z}(y)}{\tilde{f}} \\
R\Lambda(f;h) &= \left[\frac{\int_{-\infty}^{y} hds - \tilde{h}F_{Y|Z=z}(y)}{\tilde{f}}\right]\left[\frac{\tilde{h}}{\tilde{f}+\tilde{h}}\right] \\
D\tilde{\Lambda}(f;h) &= \frac{\int_{-\infty}^{x} \tilde{h}ds - \tilde{h}_W F_{X_1|W=w}(x)}{\tilde{f}_W} \\
R\tilde{\Lambda}(f;h) &= \left[\frac{\int_{-\infty}^{x} \tilde{h}ds - \tilde{h}_W F_{X_1|W=w}(x)}{\tilde{f}_W}\right]\left[\frac{\tilde{h}_W}{\tilde{f}_W+\tilde{h}_W}\right].
\end{aligned}
$$

By the boundedness of $\mathbb{S}_{(Y,Z)}$, there exist finite constants C_X, C_Y such that for all $h \in \mathrm{F}$,

$$
\begin{aligned}
&\sup_{z\in\mathbb{S}_Z}\left|\tilde{h}(z)\right| \le \sup_{z\in\mathbb{S}_Z}\int |h(y,z)|\,dy \le C_Y \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}} |h| \\
&\sup_{w\in\mathbb{S}_W}\left|\tilde{h}_W(w)\right| = \sup_{w\in\mathbb{S}_W}\int \left|\tilde{h}(x,w)\right| dx \le C_X C_Y \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}} |h|
\end{aligned}
$$

By the definition of set $\mathbb{S}_\tau$, I have that $\inf_{(y,z)\in\mathbb{S}_\tau}\tilde{f}(z) > \tau$ and that $\inf_{(y,z)\in\mathbb{S}_\tau}\tilde{f}_W(w) = \inf_{(y,z)\in\mathbb{S}_\tau}\int\tilde{f}(s,w)ds > \tau C_X$ Let $\epsilon_0 \equiv \min\{\frac{\tau}{2C_Y}, 1\}$, For any h in F that $\sup_{(y,z)\in\mathbb{S}_{(Y,Z)}}|h| \le \epsilon_0$, I have that $\sup_{z\in\mathbb{S}_Z}\left|\tilde{h}(z)\right| \le \tau/2$, $\sup_{w\in\mathbb{S}_W}\left|\tilde{h}_W(w)\right| \le C_X\tau/2$, so that $\inf_{(y,z)\in\mathbb{S}_\tau}(\tilde{f}(z) + \tilde{h}(z)) \ge \tau/2$, and $\inf_{(y,z)\in\mathbb{S}_\tau}(\tilde{f}_W(w) + \tilde{h}_W(w)) \ge C_X\tau/2$. Then there exists a finite constant a_0 such that

$$
\begin{aligned}
&\sup_{(y,z)\in\mathbb{S}_\tau}|D\Lambda(f;h)| \le \frac{a_0}{\tau}\sup_{(y,z)\in\mathbb{S}_{(Y,Z)}}|h| \\
&\sup_{(y,z)\in\mathbb{S}_\tau}|R\Lambda(f;h)| \le \frac{a_0}{\tau}\sup_{(y,z)\in\mathbb{S}_{(Y,Z)}}|h| \times \frac{2\sup_{(y,z)\in\mathbb{S}_{(Y,Z)}}|h|}{\tau} = \frac{2a_0}{\tau^2}\sup_{(y,z)\in\mathbb{S}_{(Y,Z)}}|h|^2.
\end{aligned}
$$

Hence let $a_1 \equiv \max\{a_0/\tau, 2a_0/\tau^2\}$, I have that

$$
\begin{aligned}
&\sup_{(y,z)\in\mathbb{S}_\tau}|D\Lambda(f;h)| \le a_1\sup_{(y,z)\in\mathbb{S}_{(Y,Z)}}|h| \\
&\sup_{(y,z)\in\mathbb{S}_\tau}|R\Lambda(f;h)| \le a_1\sup_{(y,z)\in\mathbb{S}_{(Y,Z)}}|h|^2.
\end{aligned}
$$

Similarly, I can show that there exists a constant $\tilde{a}_1$ such that

$$\sup_{(y,z)\in\mathbb{S}_\tau}\left|D\tilde{\Lambda}(f;h)\right| \le \tilde{a}_1 \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}}|h|$$
$$\sup_{(y,z)\in\mathbb{S}_\tau}\left|R\tilde{\Lambda}(f;h)\right| \le \tilde{a}_1 \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}}|h|^2.$$

$\blacksquare$

Lemma 8. *For any h in* F *such that $\sup_{(y,z)\in\mathbb{S}_{(Y,Z)}}|h|$ is small enough, I have that*

$$\alpha(f+h)-\alpha(f) = D\alpha(f;h) + R\alpha(f;h),$$

where

$$D\alpha(f;h) \equiv \frac{\tilde{h}(z)\int^{\alpha(f)} f(y,z)dy - \tilde{f}(z)\int^{\alpha(f)} h(y,z)dy}{\tilde{f}(z)f(\alpha(f),z)}$$
$$R\alpha(f;h) \equiv -\left[\frac{\frac{\partial f(r_f',z)}{\partial y}\left(r_f-\alpha(f)\right)+h(r_h,z)}{f(r_f,z)+h\left(r_h,z\right)}\right]D\alpha(f;h)$$

for some r_f and r_f' between $\alpha(f+h)$ and $\alpha(f)$ defined in the proof. Moreover, for some $0<a<\infty$,

$$\sup_{(\delta,z)\in\mathcal{S}_\tau}|D\alpha(f;h)| \le a \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}}|h|$$
$$\sup_{(\delta,z)\in\mathcal{S}_\tau}|R\alpha(f;h)| \le a \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}}|h|^2.$$

Proof. First I show that for some $0<a_2<\infty$ and a_1 mentioned in Lemma 7,

$$\sup_{(\delta,z)\in\mathcal{S}_\tau}|\alpha(f+h)-\alpha(f)| \le \frac{2a_1a_2}{\tau}\sup_{(y,z)\in\mathbb{S}_{(Y,Z)}}|h|.$$

By the Mean Value Theorem, there exist r_1 between $\alpha(f+h)$ and $\alpha(f)$, such that

$$\begin{aligned}&(F+H)_{Y|Z=z}\left(\alpha(f+h)\right)-(F+H)_{Y|Z=z}\left(\alpha(f)\right)\\ &=(f+h)_{Y|Z=z}(r_1)\left(\alpha(f+h)-\alpha(f)\right)\end{aligned}$$

Hence since

$$\begin{aligned}
&(F+H)_{Y|Z=z}(\alpha(f+h)) \\
&= (F+H)_{Y|Z=z}\left((F+H)^{-1}_{Y|Z=z}(\delta)\right) = \delta = F_{Y|Z=z}\left(F^{-1}_{Y|Z=z}(\delta)\right)
\end{aligned}$$

it follows that

$$\alpha(f+h)-\alpha(f) = \frac{F_{Y|Z=z}\left(F^{-1}_{Y|Z=z}(\delta)\right) - (F+H)_{Y|Z=z}\left(F^{-1}_{Y|Z=z}(\delta)\right)}{(f+h)_{Y|Z=z}(r_1)}.$$

By the compactness of $\mathbb{S}_{Y,Z}$, for any h in F, there exist some finite constant a_2 such that $\sup_{z\in\mathbb{S}_Z}(\tilde{f}(z)+\tilde{h}(z)) \leq a_2$. By the definition of $\mathbb{S}_\tau$, $\inf_{(y,z)\in\mathbb{S}_\tau} f(y,z) > \tau$. By similar argument as in Lemma 7, for any h in F such that $\sup_{(y,z)\in\mathbb{S}_{(Y,Z)}}|h| \leq \min\{\tau/2, 1\}$, I have $\inf_{(y,z)\in\mathbb{S}_\tau}(f(y,z)+h(y,z)) > \tau/2$. Then $\inf_{(y,z)\in\mathbb{S}_\tau}(f+h)_{Y|Z=z}(r_1) = \frac{\inf_{(y,z)\in\mathbb{S}_\tau} f(r_1,z)+h(r_1,z)}{\sup_{(y,z)\in\mathbb{S}_\tau}\tilde{f}(z)+\tilde{h}(z)} > \frac{\tau}{2a_2}$.

By Lemma 7,

$$\sup_{(\delta,z)\in\mathcal{S}_\tau}\left|F_{Y|Z=z}\left(F^{-1}_{Y|Z=z}(\delta)\right) - (F+H)_{Y|Z=z}\left(F^{-1}_{Y|Z=z}(\delta)\right)\right| \leq a_1 \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}}|h|$$

for some $0 < a_1 < \infty$. Then I have that

$$\sup_{(\delta,z)\in\mathcal{S}_\tau}|\alpha(f+h)-\alpha(f)| \leq \frac{2a_1a_2}{\tau}\sup_{(y,z)\in\mathbb{S}_{(Y,Z)}}|h|. \tag{22}$$

Next, I will obtain a first-order expansion for $\alpha(f+h)$. By the fact that $(F+H)_{Y|Z=z}(\alpha(f+h)) = \delta = F_{Y|Z=z}(\alpha(f))$, I have

$$\begin{aligned}
&\frac{\int_{-\infty}^{\alpha(f)} f(y,z)dy}{\tilde{f}(z)} = \frac{\int^{\alpha(f+h)}(f(y,z)+h(y,z))dy}{\widetilde{f}(z)+\widetilde{h}(z)} \\
&\Longrightarrow (\tilde{f}(z)+\tilde{h}(z))\int^{\alpha(f)} f(y,z)dy = \tilde{f}(z)\int^{\alpha(f+h)}(f(y,z)+h(y,z))\,dy \\
&\Longrightarrow \tilde{h}(z)\int^{\alpha(f)} f(y,z)dy - \tilde{f}(z)\int^{\alpha(f)} h(y,z)dy \\
&\qquad = \tilde{f}(z)\int_{\alpha(f)}^{\alpha(f+h)} h(y,z)dy + \tilde{f}(z)\int_{\alpha(f)}^{\alpha(f+h)} f(y,z)dy
\end{aligned}$$

By the Mean Value Theorem, there exist r_f and r_h, between $\alpha(f)$ and $\alpha(f+h)$, such that

$$\int_{\alpha(f)}^{\alpha(f+h)} f(y,z)dy = f(r_f,z)\left(\alpha(f+h)-\alpha(f)\right) \text{ and}$$
$$\int_{\alpha(f)}^{\alpha(f+h)} h(y,z)dy = h(r_h,z)\left(\alpha(f+h)-\alpha(f)\right).$$

Denote $Az \equiv \tilde{h}(z)\int^{\alpha(f)} f(y,z)dy - \tilde{f}(z)\int^{\alpha(f)} h(y,z)dy$. Then

$$Az = \tilde{f}(z)\left[f(r_f,z)+h(r_h,z)\right]\left(\alpha(f+h)-\alpha(f)\right)$$
$$\Longrightarrow \alpha(f+h)-\alpha(f) = \frac{Az}{\tilde{f}(z)\left[f(r_f,z)+h(r_h,z)\right]}.$$

By the Mean Value Theorem, there exist r'_f between $\alpha(f)$ and r_f such that $f(r_f,z) - f(\alpha(f),z) = [\partial f(r'_f,z)/\partial y](r_f-\alpha(f))$. Hence

$$\begin{aligned}\alpha(f+h)-\alpha(f) &= \frac{Az}{\tilde{f}(z)\left(f(\alpha(f),z)+\frac{\partial f\left(r'_f,z\right)}{\partial y}\left(r_f-\alpha(f)\right)+h\left(r_h,z\right)\right)}\\ &= \frac{Az}{\tilde{f}(z)f(\alpha(f),z)} - \left[\frac{\frac{\partial f(r'_f,z)}{\partial y}\left(r_f-\alpha(f)\right)+h(r_h,z)}{f(\alpha(f),z)+\frac{\partial f\left(r'_f,z\right)}{\partial y}\left(r_f-\alpha(f)\right)+h\left(r_h,z\right)}\right]\frac{Az}{\tilde{f}(z)f(\alpha(f),z)}\end{aligned}$$

Denote

$$D\alpha(f;h) \equiv \frac{Az}{\tilde{f}(z)f(\alpha(f),z)} \text{ and}$$
$$R\alpha(f;h) \equiv -\left[\frac{\frac{\partial f(r'_f,z)}{\partial y}\left(r_f-\alpha(f)\right)+h(r_h,z)}{f(r_f,z)+h\left(r_h,z\right)}\right]\frac{Az}{\tilde{f}(z)f(\alpha(f),z)}.$$

Then

$$\alpha(f+h)-\alpha(f) = D\alpha(f;h)+R\alpha(f;h).$$

By the definition of r_f and by (22), $\sup_{(\delta,z)\in\mathcal{S}_\tau}|r_f-\alpha(f)| \leq \sup_{(\delta,z)\in\mathcal{S}_\tau}|\alpha(f+h)-\alpha(f)| \leq \frac{2a_1a_2}{\tau}\sup_{(y,z)\in\mathbb{S}_{(Y,Z)}}|h|$. It follows by the continuity of $\partial f/\partial y$ and the compactness of $\mathbb{S}_{(Y,Z)}$ that

$$\sup_{(\delta,z)\in\mathcal{S}_\tau}\left|\frac{\partial f(r'_f,z)}{\partial y}\left(r_f-\alpha(f)\right)+h(r_h,z)\right| \leq d_1 \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}}|h|$$

for some finite constant d_1. Then for all $h \in \mathrm{F}$ such that $\sup_{(y,z)\in\mathbb{S}_{(Y,Z)}} |h| \leq \min\{\tau/2, 1\}$, I have $\inf_{(\delta,z)\in\mathcal{S}_\tau} f(r_f, z) + h(r_h, z) > \tau/2$. Then there exists a finite constant a such that

$$\sup_{(\delta,z)\in\mathcal{S}_\tau} |D\alpha(f;h)| \leq a \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}} |h| \text{ and}$$
$$\sup_{(\delta,z)\in\mathcal{S}_\tau} |R\alpha(f;h)| \leq a \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}} |h|^2.$$

■

Lemma 9. *For any value $z \in \mathbb{S}_Z$, and any value $\delta \in (0,1)$, define functional $\Phi_{1,(j)}(\cdot)$ as $\Phi_{1,(j)}(g) \equiv \int_{-\infty}^{\alpha(g)} \frac{\partial g(y,z)}{\partial z_j} dy$. For simplicity, I leave the argument (z,δ) implicit. For any h in* F *that $\sup_{(y,z)\in\mathbb{S}_{(Y,Z)}} |h|$ is sufficiently small, I have that*

$$\Phi_{1,(j)}(f+h) - \Phi_{1,(j)}(f) = D\Phi_{1,(j)}(f;h) + R\Phi_{1,(j)}(f;h),$$

where

$$\begin{aligned}
D\Phi_{1,(j)}(f;h) &\equiv \frac{\partial f(\alpha(f),z)}{\partial z_j} D\alpha(f;h) + \int_{-\infty}^{\alpha(f)} \frac{\partial h(y,z)}{\partial z_j} dy \\
R\Phi_{1,(j)}(f;h) &\equiv \frac{\partial f(\alpha(f),z)}{\partial z_j} R\alpha(f;h) + \frac{\partial^2 f(\bar{r}'_f,z)}{\partial y \partial z_j} (\bar{r}_f - \alpha(f)) (\alpha(f+h) - \alpha(f)) \\
&\quad + \frac{\partial h(\bar{r}_h,z)}{\partial z_j} (\alpha(f+h) - \alpha(f)),
\end{aligned}$$

for some $\bar{r}_f$ and $\bar{r}_h$, and $\bar{r}'_f$ between $\alpha(f+h)$ and $\alpha(f)$ defined in the proof. Moreover, for some $0 < b_1 < \infty$,

$$\begin{aligned}
\sup_{(\delta,z)\in\mathcal{S}_\tau} \left|D\Phi_{1,(j)}(f;h)\right| &\leq b_1 \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}} |h| + b_1 \sup_{(y,z)\in\mathbb{S}_{Y,Z}} \left|\frac{\partial h(y,z)}{\partial z_j}\right| \\
\sup_{(\delta,z)\in\mathcal{S}_\tau} \left|R\Phi_{1,(j)}(f;h)\right| &\leq b_1 \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}} |h|^2 + b_1 \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}} |h| \sup_{(y,z)\in\mathbb{S}_{Y,Z}} \left|\frac{\partial h(y,z)}{\partial z_j}\right|.
\end{aligned}$$

Proof.

$$\begin{aligned}
\Phi_{1,(j)}(f+h) - \Phi_{1,(j)}(f) &= \int_{-\infty}^{\alpha(f+h)} \left(\frac{\partial f(y,z)}{\partial z_j} + \frac{\partial h(y,z)}{\partial z_j}\right) dy - \int_{-\infty}^{\alpha(f)} \frac{\partial f(y,z)}{\partial z_j} dy \\
&= \int_{-\infty}^{\alpha(f)} \frac{\partial h(y,z)}{\partial z_j} dy + \int_{\alpha(f)}^{\alpha(f+h)} \frac{\partial h(y,z)}{\partial z_j} dy + \int_{\alpha(f)}^{\alpha(f+h)} \frac{\partial f(y,z)}{\partial z_j} dy
\end{aligned}$$

By the Mean Value Theorem, there exist $\bar{r}_f$ and $\bar{r}_h$ between $\alpha(f)$ and $\alpha(f+h)$ such that

$$\int_{\alpha(f)}^{\alpha(f+h)} \frac{\partial f(y,z)}{\partial z_j} dy = \frac{\partial f(\bar{r}_f, z)}{\partial z_j} \left(\alpha(f+h) - \alpha(f)\right)$$
$$\int_{\alpha(f)}^{\alpha(f+h)} \frac{\partial h(y,z)}{\partial z_j} dy = \frac{\partial h(\bar{r}_h, z)}{\partial z_j} \left(\alpha(f+h) - \alpha(f)\right)$$

Apply the Mean Value Theorem again. I have that there exist $\bar{r}'_f$ between $\alpha(f)$ and $\alpha(f+h)$ such that

$$\int_{\alpha(f)}^{\alpha(f+h)} \frac{\partial f(y,z)}{\partial z_j} dy = \frac{\partial f(\alpha(f),z)}{\partial z_j} \left(\alpha(f+h) - \alpha(f)\right) + \frac{\partial^2 f(\bar{r}'_f, z)}{\partial y \partial z_j} \left(\bar{r}_f - \alpha(f)\right)\left(\alpha(f+h) - \alpha(f)\right).$$

Let

$$D\Phi_{1,(j)}(f;h) \equiv \frac{\partial f(\alpha(f),z)}{\partial z_j} D\alpha(f;h) + \int_{-\infty}^{\alpha(f)} \frac{\partial h(y,z)}{\partial z_j} dy$$
$$R\Phi_{1,(j)}(f;h) \equiv \frac{\partial f(\alpha(f),z)}{\partial z_j} R\alpha(f;h) + \frac{\partial^2 f(\bar{r}'_f, z)}{\partial y \partial z_j} \left(\bar{r}_f - \alpha(f)\right)\left(\alpha(f+h) - \alpha(f)\right) + \frac{\partial h(\bar{r}_h, z)}{\partial z_j} \left(\alpha(f+h) - \alpha(f)\right).$$

By the boundedness of $\mathbb{S}_{Y,Z}$, continuity of $\partial f/\partial z_j$, $\partial^2 f/\partial y \partial z_j$, $\partial h/\partial z_j$ and Lemma 8 there exist finite constants d_2, d_3, d_4, d_5 such that

$$\sup_{(\delta,z)\in\mathcal{S}_\tau} \left|D\Phi_{1,(j)}(f;h)\right| \leq d_2 \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}} |h| + d_3 \sup_{(y,z)\in\mathbb{S}_{Y,Z}} \left|\frac{\partial h(y,z)}{\partial z_j}\right|$$
$$\sup_{(\delta,z)\in\mathcal{S}_\tau} \left|R\Phi_{1,(j)}(f;h)\right| \leq d_4 \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}} |h|^2 + d_5 \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}} |h| \sup_{(y,z)\in\mathbb{S}_{Y,Z}} \left|\frac{\partial h(y,z)}{\partial z_j}\right|$$

Let $b_1 \equiv \max\{d_2, d_3, d_4, d_5\}$, then

$$\sup_{(\delta,z)\in\mathcal{S}_\tau} \left|D\Phi_{1,(j)}(f;h)\right| \leq b_1 \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}} |h| + b_1 \sup_{(y,z)\in\mathbb{S}_{Y,Z}} \left|\frac{\partial h(y,z)}{\partial z_j}\right|$$
$$\sup_{(\delta,z)\in\mathcal{S}_\tau} \left|R\Phi_{1,(j)}(f;h)\right| \leq b_1 \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}} |h|^2 + b_1 \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}} |h| \sup_{(y,z)\in\mathbb{S}_{Y,Z}} \left|\frac{\partial h(y,z)}{\partial z_j}\right|.$$

∎

Lemma 10. *Define functional* $\Phi_2()$ *on* F *by* $\Phi_2(g) = g(\alpha(g), z)$. *For any* h *in* F *that*

$\sup_{(y,z)\in\mathbb{S}_{(Y,Z)}} |h|$ *is small enough, I have that*

$$\Phi_2(f+h) - \Phi_2(f) = D\Phi_2(f;h) + R\Phi_2(f;h),$$

where

$$\begin{aligned}
D\Phi_2(f;h) &\equiv \frac{\partial f(\alpha(f),z)}{\partial y} D\alpha(f;h) + h(\alpha(f),z) \\
R\Phi_2(f;h) &\equiv \frac{\partial f(\alpha(f),z)}{\partial y} R\alpha(f;h) + \frac{\partial^2 f(\tilde{r}'_f,z)}{\partial y^2}\left(\tilde{r}_f - \alpha(f)\right)\left(\alpha(f+h)-\alpha(f)\right) + \\
&\quad \frac{\partial h(\tilde{r}_h,z)}{\partial y}\left(\alpha(f+h)-\alpha(f)\right).
\end{aligned}$$

for some $\tilde{r}_f$ and $\tilde{r}_h$, and $\tilde{r}'_f$ between $\alpha(f+h)$ and $\alpha(f)$ defined in the proof. Moreover, for some $0 < b_2 < \infty$,

$$\begin{aligned}
\sup_{(\delta,z)\in\mathcal{S}_\tau} |D\Phi_2(f;h)| &\le b_2 \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}} |h|. \\
\sup_{(\delta,z)\in\mathcal{S}_\tau} |R\Phi_2(f;h)| &\le b_2 \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}} |h|^2 + b_2 \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}} |h| \sup_{(y,z)\in\mathbb{S}_{Y,X}} \left|\frac{\partial h(y,z)}{\partial y}\right|.
\end{aligned}$$

Proof.

$$\begin{aligned}
&\Phi_2(f+h) - \Phi_2(f) \\
&= (f+h)\left(\alpha(f+h),z\right) - f\left(\alpha(f),z\right) \\
&= f\left(\alpha(f+h),z\right) - f\left(\alpha(f),z\right) + h\left(\alpha(f+h),z\right) - h\left(\alpha(f),z\right) + h\left(\alpha(f),z\right)
\end{aligned}$$

By the Mean Value Theorem, there exist $\tilde{r}_f$ and $\tilde{r}_h$ between $\alpha(f+h)$ and $\alpha(f)$, such that

$$\begin{aligned}
f\left(\alpha(f+h),z\right) - f\left(\alpha(f),z\right) &= \frac{\partial f(\tilde{r}_f,z)}{\partial y}\left(\alpha(f+h)-\alpha(f)\right) \\
h\left(\alpha(f+h),z\right) - h\left(\alpha(f),z\right) &= \frac{\partial h(\tilde{r}_h,z)}{\partial y}\left(\alpha(f+h)-\alpha(f)\right).
\end{aligned}$$

Apply the Mean Value Theorem again. There exist $\tilde{r}'_f$ between $\tilde{r}_f$ and $\alpha(f)$ such that

$$f\left(\alpha(f+h),z\right) - f\left(\alpha(f),z\right)$$

$$= \frac{\partial f(\alpha(f), z)}{\partial y} (\alpha(f+h) - \alpha(f)) + \frac{\partial^2 f(\tilde{r}'_f, z)}{\partial y^2} (\tilde{r}_f - \alpha(f)) (\alpha(f+h) - \alpha(f))$$

Then by Lemma 8 I have that

$$\begin{aligned}
&\Phi_2(f+h) - \Phi_2(f) \\
&= \frac{\partial f(\alpha(f), z)}{\partial y} (D\alpha(f;h) + R\alpha(f;h)) + \frac{\partial^2 f(\tilde{r}'_f, z)}{\partial y^2} (\tilde{r}_f - \alpha(f)) (\alpha(f+h) - \alpha(f)) + \\
&\qquad \frac{\partial h(\tilde{r}_h, z)}{\partial y} (\alpha(f+h) - \alpha(f)) + h(\alpha(f), z),
\end{aligned}$$

where for some $0 < a < \infty$ and $\sup_{(y,z) \in \mathbb{S}_{(Y,Z)}} |h| \leq \epsilon_0 \equiv \min\{\tau/2, 1\}$,

$$\begin{aligned}
\sup_{(\delta,z) \in \mathcal{S}_\tau} |D\alpha(f;h)| &\leq a \sup_{(y,z) \in \mathbb{S}_{(Y,Z)}} |h| \\
\sup_{(\delta,z) \in \mathcal{S}_\tau} |R\alpha(f;h)| &\leq a \sup_{(y,z) \in \mathbb{S}_{(Y,Z)}} |h|^2.
\end{aligned}$$

Define

$$\begin{aligned}
D\Phi_2(f;h) &\equiv \frac{\partial f(\alpha(f), z)}{\partial y} D\alpha(f;h) + h(\alpha(f), z) \\
R\Phi_2(f;h) &\equiv \frac{\partial f(\alpha(f), z)}{\partial y} R\alpha(f;h) + \frac{\partial^2 f(\tilde{r}'_f, z)}{\partial y^2} (\tilde{r}_f - \alpha(f)) (\alpha(f+h) - \alpha(f)) + \\
&\qquad \frac{\partial h(\tilde{r}_h, z)}{\partial y} (\alpha(f+h) - \alpha(f)).
\end{aligned}$$

Then

$$\Phi_2(f+h) - \Phi_2(f) = D\Phi_2(f;h) + R\Phi_2(f;h).$$

By the compactness of $\mathbb{S}_{(Y,X)}$ and continuity of $\partial f / \partial y$, $\partial^2 f / \partial y^2$, and $\partial h / \partial y$, I have that there exist some finite constant d_6 such that

$$\sup_{(\delta,z) \in \mathcal{S}_\tau} |D\Phi_2(f;h)| \leq d_6 \sup_{(y,z) \in \mathbb{S}_{(Y,Z)}} |h|.$$

holds for all $h \in \mathrm{F}$. On the other hand, by Lemma 8, there exist some finite constant a_1, a_2 such that $\sup_{(\delta,z) \in \mathcal{S}_\tau} |\tilde{r}_f - \alpha(f)| \leq \sup_{(\delta,z) \in \mathcal{S}_\tau} |\alpha(f+h) - \alpha(f)| \leq \frac{2a_1 a_2}{\tau} \sup_{(y,z) \in \mathbb{S}_{(Y,Z)}} |h|$.

Then there exist some finite constants d_7, d_8 such that

$$\sup_{(\delta,z)\in\mathcal{S}_\tau} |R\Phi_2(f;h)| \leq d_7 \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}} |h|^2 + d_8 \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}} |h| \sup_{(y,z)\in\mathbb{S}_{Y,Z}} \left|\frac{\partial h(y,z)}{\partial y}\right|$$

Let $b_2 \equiv \max\{d_6, d_7, d_8\}$. Then

$$\sup_{(\delta,z)\in\mathcal{S}_\tau} |D\Phi_2(f;h)| \leq b_2 \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}} |h|.$$
$$\sup_{(\delta,z)\in\mathcal{S}_\tau} |R\Phi_2(f;h)| \leq b_2 \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}} |h|^2 + b_2 \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}} |h| \sup_{(y,z)\in\mathbb{S}_{Y,Z}} \left|\frac{\partial h(y,z)}{\partial y}\right|.$$

■

Lemma 11. *For any value $z \in \mathbb{S}_Z$, and any value $\delta \in (0,1)$, define functionals $\Psi_{1,(j)}(\cdot)$ and $\tilde{\Psi}_1(\cdot)$ as $\Psi_{1,(j)}(g) \equiv \int_{-\infty}^{y} \frac{\partial g(s,z)}{\partial z_j} ds$ and $\tilde{\Psi}_1(g) \equiv \int_{-\infty}^{x} \frac{\partial \tilde{g}(s,w)}{\partial w} ds$. For simplicity, I leave the argument z implicit. For any h in F that $\sup_{(y,z)\in\mathbb{S}_{(Y,Z)}} |h|$ is sufficiently small, I have that*

$$\Phi_{(j)}(f+h) - \Phi_{(j)}(f) = D\Phi_{(j)}(f;h) + R\Phi_{(j)}(f;h),$$
$$\Psi_{(j)}(f+h) - \Psi_{(j)}(f) = D\Psi_{(j)}(f;h) + R\Psi_{(j)}(f;h),$$
$$\tilde{\Psi}(f+h) - \tilde{\Psi}(f) = D\tilde{\Psi}(f;h) + R\tilde{\Psi}(f;h),$$

where

$$D\Phi_{(j)}(f;h) \equiv \frac{\left(\delta\frac{\partial \tilde{h}(z)}{\partial z_j} - D\Phi_{1,(j)}(f;h)\right)\Phi_2(f) - \left(\delta\frac{\partial \tilde{f}(z)}{\partial z_j} - \Phi_{1,(j)}(f)\right) D\Phi_2(f;h)}{\Phi_2^2(f)}$$
$$R\Phi_{(j)}(f;h) \equiv -\frac{\Phi_2(f)R\Phi_{1,(j)}(f;h) + \left(\delta\frac{\partial \tilde{f}(z)}{\partial z_j} - \Phi_{1,(j)}(f)\right) R\Phi_2(f;h)}{\Phi_2^2(f)}$$
$$- \frac{\Delta\Phi_2(f;h)\left[\left(\delta\frac{\partial \tilde{h}(z)}{\partial z_j} - \Delta\Phi_{1,(j)}(f;h)\right)\Phi_2(f) - \left(\delta\frac{\partial \tilde{f}(z)}{\partial z_j} - \Phi_{1,(j)}(f)\right)\Delta\Phi_2(f;h)\right]}{\Phi_2(f+h)\Phi_2^2(f)}$$
$$D\Psi_{(j)}(f;h) \equiv \frac{\left(D\Lambda(f;h)\frac{\partial \tilde{f}(z)}{\partial z_j} + \Lambda(f)\frac{\partial \tilde{h}(z)}{\partial z_j} - \int_{-\infty}^{y}\frac{\partial h(s,z)}{\partial z_j}ds\right)\times f(y,z) - \left(\Lambda(f)\frac{\partial \tilde{f}(z)}{\partial z_j} - \Psi_{1,(j)}(f)\right)\times h(y,z)}{f^2(y,z)}$$
$$R\Psi_{(j)}(f;h) \equiv \frac{\left(R\Lambda(f;h)\frac{\partial \tilde{f}(z)}{\partial z_j} + \frac{\partial \tilde{h}(z)}{\partial z_j}\left(D\Lambda(f;h) + R\Lambda(f;h)\right)\right)}{f(y,z)}$$
$$- h(y,z) \times \left[\frac{\Lambda(f)\frac{\partial \tilde{h}(z)}{\partial z_j} f(y,z) + \Delta\Lambda(f;h)\left(\frac{\partial \tilde{f}(z)}{\partial z_j} + \frac{\partial \tilde{h}(z)}{\partial z_j}\right) f(y,z)}{\left(f(y,z)+h(y,z)\right) f^2(y,z)}\right.$$
$$\left. - \frac{\Lambda(f)\frac{\partial \tilde{f}(z)}{\partial z_j} h(y,z) + \Delta\Psi_{1,(j)}(f;h) f(y,z) - \Psi_{1,(j)}(f;h) h(y,z)}{\left(f(y,z)+h(y,z)\right) f^2(y,z)}\right]$$

$$D\tilde{\Psi}(f;h) \equiv \frac{\left(D\tilde{\Lambda}(f;h)\frac{\partial \tilde{f}_W(w)}{\partial w} + \tilde{\Lambda}(f)\frac{\partial \tilde{h}_W(w)}{\partial w} - \int_{-\infty}^{x}\frac{\partial \tilde{h}(s,w)}{\partial w}ds\right) \times \tilde{f}(x,w) - \left(\tilde{\Lambda}(f)\frac{\partial \tilde{f}_W(w)}{\partial w} - \tilde{\Psi}_1(f)\right) \times \tilde{h}(x,w)}{\tilde{f}^2(x,w)}$$

$$\begin{aligned} R\tilde{\Psi}(f;h) \equiv & \frac{\left(R\tilde{\Lambda}(f;h)\frac{\partial \tilde{f}_W(w)}{\partial w} + \frac{\partial \tilde{h}_W(w)}{\partial w}\left(D\tilde{\Lambda}(f;h) + R\tilde{\Lambda}(f;h)\right)\right)}{\tilde{f}(x,w)} \\ & - \tilde{h}(x,w) \times \left[\frac{\tilde{\Lambda}(f)\frac{\partial \tilde{h}_W(w)}{\partial w}\tilde{f}(x,w) + \Delta\tilde{\Lambda}(f;h)\left(\frac{\partial \tilde{f}_W(w)}{\partial w} + \frac{\partial \tilde{h}_W(w)}{\partial w}\right)\tilde{f}(x,w)}{\left(\tilde{f}(x,w) + \tilde{h}(x,w)\right)\tilde{f}^2(x,w)}\right. \\ & \left. - \frac{\tilde{\Lambda}(f)\frac{\partial \tilde{f}_W(w)}{\partial w}\tilde{h}(x,w) + \Delta\tilde{\Psi}_{1,(j)}(f;h)\tilde{f}(x,w) - \tilde{\Psi}_1(f;h)\tilde{h}(x,w)}{\left(\tilde{f}(x,w) + \tilde{h}(x,w)\right)\tilde{f}^2(x,w)}\right], \end{aligned}$$

with $\Delta\Phi_l(f;h) \equiv D\Phi_l(f;h) + R\Phi_l(f;h)$ *for* $l = 1,(j)$ *and* $l = 2$, $\Delta\Psi_{1,(j)}(f;h) \equiv D\Psi_{1,(j)}(f;h) + R\Psi_{1,(j)}(f;h)$ *and* $\Delta\tilde{\Psi}_1(f;h) \equiv D\tilde{\Psi}_1(f;h) + R\tilde{\Psi}_1(f;h)$. *Moreover, for some* $c_1, c_2, c_3 < \infty$,

$$\begin{aligned} \sup_{(\delta,z)\in\mathcal{S}_\tau}\left|D\Phi_{(j)}(f;h)\right| &\le c_1 \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}}|h| + c_1 \sup_{(y,z)\in\mathbb{S}_{Y,Z}}\left|\frac{\partial h(y,z)}{\partial z_j}\right| \\ \sup_{(\delta,z)\in\mathcal{S}_\tau}\left|R\Phi_{(j)}(f;h)\right| &\le c_1 \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}}|h|^2 + c_1 \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}}|h| \sup_{(y,z)\in\mathbb{S}_{Y,Z}}\left|\frac{\partial h(y,z)}{\partial y}\right| \\ &\quad + c_1 \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}}|h| \sup_{(y,z)\in\mathbb{S}_{Y,Z}}\left|\frac{\partial h(y,z)}{\partial z_j}\right| \\ \sup_{(y,z)\in\mathbb{S}_\tau}\left|D\Psi_{(j)}(f;h)\right| &\le c_2 \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}}|h| + c_2 \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}}\left|\frac{\partial h(y,z)}{\partial z_j}\right| \\ \sup_{(y,z)\in\mathbb{S}_\tau}\left|R\Psi_{(j)}(f;h)\right| &\le c_2 \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}}|h|^2 + c_2 \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}}|h| \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}}\left|\frac{\partial h(y,z)}{\partial z_j}\right| \\ \sup_{(y,z)\in\mathbb{S}_\tau}\left|D\tilde{\Psi}(f;h)\right| &\le c_3 \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}}|h| + c_3 \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}}\left|\frac{\partial h(y,z)}{\partial z_{j+1}}\right| \\ \sup_{(y,z)\in\mathbb{S}_\tau}\left|R\tilde{\Psi}(f;h)\right| &\le c_3 \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}}|h|^2 + c_3 \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}}|h| \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}}\left|\frac{\partial h(y,z)}{\partial z_{j+1}}\right|. \end{aligned}$$

Proof. By Lemma 6,

$$\begin{aligned} \Phi_{(j)}(f) &= \frac{\delta\frac{\partial \tilde{f}(z)}{\partial z_j} - \Phi_{1,(j)}(f)}{\Phi_2(f)} \\ \Psi_{(j)}(f) &= \frac{\Lambda(f)\frac{\partial \tilde{f}(z)}{\partial z_j} - \Psi_{1,(j)}(f)}{f(y,z)} \text{ and} \\ \tilde{\Psi}(f) &= \frac{\tilde{\Lambda}(f)\frac{\partial \tilde{f}_W(w)}{\partial w} - \tilde{\Psi}_1(f)}{\tilde{f}(x,w)} \end{aligned}$$

For all $h \in \mathrm{F}$,

$$\begin{aligned} \Phi_{(j)}(f+h) - \Phi_{(j)}(f) &= \frac{\delta\frac{\partial \tilde{f}(z)}{\partial z_j} + \delta\frac{\partial \tilde{h}(z)}{\partial z_j} - \Phi_{1,(j)}(f+h)}{\Phi_2(f+h)} - \frac{\delta\frac{\partial \tilde{f}(z)}{\partial z_j} - \Phi_{1,(j)}(f)}{\Phi_2(f)} \\ &\equiv \frac{N_1'}{D_1'} - \frac{N_1}{D_1} \end{aligned}$$

$$\Psi_{(j)}(f+h) - \Psi_{(j)}(f) = \frac{\Lambda(f+h)\left(\frac{\partial \tilde{f}(z)}{\partial z_j} + \frac{\partial \tilde{h}(z)}{\partial z_j}\right) - \Psi_{1,(j)}(f+h)}{f(y,z)+h(y,z)} - \frac{\Lambda(f)\frac{\partial \tilde{f}(z)}{\partial z_j} - \Psi_{1,(j)}(f)}{f(y,z)}$$
$$\equiv \frac{N_2'}{D_2'} - \frac{N_2}{D_2}$$
$$\tilde{\Psi}(f+h) - \tilde{\Psi}(f) = \frac{\tilde{\Lambda}(f+h)\left(\tilde{f}_W(w) + \tilde{h}_W(w)\right) - \tilde{\Psi}_1(f+h)}{\tilde{f}(x,w)+\tilde{h}(x,w)} - \frac{\tilde{\Lambda}(f)\tilde{f}_W(w) - \tilde{\Psi}_1(f)}{\tilde{f}(x,w)}$$
$$\equiv \frac{N_3'}{D_3'} - \frac{N_3}{D_3},$$

where I denote

$$N_1' = \delta\frac{\partial \tilde{f}(z)}{\partial z_j} + \delta\frac{\partial \tilde{h}(z)}{\partial z_j} - \Phi_{1,(j)}(f+h)$$
$$N_1 = \delta\frac{\partial \tilde{f}(z)}{\partial z_j} - \Phi_{1,(j)}(f)$$
$$N_2' = \Lambda(f+h)\left(\frac{\partial \tilde{f}(z)}{\partial z_j} + \frac{\partial \tilde{h}(z)}{\partial z_j}\right) - \Psi_{1,(j)}(f+h)$$
$$N_2 = \Lambda(f)\frac{\partial \tilde{f}(z)}{\partial z_j} - \Psi_{1,(j)}(f)$$
$$D_1' = \Phi_2(f+h)$$
$$D_1 = \Phi_2(f)$$
$$D_2' = f(y,z)+h(y,z)$$
$$D_2 = f(y,z).$$

Note that

$$N_1' - N_1 = \delta\frac{\partial \tilde{h}(z)}{\partial z_j} - D\Phi_{1,(j)}(f;h) - R\Phi_{1,(j)}(f;h) \equiv DN_1 + RN_1$$
$$N_2' - N_2 = (\Lambda(f+h) - \Lambda(f))\left(\frac{\partial \tilde{f}(z)}{\partial z_j} + \frac{\partial \tilde{h}(z)}{\partial z_j}\right) + \Lambda(f)\frac{\partial \tilde{h}(z)}{\partial z_j} - \int_{-\infty}^{y} \frac{\partial h(s,z)}{\partial z_j} ds$$
$$= (D\Lambda(f;h) + R\Lambda(f;h))\left(\frac{\partial \tilde{f}(z)}{\partial z_j} + \frac{\partial \tilde{h}(z)}{\partial z_j}\right) + \Lambda(f)\frac{\partial \tilde{h}(z)}{\partial z_j} - \int_{-\infty}^{y} \frac{\partial h(s,z)}{\partial z_j} ds$$
$$\equiv DN_2 + RN_2$$
$$D_1' - D_1 = D\Phi_2(f;h) + R\Phi_2(f;h)$$
$$D_2' - D_2 = h(y,z),$$

where

$$
\begin{aligned}
DN_1 &\equiv \delta \frac{\partial \tilde{h}(z)}{\partial z_j} - D\Phi_{1,(j)}(f;h) \\
RN_1 &\equiv -R\Phi_{1,(j)}(f;h) \\
DN_2 &\equiv D\Lambda(f;h)\frac{\partial \tilde{f}(z)}{\partial z_j} + \Lambda(f)\frac{\partial \tilde{h}(z)}{\partial z_j} - \int_{-\infty}^{y} \frac{\partial h(s,z)}{\partial z_j} ds \\
RN_2 &\equiv R\Lambda(f;h)\frac{\partial \tilde{f}(z)}{\partial z_j} + \frac{\partial \tilde{h}(z)}{\partial z_j}\left(D\Lambda(f;h) + R\Lambda(f;h)\right).
\end{aligned}
$$

I will make use of the equation:

$$
\begin{aligned}
\frac{N_j'}{D_j'} - \frac{N_j}{D_j} &= \frac{N_j' D_j - N_j D_j'}{D_j^2} - \frac{(D_j' - D_j)(N_j' D_j - N_j D_j')}{D_j' D_j^2} \\
&= \frac{(N_j' - RN_j)D_j - N_j\left(D_j' - RD_j\right)}{D_j^2} \\
&\quad + \frac{D_j \times RN_j - N_j \times RD_j}{D_j^2} - \frac{(D_j' - D_j)(N_j' D_j - N_j D_j')}{D_j' D_j^2}.
\end{aligned}
$$

Then

$$
\begin{aligned}
&\Phi_{(j)}(f+h) - \Phi_{(j)}(f) \\
&= \frac{\left(\delta \frac{\partial \tilde{h}(z)}{\partial z_j} - D\Phi_{1,(j)}(f;h)\right)\Phi_2(f) - \left(\delta \frac{\partial \tilde{f}(z)}{\partial z_j} - \Phi_{1,(j)}(f)\right) D\Phi_2(f;h)}{\Phi_2^2(f)} \\
&\quad - \frac{\Phi_2(f) R\Phi_{1,(j)}(f;h) + \left(\delta \frac{\partial \tilde{f}(z)}{\partial z_j} - \Phi_{1,(j)}(f)\right) R\Phi_2(f;h)}{\Phi_2^2(f)} \\
&\quad - \frac{(D\Phi_2(f;h) + R\Phi_2(f;h))\left(\left(\delta \frac{\partial \tilde{f}(z)}{\partial z_j} + \delta \frac{\partial \tilde{h}(z)}{\partial z_j} - \Phi_{1,(j)}(f+h)\right)\Phi_2(f) - \left(\delta \frac{\partial \tilde{f}(z)}{\partial z_j} - \Phi_{1,(j)}(f)\right)\Phi_2(f+h)\right)}{\Phi_2(f+h)\Phi_2^2(f)}.
\end{aligned}
$$

$$
\begin{aligned}
&\Psi_{(j)}(f+h) - \Psi_{(j)}(f) \\
&= \frac{DN_2 \times D_2 - N_2 \times h(y,z)}{D_2^2} + \frac{D_2 \times RN_2}{D_2^2} - \frac{h(y,z)(N_2' D_2 - N_2 D_2')}{D_2' D_2^2} \\
&= \frac{\left(D\Lambda(f;h)\frac{\partial \tilde{f}(z)}{\partial z_j} + \Lambda(f)\frac{\partial \tilde{h}(z)}{\partial z_j} - \int_{-\infty}^{y} \frac{\partial h(s,z)}{\partial z_j} ds\right) \times f(y,z) - \left(\Lambda(f)\frac{\partial \tilde{f}(z)}{\partial z_j} - \Psi_{1,(j)}(f)\right) \times h(y,z)}{f^2(y,z)} \\
&\quad + \frac{\left(R\Lambda(f;h)\frac{\partial \tilde{f}(z)}{\partial z_j} + \frac{\partial \tilde{h}(z)}{\partial z_j}\left(D\Lambda(f;h) + R\Lambda(f;h)\right)\right)}{f(y,z)} \\
&\quad - \frac{h(y,z)\left(\left(\Lambda(f+h)\left(\frac{\partial \tilde{f}(z)}{\partial z_j} + \frac{\partial \tilde{h}(z)}{\partial z_j}\right) - \Psi_{1,(j)}(f+h)\right) f(y,z) - \left(\Lambda(f)\frac{\partial \tilde{f}(z)}{\partial z_j} - \Psi_{1,(j)}(f)\right)(f(y,z) + h(y,z))\right)}{(f(y,z) + h(y,z))\, f^2(y,z)}.
\end{aligned}
$$

Denote $D\Phi_{(j)}(f;h)$, $D\Psi_{(j)}(f;h)$, $R\Phi_{(j)}(f;h)$, and $R\Psi_{(j)}(f;h)$ by

$$
\begin{aligned}
D\Phi_{(j)}(f;h) &\equiv \frac{\left(\delta\frac{\partial\tilde{h}(z)}{\partial z_j} - D\Phi_{1,(j)}(f;h)\right)\Phi_2(f) - \left(\delta\frac{\partial\tilde{f}(z)}{\partial z_j} - \Phi_{1,(j)}(f)\right)D\Phi_2(f;h)}{\Phi_2^2(f)} \\
R\Phi_{(j)}(f;h) &\equiv -\frac{\Phi_2(f)R\Phi_{1,(j)}(f;h) + \left(\delta\frac{\partial\tilde{f}(z)}{\partial z_j} - \Phi_{1,(j)}(f)\right)R\Phi_2(f;h)}{\Phi_2^2(f)} \\
&\quad - \frac{(D\Phi_2(f;h)+R\Phi_2(f;h))\left(\left(\delta\frac{\partial\tilde{f}(z)}{\partial z_j} + \delta\frac{\partial\tilde{h}(z)}{\partial z_j} - \Phi_{1,(j)}(f+h)\right)\Phi_2(f) - \left(\delta\frac{\partial\tilde{f}(z)}{\partial z_j} - \Phi_{1,(j)}(f)\right)\Phi_2(f+h)\right)}{\Phi_2(f+h)\Phi_2^2(f)} \\
&= -\frac{\Phi_2(f)R\Phi_{1,(j)}(f;h) + \left(\delta\frac{\partial\tilde{f}(z)}{\partial z_j} - \Phi_{1,(j)}(f)\right)R\Phi_2(f;h)}{\Phi_2^2(f)} \\
&\quad - \frac{\Delta\Phi_2(f;h)\left[\left(\delta\frac{\partial\tilde{h}(z)}{\partial z_j} - \Delta\Phi_{1,(j)}(f;h)\right)\Phi_2(f) - \left(\delta\frac{\partial\tilde{f}(z)}{\partial z_j} - \Phi_{1,(j)}(f)\right)\Delta\Phi_2(f;h)\right]}{\Phi_2(f+h)\Phi_2^2(f)} \\
D\Psi_{(j)}(f;h) &\equiv \frac{\left(D\Lambda(f;h)\frac{\partial\tilde{f}(z)}{\partial z_j} + \Lambda(f)\frac{\partial\tilde{h}(z)}{\partial z_j} - \int_{-\infty}^{y}\frac{\partial h(s,z)}{\partial z_j}ds\right)\times f(y,z) - \left(\Lambda(f)\frac{\partial\tilde{f}(z)}{\partial z_j} - \Psi_{1,(j)}(f)\right)\times h(y,z)}{f^2(y,z)} \\
R\Psi_{(j)}(f;h) &\equiv \frac{\left(R\Lambda(f;h)\frac{\partial\tilde{f}(z)}{\partial z_j} + \frac{\partial\tilde{h}(z)}{\partial z_j}\left(D\Lambda(f;h)+R\Lambda(f;h)\right)\right)}{f(y,z)} \\
&\quad - \frac{h(y,z)\left(\left(\Lambda(f+h)\left(\frac{\partial\tilde{f}(z)}{\partial z_j} + \frac{\partial\tilde{h}(z)}{\partial z_j}\right) - \Psi_{1,(j)}(f+h)\right)f(y,z) - \left(\Lambda(f)\frac{\partial\tilde{f}(z)}{\partial z_j} - \Psi_{1,(j)}(f)\right)(f(y,z)+h(y,z))\right)}{(f(y,z)+h(y,z))\,f^2(y,z)} \\
&= \frac{\left(R\Lambda(f;h)\frac{\partial\tilde{f}(z)}{\partial z_j} + \frac{\partial\tilde{h}(z)}{\partial z_j}\left(D\Lambda(f;h)+R\Lambda(f;h)\right)\right)}{f(y,z)} \\
&\quad - h(y,z)\times\left[\frac{\Lambda(f)\frac{\partial\tilde{h}(z)}{\partial z_j}f(y,z) + \Delta\Lambda(f;h)\left(\frac{\partial\tilde{f}(z)}{\partial z_j} + \frac{\partial\tilde{h}(z)}{\partial z_j}\right)f(y,z)}{(f(y,z)+h(y,z))\,f^2(y,z)}\right. \\
&\qquad\left. - \frac{\Lambda(f)\frac{\partial\tilde{f}(z)}{\partial z_j}h(y,z) + \Delta\Psi_{1,(j)}(f;h)f(y,z) - \Psi_{1,(j)}(f;h)h(y,z)}{(f(y,z)+h(y,z))\,f^2(y,z)}\right],
\end{aligned}
$$

where $\Delta\Phi_l(f;h) \equiv D\Phi_l(f;h) + R\Phi_l(f;h)$ for $l = 1,(j)$ and $l = 2$, and $\Delta\Psi_{1,(j)}(f;h) \equiv D\Psi_{1,(j)}(f;h) + R\Psi_{1,(j)}(f;h)$. Then

$$
\begin{aligned}
\Phi_{(j)}(f+h) - \Phi_{(j)}(f) &= D\Phi_{(j)}(f;h) + R\Phi_{(j)}(f;h) \\
\Psi_{(j)}(f+h) - \Psi_{(j)}(f) &= D\Psi_{(j)}(f;h) + R\Psi_{(j)}(f;h).
\end{aligned}
$$

By the same logic, I can write

$$
\tilde{\Psi}(f+h) - \tilde{\Psi}(f) = D\tilde{\Psi}(f;h) + R\tilde{\Psi}(f;h),
$$

where $D\tilde{\Psi}(f;h)$ and $R\tilde{\Psi}(f;h)$ are as defined in the Lemma.

By compactness of $\mathbb{S}_{(Y,Z)}$ and continuity of f, $\partial f/\partial z$, I have that there exist finite constant

d_9 such that

$$\sup_{(\delta,z)\in[0,1]\times\mathbb{S}_\tau} \left|\Phi_{1,(j)}(f)\right| = \sup_{(\delta,z)\in[0,1]\times\mathbb{S}_\tau} \left|\int_{-\infty}^{\alpha(f)} \frac{\partial f(y,z)}{\partial z_j} dy\right| \leq d_9 \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}} \left|\frac{\partial f(y,z)}{\partial z_j}\right| < \infty$$

$$\sup_{(\delta,z)\in[0,1]\times\mathbb{S}_\tau} |\Phi_2(f)| \leq \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}} |f(y,z)| < \infty \tag{23}$$

Also, by the definition of $\mathbb{S}_\tau$,

$$\inf_{(\delta,z)\in\mathcal{S}_\tau} \Phi_2^2(f) \geq \inf_{(y,z)\in\mathbb{S}_\tau} f^2(y,z) > \tau^2. \tag{24}$$

For $\sup_{(y,z)\in\mathbb{S}_{(Y,Z)}} |h| \leq \min\{\tau/2, 1\}$, I have

$$\begin{aligned}\inf_{(y,z)\in\mathbb{S}_\tau} \Phi_2(f+h) &= \inf_{(y,z)\in\mathbb{S}_\tau} f(\alpha(f+h),z) + h(\alpha(f+h),z) \\ &> \tau - \frac{\tau}{2} = \frac{\tau}{2}.\end{aligned} \tag{25}$$

By Lemma 9, 10 there exist finite constants $d_{10}, d_{11}, d_{12}, d_{13}, d_{14}, d_{15}$ such that

$$\begin{aligned}
&\sup_{(\delta,z)\in\mathcal{S}_\tau} \left|\delta\frac{\partial\tilde{h}(z)}{\partial z_j} - D\Phi_{1,(j)}(f;h)\right| \leq d_{10} \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}} |h| + d_{10} \sup_{(y,z)\in\mathbb{S}_{Y,Z}} \left|\frac{\partial h(y,z)}{\partial z_j}\right| \\
&\sup_{(\delta,z)\in\mathcal{S}_\tau} \left|\left(\delta\frac{\partial\tilde{f}(z)}{\partial z_j} - \Phi_{1,(j)}(f)\right) D\Phi_2(f;h)\right| \leq d_{11} \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}} |h| \\
&\sup_{(\delta,z)\in\mathcal{S}_\tau} \left|\left(\delta\frac{\partial\tilde{f}(z)}{\partial z_j} - \Phi_{1,(j)}(f)\right) R\Phi_2(f;h)\right| \leq d_{12} \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}} |h|^2 \\
&\qquad\qquad + d_{12} \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}} |h| \sup_{(y,z)\in\mathbb{S}_{Y,Z}} \left|\frac{\partial h(y,z)}{\partial y}\right| \\
&\sup_{(\delta,z)\in\mathcal{S}_\tau} |\Delta\Phi_2(f;h)| \leq d_{13} \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}} |h| \\
&\sup_{(\delta,z)\in\mathcal{S}_\tau} \left|\delta\frac{\partial\tilde{h}(z)}{\partial z_j} - \Delta\Phi_{1,(j)}(f;h)\right| \leq d_{14} \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}} |h| + d_{14} \sup_{(y,z)\in\mathbb{S}_{Y,Z}} \left|\frac{\partial h(y,z)}{\partial z_j}\right| \\
&\sup_{(\delta,z)\in\mathcal{S}_\tau} \left|\left(\delta\frac{\partial\tilde{f}(z)}{\partial z_j} - \Phi_1(f)\right) \Delta\Phi_2(f;h)\right| \leq d_{15} \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}} |h|.
\end{aligned} \tag{26}$$

Combining results from (24), (25), and (26), I have that for $\sup_{(y,z)\in\mathbb{S}_{(Y,Z)}} |h| \leq \min\{\tau/2, 1\}$ and some finite constants d_{16}, d_{17}

$$\sup_{(\delta,z)\in\mathcal{S}_\tau} \left|D\Phi_{(j)}(f;h)\right| \leq d_{16} \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}} |h| + d_{16} \sup_{(y,z)\in\mathbb{S}_{Y,Z}} \left|\frac{\partial h(y,z)}{\partial z_j}\right|$$

$$\sup_{(\delta,z)\in\mathcal{S}_\tau} \left|R\Phi_{(j)}(f;h)\right| \leq d_{17} \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}} |h|^2 + d_{17} \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}} |h| \sup_{(y,z)\in\mathbb{S}_{Y,Z}} \left|\frac{\partial h(y,z)}{\partial y}\right|$$

$$+ d_{17} \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}} |h| \sup_{(y,z)\in\mathbb{S}_{Y,Z}} \left|\frac{\partial h(y,z)}{\partial z_j}\right|.$$

Let $c_1 \equiv \max\{d_{16}, d_{17}\}$ gives the desired result. By the same logic, it can be shown that

$$\begin{aligned}
\sup_{(y,z)\in\mathbb{S}_\tau} \left|D\Psi_{(j)}(f;h)\right| &\leq d_{18} \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}} |h| + d_{18} \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}} \left|\frac{\partial h(y,z)}{\partial z_j}\right| \\
\sup_{(y,z)\in\mathbb{S}_\tau} \left|R\Psi_{(j)}(f;h)\right| &\leq d_{19} \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}} |h|^2 + d_{19} \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}} |h| \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}} \left|\frac{\partial h(y,z)}{\partial z_j}\right| \\
\sup_{(y,z)\in\mathbb{S}_\tau} \left|D\tilde{\Psi}(f;h)\right| &\leq d_{20} \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}} |h| + d_{20} \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}} \left|\frac{\partial h(y,z)}{\partial z_{j+1}}\right| \\
\sup_{(y,z)\in\mathbb{S}_\tau} \left|R\tilde{\Psi}(f;h)\right| &\leq d_{21} \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}} |h|^2 + d_{21} \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}} |h| \sup_{(y,z)\in\mathbb{S}_{(Y,Z)}} \left|\frac{\partial h(y,z)}{\partial z_{j+1}}\right|.
\end{aligned}$$

Let $c_2 \equiv \max\{d_{18}, d_{19}\}$, $c_3 \equiv \max\{d_{20}, d_{21}\}$ gives the desired result. ■

Lemma 12. *For any value $(y,z) \in \mathbb{S}_{(Y,Z)}$, and any value $\delta \in [0,1]$, define functional Ξ and $\tilde{\Xi}$ on* F *by $\Xi(g) \equiv \Phi_{(1)}(g) + \frac{\Phi_{(2)}(g)}{\tilde{\Psi}(g)}$ and $\tilde{\Xi}(g) \equiv \Psi_{(1)}(g) + \frac{\Psi_{(2)}(g)}{\tilde{\Psi}(g)}$. Then I have that*

$$\begin{aligned}
\Xi(f+h) - \Xi(f) &= D\Xi(f;h) + R\Xi(f;h) \\
\tilde{\Xi}(f+h) - \tilde{\Xi}(f) &= D\tilde{\Xi}(f;h) + R\tilde{\Xi}(f;h),
\end{aligned}$$

and that there exist constants e_1, e_2 such that

$$\begin{aligned}
\sup_{(\delta,x,w)\in\mathcal{S}_\tau} |D\Xi(f;h)| &\leq \frac{e_1}{\tau^2}\|h\| + e_1 \sup_{(y,x,w)\in\mathbb{S}_{Y,X,W}} \left|\frac{\partial h(y,x,w)}{\partial x}\right| + \frac{e_1}{\tau^2} \sup_{(y,x,w)\in\mathbb{S}_{Y,X,W}} \left|\frac{\partial h(y,x,w)}{\partial w}\right| \\
\sup_{(\delta,x,w)\in\mathcal{S}_\tau} |R\Xi(f;h)| &\leq \frac{e_1}{\tau^3}\|h\|^2 + \frac{e_1}{\tau^3}\|h\| \sup_{(y,x,w)\in\mathbb{S}_{Y,X,W}} \left|\frac{\partial h(y,x,w)}{\partial w}\right| \\
&\quad + \frac{e_1}{\tau}\|h\| \sup_{(y,x,w)\in\mathbb{S}_{Y,X,W}} \left|\frac{\partial h(y,x,w)}{\partial y}\right| + e_1\|h\| \sup_{(y,x,w)\in\mathbb{S}_{Y,X,W}} \left|\frac{\partial h(y,x,w)}{\partial x}\right| \\
&\quad + \frac{e_1}{\tau^3} \sup_{(y,x,w)\in\mathbb{S}_{Y,X,W}} \left|\frac{\partial h(y,x,w)}{\partial w}\right|^2 \\
\sup_{(y,x,w)\in\mathbb{S}_\tau} \left|D\tilde{\Xi}(f;h)\right| &\leq \frac{e_2}{\tau^2}\|h\| + e_2 \sup_{(y,x,w)\in\mathbb{S}_{Y,X,W}} \left|\frac{\partial h(y,x,w)}{\partial x}\right| + \frac{e_2}{\tau^2} \sup_{(y,x,w)\in\mathbb{S}_{Y,X,W}} \left|\frac{\partial h(y,x,w)}{\partial w}\right| \\
\sup_{(y,x,w)\in\mathbb{S}_\tau} \left|R\tilde{\Xi}(f;h)\right| &\leq \frac{e_2}{\tau^3}\|h\|^2 + \frac{e_2}{\tau^3}\|h\| \sup_{(y,x,w)\in\mathbb{S}_{Y,X,W}} \left|\frac{\partial h(y,x,w)}{\partial w}\right| \\
&\quad + e_2\|h\| \sup_{(y,x,w)\in\mathbb{S}_{Y,X,W}} \left|\frac{\partial h(y,x,w)}{\partial x}\right| + \frac{e_2}{\tau^3} \sup_{(y,x,w)\in\mathbb{S}_{Y,X,W}} \left|\frac{\partial h(y,x,w)}{\partial w}\right|^2.
\end{aligned}$$

Proof.

$$\Xi(f+h) - \Xi(f)$$

$$
\begin{aligned}
&= \left(\Phi_{(1)}(f+h) + \frac{\Phi_{(2)}(f+h)}{\tilde{\Psi}_{(1)}(f+h)}\right) - \left(\Phi_{(1)}(f) + \frac{\Phi_{(2)}(f)}{\tilde{\Psi}_{(1)}(f)}\right) \\
&= \Delta\Phi_{(1)}(f;h) + \frac{D\Phi_{(2)}(f;h)\tilde{\Psi}_{(1)}(f) - D\tilde{\Psi}_{(1)}(f;h)\Phi_{(2)}(f)}{\tilde{\Psi}^2_{(1)}(f)} \\
&\quad - \frac{D\Phi_{(2)}(f;h)\tilde{\Psi}_{(1)}(f)\Delta\tilde{\Psi}_{(1)}(f;h) - D\tilde{\Psi}_{(1)}(f;h)\Phi_{(2)}(f)\Delta\tilde{\Psi}_{(1)}(f;h)}{\tilde{\Psi}^2_{(1)}(f)\tilde{\Psi}_{(1)}(f+h)} \\
&\quad + \frac{R\Phi_{(2)}(f;h)\tilde{\Psi}_{(1)}(f) - R\tilde{\Psi}_{(1)}(f;h)\Phi_{(2)}(f)}{\tilde{\Psi}_{(1)}(f+h)\tilde{\Psi}_{(1)}(f)}
\end{aligned}
$$

$$
\begin{aligned}
&\tilde{\Xi}(f+h) - \tilde{\Xi}(f) \\
&= \left(\Psi_{(1)}(f+h) + \frac{\Psi_{(2)}(f+h)}{\tilde{\Psi}_{(1)}(f+h)}\right) - \left(\Psi_{(1)}(f) + \frac{\Psi_{(2)}(f)}{\tilde{\Psi}_{(1)}(f)}\right) \\
&= \Delta\Psi_{(1)}(f;h) + \frac{D\Psi_{(2)}(f;h)\tilde{\Psi}_{(1)}(f) - D\tilde{\Psi}_{(1)}(f;h)\Psi_{(2)}(f)}{\tilde{\Psi}^2_{(1)}(f)} \\
&\quad - \frac{D\Psi_{(2)}(f;h)\tilde{\Psi}_{(1)}(f)\Delta\tilde{\Psi}_{(1)}(f;h) - D\tilde{\Psi}_{(1)}(f;h)\Psi_{(2)}(f)\Delta\tilde{\Psi}_{(1)}(f;h)}{\tilde{\Psi}^2_{(1)}(f)\tilde{\Psi}_{(1)}(f+h)} \\
&\quad + \frac{R\Psi_{(2)}(f;h)\tilde{\Psi}_{(1)}(f) - R\tilde{\Psi}_{(1)}(f;h)\Psi_{(2)}(f)}{\tilde{\Psi}_{(1)}(f+h)\tilde{\Psi}_{(1)}(f)}
\end{aligned}
$$

Define

$$
\begin{aligned}
D\Xi(f;h) &\equiv D\Phi_{(1)}(f;h) + \frac{D\Phi_{(2)}(f;h)\tilde{\Psi}_{(1)}(f) - D\tilde{\Psi}_{(1)}(f;h)\Phi_{(2)}(f)}{\tilde{\Psi}^2_{(1)}(f)} \\
R\Xi(f;h) &\equiv R\Phi_{(1)}(f;h) - \frac{D\Phi_{(2)}(f;h)\tilde{\Psi}_{(1)}(f)\Delta\tilde{\Psi}_{(1)}(f;h) - D\tilde{\Psi}_{(1)}(f;h)\Phi_{(2)}(f)\Delta\tilde{\Psi}_{(1)}(f;h)}{\tilde{\Psi}^2_{(1)}(f)\tilde{\Psi}_{(1)}(f+h)} \\
&\qquad + \frac{R\Phi_{(2)}(f;h)\tilde{\Psi}_{(1)}(f) - R\tilde{\Psi}_{(1)}(f;h)\Phi_{(2)}(f)}{\tilde{\Psi}_{(1)}(f+h)\tilde{\Psi}_{(1)}(f)} \\
D\tilde{\Xi}(f;h) &\equiv D\Psi_{(1)}(f;h) + \frac{D\Psi_{(2)}(f;h)\tilde{\Psi}_{(1)}(f) - D\tilde{\Psi}_{(1)}(f;h)\Psi_{(2)}(f)}{\tilde{\Psi}^2_{(1)}(f)} \\
R\tilde{\Xi}(f;h) &\equiv R\Psi_{(1)}(f;h) - \frac{D\Psi_{(2)}(f;h)\tilde{\Psi}_{(1)}(f)\Delta\tilde{\Psi}_{(1)}(f;h) - D\tilde{\Psi}_{(1)}(f;h)\Psi_{(2)}(f)\Delta\tilde{\Psi}_{(1)}(f;h)}{\tilde{\Psi}^2_{(1)}(f)\tilde{\Psi}_{(1)}(f+h)} \\
&\qquad + \frac{R\Psi_{(2)}(f;h)\tilde{\Psi}_{(1)}(f) - R\tilde{\Psi}_{(1)}(f;h)\Psi_{(2)}(f)}{\tilde{\Psi}_{(1)}(f+h)\tilde{\Psi}_{(1)}(f)}.
\end{aligned}
$$

Then

$$
\begin{aligned}
\Xi(f+h) - \Xi(f) &= D\Xi(f;h) + R\Xi(f;h) \\
\tilde{\Xi}(f+h) - \tilde{\Xi}(f) &= D\tilde{\Xi}(f;h) + R\tilde{\Xi}(f;h).
\end{aligned}
$$

Note that there exist some constant d_{14} such that

$$
\sup_{(\delta,x,w)\in\mathcal{S}_\tau} \left|\Phi_{(2)}(f)\right| = \sup_{(\delta,x,w)\in\mathcal{S}_\tau} \left| \frac{\delta\frac{\partial \tilde{f}(x,w)}{\partial w} - \Phi_{1,(2)}(f)}{\Phi_2(f)} \right| \leq \frac{\sup_{(\delta,x,w)\in\mathcal{S}_\tau} \left|\delta\frac{\partial \tilde{f}(x,w)}{\partial w} - \Phi_{1,(2)}(f)\right|}{\inf_{(\delta,x,w)\in\mathcal{S}_\tau} |f(\alpha(f),x,w)|} \leq d_{14}
$$

$$\sup_{(y,x,w)\in\mathbb{S}_\tau}\left|\Psi_{(2)}(f)\right| = \sup_{(y,x,w)\in\mathbb{S}_\tau}\left|\frac{\Lambda(f)\frac{\partial\tilde{f}(x,w)}{\partial w} - \Psi_{1,(2)}(f)}{f(y,x,w)}\right|$$
$$\leq \frac{\sup_{(y,x,w)\in\mathbb{S}_\tau}\left|\Lambda(f)\frac{\partial\tilde{f}(x,w)}{\partial w} - \Psi_{1,(2)}(f)\right|}{\inf_{(y,x,w)\in\mathbb{S}_\tau}|f(y,x,w)|} \leq d_{14}.$$

Application of Lemma 11 yields the desired result. ■

Lemma 13. *In Lemma 6, let the dimension of vector $L = 2$. For any value $(y,z) \in \mathbb{S}_{(Y,Z)}$, and any value $\delta \in [0,1]$, define functional $\tilde{\Gamma}$ on* F *by*

$$\tilde{\Gamma}(g) \equiv \int\int_0^1 \omega(x,w^*,s(r(x,w^*),\delta))\Xi(g)\frac{\int g(y,x.w)dy}{\int\int g(y,x,w)dydw}d\delta dw,$$

where $\omega()$ is a known[4] *compactly supported weighting function as defined in (19), and $s()$ and $r()$ are as defined in (5) and (6). Then I have that*

$$\tilde{\Gamma}(f+h) - \tilde{\Gamma}(f) = D\tilde{\Gamma}(f;h) + R\tilde{\Gamma}(f;h),$$

and there exists a constant e_4 such that

$$\sup_{x\in\mathbb{S}_X}\left|D\tilde{\Gamma}(f;h)\right| \leq e_4\|h\| + e_4\sup_{(y,x,w)\in\mathbb{S}_{Y,X,W}}\left|\frac{\partial h(y,x,w)}{\partial x}\right| + e_4\sup_{(y,x,w)\in\mathbb{S}_{Y,X,W}}\left|\frac{\partial h(y,x,w)}{\partial w}\right|$$
$$\sup_{x\in\mathbb{S}_X}\left|R\tilde{\Gamma}(f;h)\right| \leq e_4\|h\|^2 + e_4\|h\|\sup_{(y,x,w)\in\mathbb{S}_{Y,X,W}}\left|\frac{\partial h(y,x,w)}{\partial w}\right| + e_4\|h\|\sup_{(y,x,w)\in\mathbb{S}_{Y,X,W}}\left|\frac{\partial h(y,x,w)}{\partial y}\right|.$$

Proof. Define functional Γ on F by $\Gamma(g) \equiv \Xi(g)\frac{\int g(y,x.w)dy}{\int\int g(y,x,w)dydw}$, then

$$\Gamma(f+h) - \Gamma(f) = \frac{\Xi(f+h)\int f(y,x.w) + h(y,x.w)dy}{\int\int f(y,x,w) + h(y,x.w)dydw} - \frac{\Xi(f)\int f(y,x.w)dy}{\int\int f(y,x,w)dydw}$$
$$\equiv \frac{N'}{D'} - \frac{N}{D}$$

where I've defined

$$N' \equiv \Xi(f+h)\int f(y,x.w) + h(y,x.w)dy$$
$$N \equiv \Xi(f)\int f(y,x.w)dy$$

[4] The same logic carries over if $\omega()$ is estimated. Take the example mentioned in the main text for instance. When $\tilde{\omega}(x,w^*)$ is a constant equal to $\left[(\tau_u - \tau_l)\int_{\underline{w}^*}^{\bar{w}^*} f_{W^*|X=x}(w^*)dw^*\right]^{-1}$, and $f_{W^*|X=x}(w^*)$ is estimated by a plug-in estimator, the same argument in this proof will follow by replacing $\int\int f(y,x,w)dydw$ with $\int_{\underline{w}^*}^{\bar{w}^*}\int f(y,x,w)dydw$ in the definition of Γ.

$$D' \equiv \int\int f(y,x,w)+h(y,x.w)dydw$$
$$D \equiv \int\int f(y,x,w)dydw$$

Note that

$$\begin{aligned} N'-N &= D\Xi(f;h)\int f(y,x,w)dy + \Xi(f)\int h(y,x,w)dy \\ &\quad + D\Xi(f;h)\int h(y,x,w)dy + R\Xi(f;h)\int (f(y,x,w)+h(y,x,w))dy \\ &\equiv DN + RN \\ D'-D &= \int\int h(y,x,w)dydw \end{aligned}$$

where $DN \equiv D\Xi(f;h)\int f(y,x,w)dy + \Xi(f)\int h(y,x,w)dy$ and $RN \equiv D\Xi(f;h)\int h(y,x,w)dy + R\Xi(f;h)\int(f(y,x,w)+h(y,x,w))dy$, and $D\Xi(f;h)$ and $R\Xi(f;h)$ are as defined in Lemma 12. Then by the same logic as Lemma 11,

$$\begin{aligned} &\Gamma(f+h)-\Gamma(f) \\ &= \frac{\left(D\Xi(f;h)\int f(y,x,w)dy + \Xi(f)\int h(y,x,w)dy\right)\int\int f(y,x,w)dydw - \Xi(f)\int f(y,x.w)dy\left(\int\int h(y,x.w)dydw\right)}{\left(\int\int f(y,x,w)dydw\right)^2} \\ &\quad + \frac{\int\int f(y,x,w)dydw \times \left(D\Xi(f;h)\int h(y,x,w)dy + R\Xi(f;h)\int(f(y,x,w)+h(y,x,w))dy\right)}{\left(\int\int f(y,x,w)dydw\right)^2} \\ &\quad - \int\int h(y,x,w)dydw \times \\ &\quad \frac{\left((\Xi(f+h)\int f(y,x.w)+h(y,x.w)dy\right)\int\int f(y,x,w)dydw - \Xi(f)\int f(y,x.w)dy\left(\int\int f(y,x,w)+h(y,x.w)dydw\right))}{\left(\int\int f(y,x,w)+h(y,x.w)dydw\right)\left(\int\int f(y,x,w)dydw\right)^2} \\ &= D\Gamma(f;h)+R\Gamma(f;h). \end{aligned}$$

where I've defined

$$\begin{aligned} &D\Gamma(f;h) \\ &\quad \equiv \frac{\left(D\Xi(f;h)\int f(y,x,w)dy + \Xi(f)\int h(y,x,w)dy\right)\int\int f(y,x,w)dydw - \Xi(f)\int f(y,x.w)dy\left(\int\int h(y,x.w)dydw\right)}{\left(\int\int f(y,x,w)dydw\right)^2} \\ &R\Gamma(f;h) \\ &\quad \equiv + \frac{\int\int f(y,x,w)dydw \times \left(D\Xi(f;h)\int h(y,x,w)dy + R\Xi(f;h)\int(f(y,x,w)+h(y,x,w))dy\right)}{\left(\int\int f(y,x,w)dydw\right)^2} \\ &\quad - \int\int h(y,x,w)dydw \times \\ &\quad \frac{\left((\Xi(f+h)\int f(y,x.w)+h(y,x.w)dy\right)\int\int f(y,x,w)dydw - \Xi(f)\int f(y,x.w)dy\left(\int\int f(y,x,w)+h(y,x.w)dydw\right))}{\left(\int\int f(y,x,w)+h(y,x.w)dydw\right)\left(\int\int f(y,x,w)dydw\right)^2}. \end{aligned}$$

By the boundedness of support, the definition of $\mathcal{S}_\tau$ and Lemma 12, there exists a constant

e_3 such that

$$\sup_{(\delta,x,w)\in\mathcal{S}_\tau}|D\Gamma(f;h)| \leq e_3\|h\| + e_3 \sup_{(y,x,w)\in\mathbb{S}_{Y,X,W}}\left|\frac{\partial h(y,x,w)}{\partial x}\right| + e_3 \sup_{(y,x,w)\in\mathbb{S}_{Y,X,W}}\left|\frac{\partial h(y,x,w)}{\partial w}\right|$$
$$\sup_{(\delta,x,w)\in\mathcal{S}_\tau}|R\Gamma(f;h)| \leq e_3\|h\|^2 + e_3\|h\| \sup_{(y,x,w)\in\mathbb{S}_{Y,X,W}}\left|\frac{\partial h(y,x,w)}{\partial w}\right| + e_3\|h\| \sup_{(y,x,w)\in\mathbb{S}_{Y,X,W}}\left|\frac{\partial h(y,x,w)}{\partial y}\right|$$

Denote the support of $\omega(x, w^*, s(r(x, w^*), \delta))$ as $\mathcal{M}$, then by the argument stated at the beginning of Appendix A, there exists $\tau > 0$ such that $\mathcal{M} \subset \mathcal{S}_\tau$ so that I have

$$\sup_{(\delta,x,w)\in\mathcal{M}}|D\Gamma(f;h)| \leq e_3\|h\| + e_3 \sup_{(y,x,w)\in\mathbb{S}_{Y,X,W}}\left|\frac{\partial h(y,x,w)}{\partial x}\right| + e_3 \sup_{(y,x,w)\in\mathbb{S}_{Y,X,W}}\left|\frac{\partial h(y,x,w)}{\partial w}\right|$$
$$\sup_{(\delta,x,w)\in\mathcal{M}}|R\Gamma(f;h)| \leq e_3\|h\|^2 + e_3\|h\| \sup_{(y,x,w)\in\mathbb{S}_{Y,X,W}}\left|\frac{\partial h(y,x,w)}{\partial w}\right| + e_3\|h\| \sup_{(y,x,w)\in\mathbb{S}_{Y,X,W}}\left|\frac{\partial h(y,x,w)}{\partial y}\right|.$$

By definition, I have

$$\begin{aligned}
&\int\int_0^1 \omega(x,w^*,s(r(x,w^*),\delta))\Gamma(f+h)d\delta dw - \int\int_0^1 \omega(x,w^*,s(r(x,w^*),\delta))\Gamma(f)d\delta dw \\
&= \int\int_0^1 \omega(x,w^*,s(r(x,w^*),\delta))D\Gamma(f;h)d\delta dw + \int\int_0^1 \omega(x,w^*,s(r(x,w^*),\delta))R\Gamma(f;h)d\delta dw \\
&\equiv D\tilde{\Gamma}(f;h) + R\tilde{\Gamma}(f;h)
\end{aligned}$$

where I've defined $D\tilde{\Gamma}(f;h) \equiv \int\int_0^1 \omega(x,w^*,s(r(x,w^*),\delta))D\Gamma(f;h)d\delta dw$ and $R\tilde{\Gamma}(f;h) \equiv \int\int_0^1 \omega(x,w^*,s(r(x,w^*),\delta))R\Gamma(f;h)d\delta dw$. By the compactness of $\mathcal{M}$, there exist a constant e_4 such that

$$\sup_{x\in\mathbb{S}_X}\left|D\tilde{\Gamma}(f;h)\right| \leq e_4\|h\| + e_4 \sup_{(y,x,w)\in\mathbb{S}_{Y,X,W}}\left|\frac{\partial h(y,x,w)}{\partial x}\right| + e_4 \sup_{(y,x,w)\in\mathbb{S}_{Y,X,W}}\left|\frac{\partial h(y,x,w)}{\partial w}\right|$$
$$\sup_{x\in\mathbb{S}_X}\left|R\tilde{\Gamma}(f;h)\right| \leq e_4\|h\|^2 + e_4\|h\| \sup_{(y,x,w)\in\mathbb{S}_{Y,X,W}}\left|\frac{\partial h(y,x,w)}{\partial w}\right| + e_4\|h\| \sup_{(y,x,w)\in\mathbb{S}_{Y,X,W}}\left|\frac{\partial h(y,x,w)}{\partial y}\right|.$$

■

7.2 Appendix B: Proofs of Theorems and Lemmas in the Main Text

Proof of Lemma 1. Since $W^* \perp (\epsilon, \eta)$, then $\epsilon \perp W^* \mid \eta$. Define $\delta = F_{\epsilon|\eta}(\epsilon)$. Then I can write $\epsilon = F^{-1}_{\epsilon|\eta}(\delta) =: s(\eta, \delta)$. By Assumption 8, s is strictly increasing in δ. To see that $\delta \perp (X, W^*)$ and is uniform (0,1), note that

$$\begin{aligned}
F_{\delta|W^*=w^*,\eta=\bar{\eta}}(t) &= Pr\left(F_{\epsilon|\eta}(\epsilon) \leq t \mid W^* = w^*, \eta = \bar{\eta}\right) \\
&= Pr\left(\epsilon \leq F^{-1}_{\epsilon|\eta=\bar{\eta}}(t) \mid W^* = w^*, \eta = \bar{\eta}\right)
\end{aligned}$$

$$= Pr\left(\epsilon \leq F^{-1}_{\epsilon|\eta=\bar{\eta}}(t) \mid \eta = \bar{\eta}\right)$$
$$= F_{\epsilon|\eta=\bar{\eta}}(F^{-1}_{\epsilon|\eta=\bar{\eta}}(t)) = t,$$

which says that the conditional distribution of δ given W^*, η is $U(0,1)$ regardless of the values of W^* and η. In other words, $\delta \sim U(0,1)$ and is independent of (W^*, η). Since X is a function of (W^*, η), δ is independent of X as well. ■

Proof of Lemma 2. (i)First I show that $\mathbb{S}_{\epsilon|X=x} \subseteq \mathbb{S}_{\mathbf{s}|X=x}$. For any $\bar{\epsilon} \in \mathbb{S}_{\epsilon|X=x}$, let $\bar{\delta} \equiv F_{\epsilon|\eta=r(x,\bar{w}^*)}(\bar{\epsilon})$ for some $\bar{w}^* \in \mathbb{S}_{W^*|X=x}$. Then $\bar{\delta} \in [0,1]$ and thus $\bar{\epsilon} \equiv s(r(x,\bar{w}^*),\bar{\delta}) \in \mathbb{S}_{\mathbf{s}|X=x}$.

(ii) Then I show that $\mathbb{S}_{\mathbf{s}|X=x} \subseteq \mathbb{S}_{\epsilon|X=x}$. By definition of $\mathbb{S}_{\mathbf{s}|X=x}$, for each $\tilde{s} \in \mathbb{S}_{\mathbf{s}|X=x}$, there is some $(\tilde{w}^*, \tilde{\delta}) \in \mathbb{S}_{W^*|X=x} \times [0,1]$ such that $\tilde{s} = s(r(x,\tilde{w}^*),\tilde{\delta})$. Denote the upper and lower bound (potentially infinity) of $\mathbb{S}_{\epsilon|X=x}$ as $\bar{e}_x$ and $\underline{e}_x$, respectively. From Lemma 1, I know that $\epsilon \perp W^* \mid \eta$. Since X is a function of W^*, η, I have that $\epsilon \perp X \mid \eta$. Then if $\bar{e}_x \in \mathbb{S}_{\epsilon|X=x}$, then $Pr\left(\epsilon \leq \bar{e}_x \mid \eta = r(x,\tilde{w}^*)\right) = Pr\left(\epsilon \leq \bar{e}_x \mid \eta = r(x,\tilde{w}^*), X = x\right) = 1$, which implies $F^{-1}_{\epsilon|\eta=r(x,\tilde{w}^*)}(1) \leq \bar{e}_x$. If $\bar{e}_x \notin \mathbb{S}_{\epsilon|X=x}$, then $Pr\left(\epsilon < \bar{e}_x \mid \eta = r(x,\tilde{w}^*)\right) = Pr\left(\epsilon < \bar{e}_x \mid \eta = r(x,\tilde{w}^*), X = x\right) = 1$, which implies $F^{-1}_{\epsilon|\eta=r(x,\tilde{w}^*)}(1) < \bar{e}_x$. Similarly, $F^{-1}_{\epsilon|\eta=r(x,\tilde{w}^*)}(0) \geq \bar{e}_x$, and the inequality is strict if $\underline{e}_x \notin \mathbb{S}_{\epsilon|X=x}$ Since $\tilde{\delta} \in [0,1]$, by Assumption 8, $\underline{e}_x \leq F^{-1}_{\epsilon|\eta=r(x,\tilde{w}^*)}(\tilde{\delta}) \leq \bar{e}_x$, or equivalently, $\underline{e}_x \leq s(r(x,\tilde{w}^*),\tilde{\delta}) \leq \bar{e}_x$, and the inequalities are strict if $\underline{e}_x \notin \mathbb{S}_{\epsilon|X=x}$ or $\bar{e}_x \notin \mathbb{S}_{\epsilon|X=x}$, respectively. Thus $\tilde{s} \in \mathbb{S}_{\epsilon|X=x}$. Combining results from (i) and (ii) yields the desired conclusion. ■

Proof of Lemma 3.

$$\delta = F_{Y|X=x,W^*=w^*}\left(F^{-1}_{Y|X=x,W^*=w^*}(\delta)\right)$$
$$= \frac{\int_{-\infty}^{F^{-1}_{Y|X=x,W^*=w^*}(\delta)} f_{Y,X,W^*}(y,x,w^*)dy}{\int_{-\infty}^{\infty} f_{Y,X,W^*}(y,x,w^*)dy}$$
$$\Longrightarrow \delta f_{X,W^*}(x,w^*) = \int_{-\infty}^{F^{-1}_{Y|X=x,W^*=w^*}(\delta)} f_{Y,X,W^*}(y,x,w^*)dy$$

Taking derivatives on both sides with respect to w^* yields

$$f_{Y,X,W^*}\left(F^{-1}_{Y|X=x,W^*=w^*}(\delta),x,w^*\right)\frac{\partial F^{-1}_{Y|X=x,W^*=w^*}(\delta)}{\partial w^*}+\int_{-\infty}^{F^{-1}_{Y|X=x,W^*=w^*}(\delta)}\frac{\partial f_{Y,X,W^*}(y,x,w^*)}{\partial w^*}dy$$
$$=\delta\frac{\partial f_{X,W^*}(x,w^*)}{\partial w^*}$$

This implies

$$\frac{\partial F^{-1}_{Y|X=x,W^*=w^*}(\delta)}{\partial w^*}=\frac{\delta\frac{\partial f_{X,W^*}(x,w^*)}{\partial w^*}-\int_{-\infty}^{F^{-1}_{Y|X=x,W^*=w^*}(\delta)}\frac{\partial f_{Y,X,W^*}(y,x,w^*)}{\partial w^*}dy}{f_{Y,X,W^*}\left(F^{-1}_{Y|X=x,W^*=w^*}(\delta),x,w^*\right)}.$$

The conclusion for $\frac{\partial F^{-1}_{Y|X=x,W^*=w^*}(\delta)}{\partial x}$ follows the same logic. ■

Proof of Lemma 4. Assumption 5, 6, and 10 ensure the existence of

$$\phi_{f_{Y,X,W_2}(y,x,\cdot)}(t)=E\left[e^{\mathbf{i}tW_2}\mid Y=y,X=x\right]f_{Y,X}(y,x)$$
$$\phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(y,x,\cdot)}(t)=\int e^{\mathbf{i}tw_2}\frac{\partial^{\lambda_2}f_{Y,X,W_2}(y,x,w_2)}{\partial y^{\lambda}_{2,1}\partial x^{\lambda}_{2,2}}dw_2 \quad \text{and thus}$$
$$g_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h_1)\equiv\int\frac{1}{h_1}K\left(\frac{\tilde{w}^*-w^*}{h_1}\right)g_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,\tilde{w}^*)d\tilde{w}^*$$
$$=\int\frac{1}{h_1}K\left(\frac{\tilde{w}^*-w^*}{h_1}\right)\left(\frac{1}{2\pi}\int(-\mathbf{i}t)^{\lambda_1}e^{-\mathbf{i}t\tilde{w}^*}\frac{\phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(y,x,\cdot)}(t)\phi_{W^*}(t)}{\phi_{W_2}(t)}dt\right)d\tilde{w}^*.$$

Denote $k_{h_1}(v)\equiv\frac{1}{h_1}K\left(\frac{v}{h_1}\right)$, then for all $x\in\mathbb{R}$, $g_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h_1)$ is the convolution between $k_{h_1}(\cdot)$ and $g_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,\cdot)$. By the convolution theorem, it's equal to the inverse Fourier transform of the product of the Fourier transformation of $k_{h_1}(\cdot)$ and the Fourier transformation of $g_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,\cdot)$. The Fourier transformation of $k_{h_1}(\cdot)$ is $\phi_K(ht)$, and the Fourier transformation of $g_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,\cdot)$ is $(-\mathbf{i}t)^{\lambda_1}\frac{\phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(y,x,\cdot)}(t)\phi_{W^*}(t)}{\phi_{W_2}(t)}$. ■

Proof of Lemma 5. Write $\hat{\theta}(\zeta)\equiv\hat{E}\left[W_1e^{\mathrm{i}\zeta W_2}\right]$, $\delta\hat{\theta}(t)\equiv\hat{\theta}(t)-\theta(t)$, $\delta\hat{\phi}_Z(t)=\hat{\phi}_Z(t)-\phi_Z(t)$ for random variable Z, and $\delta\hat{\phi}_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(y,x,\cdot)}(t)=\hat{\phi}_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(y,x,\cdot)}(t)-\phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(y,x,\cdot)}(t)$.

Denote $q_{\lambda_2}(t)\equiv\frac{\phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(y,x,\cdot)}(t)}{\phi_{W_2}(t)}$, $\hat{q}_{\lambda_2}(t)\equiv\frac{\hat{\phi}_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(y,x,\cdot)}(t)}{\hat{\phi}_{W_2}(t)}$ and $\delta\hat{q}_{\lambda_2}(t)\equiv\hat{q}_{\lambda_2}(t)-q_{\lambda_2}(t)$, then $\delta\hat{q}_{\lambda_2}(t)$ can be written as

$$\delta\hat{q}_{\lambda_2}(t)=\left(\frac{\delta\hat{\phi}_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(y,x,\cdot)}(t)}{\phi_{W_2}(t)}-\frac{\phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(y,x,\cdot)}(t)\delta\hat{\phi}_{W_2}(t)}{\left(\phi_{W_2}(t)\right)^2}\right)\left(1+\frac{\delta\hat{\phi}_{W_2}(t)}{\phi_{W_2}(t)}\right)^{-1} \quad \text{or}$$

$$\delta\hat{q}_{\lambda_2}(t)=\delta_1\hat{q}_{\lambda_2}(t)+\delta_2\hat{q}_{\lambda_2}(t), \quad \text{with}$$

$$\delta_1 \hat{q}_{\lambda_2}(t) \equiv \frac{\delta\hat{\phi}_{f_{Y,X,W_2}^{(\lambda_{2,1},\lambda_{2,2})}(y,x,\cdot)}(t)}{\phi_{W_2}(t)} - \frac{\phi_{f_{Y,X,W_2}^{(\lambda_{2,1},\lambda_{2,2})}(y,x,\cdot)}(t)\delta\hat{\phi}_{W_2}(t)}{\left(\phi_{W_2}(t)\right)^2}$$

$$\delta_2 \hat{q}_{\lambda_2}(t) \equiv \frac{\phi_{f_{Y,X,W_2}^{(\lambda_{2,1},\lambda_{2,2})}(y,x,\cdot)}(t)}{\phi_{W_2}(t)} \left(\frac{\delta\hat{\phi}_{W_2}(t)}{\phi_{W_2}(t)}\right)^2 \left(1 + \frac{\delta\hat{\phi}_{W_2}(t)}{\phi_{W_2}(t)}\right)^{-1}$$
$$- \frac{\delta\hat{\phi}_{f_{Y,X,W_2}^{(\lambda_{2,1},\lambda_{2,2})}(y,x,\cdot)}(t)}{\phi_{W_2}(t)} \frac{\delta\hat{\phi}_{W_2}(t)}{\phi_{W_2}(t)} \left(1 + \frac{\delta\hat{\phi}_{W_2}(t)}{\phi_{W_2}(t)}\right)^{-1}. \tag{27}$$

Similarly, I denote $q_{W_1}(\xi) \equiv \frac{\theta(\xi)}{\phi_{W_2}(\xi)}$, $\hat{q}_{W_1}(\xi) \equiv \frac{\hat{\theta}(\xi)}{\hat{\phi}_{W_2}(\xi)}$, and $\delta\hat{q}_{W_1}(\xi) \equiv \hat{q}_{W_1}(\xi) - q_{W_1}(\xi)$. Then $\delta\hat{q}_{W_1}(\xi)$ can be written as

$$\delta\hat{q}_{W_1}(\xi) = \left(\frac{\delta\hat{\theta}(\xi)}{\phi_{W_2}(\xi)} - \frac{\theta(\xi)\delta\hat{\phi}_{W_2}(\xi)}{\left(\phi_{W_2}(\xi)\right)^2}\right)\left(1 + \frac{\delta\hat{\phi}_{W_2}(\xi)}{\phi_{W_2}(\xi)}\right)^{-1} \quad \text{or}$$

$$\delta\hat{q}_{W_1}(\xi) = \delta_1\hat{q}_{W_1}(\xi) + \delta_2\hat{q}_{W_1}(\xi), \quad \text{with}$$

$$\delta_1\hat{q}_{W_1}(\xi) \equiv \frac{\delta\hat{\theta}(\xi)}{\phi_{W_2}(\xi)} - \frac{\theta(\xi)\delta\hat{\phi}_{W_2}(\xi)}{\left(\phi_{W_2}(\xi)\right)^2}$$

$$\delta_2\hat{q}_{W_1}(\xi) \equiv \frac{\theta(\xi)}{\phi_{W_2}(\xi)}\left(\frac{\delta\hat{\phi}_{W_2}(\xi)}{\phi_{W_2}(\xi)}\right)^2\left(1 + \frac{\delta\hat{\phi}_{W_2}(\xi)}{\phi_{W_2}(\xi)}\right)^{-1}$$
$$- \frac{\delta\hat{\theta}(\xi)}{\phi_{W_2}(\xi)}\frac{\delta\hat{\phi}_{W_2}(\xi)}{\phi_{W_2}(\xi)}\left(1 + \frac{\delta\hat{\phi}_{W_2}(\xi)}{\phi_{W_2}(\xi)}\right)^{-1}. \tag{28}$$

For $Q(t) = \int_0^t \frac{\mathbf{i}\theta(\xi)}{\phi_{W_2}(\xi)}d\xi$, $\delta\hat{Q}(t) = \int_0^t \frac{\mathbf{i}\hat{\theta}(\xi)}{\hat{\phi}_{W_2}(\xi)}d\xi - Q(t)$ and some random function $\delta\bar{Q}(t)$ such that $|\delta\bar{Q}(t)| \leq |\delta\hat{Q}(t)|$ for all t,

$$\exp\left(Q(t) + \delta\hat{Q}(t)\right) = \exp(Q(t))\left(1 + \delta Q(t) + \frac{1}{2}[\exp(\delta\bar{Q}(t))]\left(\delta\hat{Q}(t)\right)^2\right) \tag{29}$$

Then I can state a useful representation for

$$\frac{\hat{\phi}_{f_{Y,X,W_2}^{(\lambda_{2,1},\lambda_{2,2})}(y,x,\cdot)}(t)\hat{\phi}_{W^*}(t)}{\hat{\phi}_{W_2}(t)}$$
$$= (q_{\lambda_2}(t) + \delta\hat{q}_{\lambda_2}(t))\exp(Q(t))\left(1 + \delta\hat{Q}(t) + \frac{1}{2}\exp\left(\delta\bar{Q}(t)\right)\left(\delta\hat{Q}(t)\right)^2\right)$$
$$= q_{\lambda_2}(t)\exp(Q(t)) + q_{\lambda_2}(t)\exp(Q(t))\left(\delta\hat{Q}(t) + \frac{1}{2}\exp\left(\delta\bar{Q}(t)\right)\left(\delta\hat{Q}(t)\right)^2\right)$$
$$+ \delta\hat{q}_{\lambda_2}(t)\exp(Q(t))\left(1 + \delta\hat{Q}(t) + \frac{1}{2}\exp\left(\delta\bar{Q}(t)\right)\left(\delta\hat{Q}(t)\right)^2\right)$$

$$
\begin{aligned}
&= q_{\lambda_2}(t)\exp(Q(t)) + q_{\lambda_2}(t)\exp(Q(t))\left[\int_0^t \mathbf{i}\delta_1\hat{q}_{W_1}(\xi)d\xi\right] + \delta_1\hat{q}_{\lambda_2}(t)\exp(Q(t)) \\
&\quad + q_{\lambda_2}(t)\exp(Q(t))\left[\int_0^t \mathbf{i}\delta_2\hat{q}_{W_1}(\xi)d\xi\right] + q_{\lambda_2}(t)\exp(Q(t))\left[\frac{1}{2}\exp\left(\delta\bar{Q}(t)\right)\left(\delta\hat{Q}(t)\right)^2\right] \\
&\quad + \delta_2\hat{q}_{\lambda_2}(t)\exp(Q(t)) + \delta\hat{q}_{\lambda_2}(t)\exp(Q(t))\left(\delta\hat{Q}(t) + \frac{1}{2}\exp\left(\delta\bar{Q}(t)\right)\left(\delta\hat{Q}(t)\right)^2\right). \qquad (30)
\end{aligned}
$$

Plug (30) into $\hat{g}_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h)$, I get

$$
\begin{aligned}
&\hat{g}_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h) - g_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h) \\
&= Dg_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h) + Rg_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h)
\end{aligned}
$$

where $Dg_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h)$ is the term linear in $\delta\hat{\theta}(t)$, $\delta\hat{\phi}_{W_2}$, and $\delta\hat{\phi}_{f_{Y,X,W_2}^{(\lambda_{2,1},\lambda_{2,2})}(y,x,\cdot)}(t)$, and $Rg_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h)$ is the higher-order remainder term. More specifically,

$$
\begin{aligned}
Dg_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h) &\equiv \frac{1}{2\pi}\int(-\mathbf{i}t)^{\lambda_1}e^{-\mathbf{i}tw^*}\phi_K(h_{1n}t) \\
&\qquad \left(q_{\lambda_2}(t)\exp(Q(t))\left[\int_0^t \mathbf{i}\delta_1\hat{q}_{W_1}(\xi)d\xi\right] + \delta_1\hat{q}_{\lambda_2}(t)\exp(Q(t))\right)dt \\
Rg_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h) &\equiv \frac{1}{2\pi}\int(-\mathbf{i}t)^{\lambda_1}e^{-\mathbf{i}tw^*}\phi_K(h_{1n}t) \\
&\qquad \Bigg\{q_{\lambda_2}(t)\exp(Q(t))\left[\int_0^t \mathbf{i}\delta_2\hat{q}_{W_1}(\xi)d\xi\right] \\
&\qquad + q_{\lambda_2}(t)\exp(Q(t))\left[\frac{1}{2}\exp\left(\delta\bar{Q}(t)\right)\left(\delta\hat{Q}(t)\right)^2\right] \\
&\qquad + \delta_2\hat{q}_{\lambda_2}(t)\exp(Q(t)) \\
&\qquad + \delta\hat{q}_{\lambda_2}(t)\exp(Q(t))\left(\delta\hat{Q}(t) + \frac{1}{2}\exp\left(\delta\bar{Q}(t)\right)\left(\delta\hat{Q}(t)\right)^2\right)\Bigg\}dt.
\end{aligned}
$$

Define $\bar{g}_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h) \equiv g_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h) + Dg_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h)$ Then I have the following expression:

$$
\begin{aligned}
&\hat{g}_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h) - g_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*) \\
&= B_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h) + L_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h) + R_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h)
\end{aligned}
$$

where

$$
\begin{aligned}
B_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h) &\equiv E\left[\bar{g}_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h)\right] - g_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*) \\
L_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h) &\equiv \bar{g}_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h) - E\left[\bar{g}_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h)\right] \\
R_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h) &\equiv \hat{g}_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h) - \bar{g}_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h).
\end{aligned}
$$

To see the explicit form of $B_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h)$, note that

$$
\begin{aligned}
&B_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h) \equiv E\left[\bar{g}_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h)\right] - g_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*) \\
&= g_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h) - g_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*) + E\left[Dg_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h)\right] \\
&= g_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h) - g_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*) \\
&\quad + \frac{1}{2\pi}\int (-\mathbf{i}t)^{\lambda_1} e^{-\mathbf{i}tw^*}\phi_K(h_{1n}t)\frac{E\left[\delta\hat{\phi}_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(y,x,\cdot)}(t)\right]\phi_{W^*}(t)}{\phi_{W_2}(t)}dt
\end{aligned}
$$

To see the explicit form of $L_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h)$, note that

$$
\begin{aligned}
&L_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h) \equiv \bar{g}_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h) - E\left[\bar{g}_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h)\right] \\
&= Dg_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h) - E\left[Dg_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h)\right] \\
&= \frac{1}{2\pi}\int (-\mathbf{i}t)^{\lambda_1} e^{-\mathbf{i}tw^*}\phi_K(h_{1n}t) \\
&\quad \left(\phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W^*}(y,x,\cdot)}(t)\left[\int_0^t \left(\frac{\mathbf{i}\delta\hat{\theta}(\xi)}{\phi_{W_2}(\xi)} - \frac{\mathbf{i}\theta(\xi)\delta\hat{\phi}_{W_2}(\xi)}{\left(\phi_{W_2}(\xi)\right)^2}\right)d\xi\right] - \frac{\delta\hat{\phi}_{W_2}(t)}{\phi_{W_2}(t)}\phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W^*}(y,x,\cdot)}(t)\right)dt \\
&\quad + \frac{1}{2\pi}\int (-\mathbf{i}t)^{\lambda_1} e^{-\mathbf{i}tw^*}\phi_K(h_{1n}t)\frac{\left(\hat{\phi}_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(y,x,\cdot)}(t) - E\left[\hat{\phi}_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(y,x,\cdot)}(t)\right]\right)\phi_{W^*}(t)}{\phi_{W_2}(t)}dt \\
&= \frac{1}{2\pi}\int (-\mathbf{i}t)^{\lambda_1} e^{-\mathbf{i}tw^*}\phi_K(h_1t)\phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W^*}(y,x,\cdot)}(t)\left[\int_0^t \left(\frac{\mathbf{i}\delta\hat{\theta}(\xi)}{\phi_{W_2}(\xi)} - \frac{\mathbf{i}\theta(\xi)\delta\hat{\phi}_{W_2}(\xi)}{\left(\phi_{W_2}(\xi)\right)^2}\right)d\xi\right]dt \\
&+ \quad \frac{1}{2\pi}\int (-\mathbf{i}t)^{\lambda_1} e^{-\mathbf{i}tw^*}\phi_K(h_1t) \\
&\times \left(\frac{\hat{\phi}_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(y,x,\cdot)}(t) - E\left[\hat{\phi}_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(y,x,\cdot)}(t)\right]}{\phi_{W_2}(t)}\phi_{W^*}(t) - \frac{\delta\hat{\phi}_{W_2}(t)}{\phi_{W_2}(t)}\phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W^*}(y,x,\cdot)}(t)\right)dt
\end{aligned}
$$

Then using the identity

$$
\begin{aligned}
\int_{-\infty}^{\infty}\int_0^{\xi} f(\xi,\zeta)d\zeta d\xi &= \int_0^{\infty}\int_{\zeta}^{\infty} f(\xi,\zeta)d\xi d\zeta + \int_{-\infty}^{0}\int_{\zeta}^{-\infty} f(\xi,\zeta)d\xi d\zeta \\
&:= \iint_{\zeta}^{\pm\infty} f(\xi,\zeta)d\xi d\zeta
\end{aligned}
$$

for any absolutely integrable function f, I obtain

$$
\begin{aligned}
&L_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h) && (31) \\
&= \bar{g}_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h) - g_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h) \\
&= \frac{1}{2\pi}\int\int_{\xi}^{\pm\infty} (-\mathbf{i}t)^{\lambda_1} e^{-\mathbf{i}tw^*}\phi_K(h_1t)\phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W^*}(y,x,\cdot)}(t)dt\left(\frac{\mathbf{i}\delta\hat{\theta}(\xi)}{\phi_{W_2}(\xi)} - \frac{\mathbf{i}\theta(\xi)\delta\hat{\phi}_{W_2}(\xi)}{\left(\phi_{W_2}(\xi)\right)^2}\right)d\xi \\
&+ \frac{1}{2\pi}\int (-\mathbf{i}t)^{\lambda_1} e^{-\mathbf{i}tw^*}\phi_K(h_1t)\times && (32)
\end{aligned}
$$

$$\left(\frac{\hat{\phi}_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(y,x,\cdot)}(t) - E\left[\hat{\phi}_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(y,x,\cdot)}(t)\right]}{\phi_{W_2}(t)} \phi_{W^*}(t) - \frac{\delta\hat{\phi}_{W_2}(t)}{\phi_{W_2}(t)} \phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W^*}(y,x,\cdot)}(t) \right) dt$$

$$= \int \Psi_{1,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi, y, x, w^*, h_1) \left(\hat{E}[W_1 e^{\mathbf{i}\xi W_2}] - E[W_1 e^{\mathbf{i}\xi W_2}]\right) d\xi$$

$$+ \int \Psi_{2,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi, y, x, w^*, h_1) \left(\hat{E}[e^{\mathbf{i}\xi W_2}] - E[e^{\mathbf{i}\xi W_2}]\right) d\xi$$

$$+ \int \Psi_{3,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi, y, x, w^*, h_1) \left(\hat{E}\left[e^{\mathbf{i}\xi W_2} \frac{1}{h_{2,1}^{1+\lambda_{2,1}} h_{2,2}^{1+\lambda_{2,2}}} G_Y\left(\frac{y-Y}{h_{2,1}}\right) G_X^{(\lambda_{2,2})}\left(\frac{x-X}{h_{2,2}}\right)\right] \right.$$

$$\left. - E\left[e^{\mathbf{i}\xi W_2} \frac{1}{h_{2,1}^{1+\lambda_{2,1}} h_{2,2}^{1+\lambda_{2,2}}} G_Y\left(\frac{y-Y}{h_{2,1}}\right) G_X^{(\lambda_{2,2})}\left(\frac{x-X}{h_{2,2}}\right)\right] \right) d\xi$$

$$= \hat{E}\left[\int \Psi_{1,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi, y, x, w^*, h_1) \left(W_1 e^{\mathbf{i}\xi W_2} - E[W_1 e^{\mathbf{i}\xi W_2}]\right) d\xi \right.$$

$$+ \int \Psi_{2,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi, y, x, w^*, h_1) \left(e^{\mathbf{i}\xi W_2} - E[e^{\mathbf{i}\xi W_2}]\right) d\xi$$

$$+ \int \Psi_{3,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi, y, x, w^*, h_1) \times \left(e^{\mathbf{i}\xi W_2} \frac{1}{h_{2,1}^{1+\lambda_{2,1}} h_{2,2}^{1+\lambda_{2,2}}} G_Y\left(\frac{y-Y}{h_{2,1}}\right) G_X^{(\lambda_{2,2})}\left(\frac{x-X}{h_{2,2}}\right) \right.$$

$$\left.\left. - E\left[e^{\mathbf{i}\xi W_2} \frac{1}{h_{2,1}^{1+\lambda_{2,1}} h_{2,2}^{1+\lambda_{2,2}}} G_Y\left(\frac{y-Y}{h_{2,1}}\right) G_X^{(\lambda_{2,2})}\left(\frac{x-X}{h_{2,2}}\right)\right] \right) d\xi \right]$$

$$= \hat{E}[l_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y, x, w^*, h; Y, X, W_1, W_2)], \tag{33}$$

where

$$\Psi_{1,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi, y, x, w^*, h_1) = \frac{1}{2\pi} \frac{\mathbf{i}}{\phi_{W_2}(\xi)} \int_{\xi}^{\pm\infty} (-\mathbf{i}t)^{\lambda_1} e^{-\mathbf{i}tw^*} \phi_K(h_1 t) \phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W^*}(y,x,\cdot)}(t) dt$$

$$\Psi_{2,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi, y, x, w^*, h_1) = -\frac{1}{2\pi} \frac{\mathbf{i}\theta(\xi)}{(\phi_{W_2}(\xi))^2} \int_{\xi}^{\pm\infty} (-\mathbf{i}t)^{\lambda_1} e^{-\mathbf{i}tw^*} \phi_K(h_1 t) \phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W^*}(y,x,\cdot)}(t) dt$$

$$- \frac{1}{2\pi} (-\mathbf{i}t)^{\lambda_1} e^{-\mathbf{i}\xi w^*} \phi_K(h_1 \xi) \frac{\phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W^*}(y,x,\cdot)}(\xi)}{\phi_{W_2}(\xi)}$$

$$\Psi_{3,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi, y, x, w^*, h_1) = \frac{1}{2\pi} (-\mathbf{i}t)^{\lambda_1} e^{-\mathbf{i}\xi w^*} \phi_K(h_1 \xi) \frac{\phi_{W^*}(\xi)}{\phi_{W_2}(\xi)}.$$

■

Definition 2. *Write $f(t) \preceq g(t)$ for $f, g : \mathbb{R} \to \mathbb{R}$ when there is a constant $C > 0$, independent of t, such that $f(t) \leq Cg(t)$ for all $t \in \mathbb{R}$ (and similarly for $\succeq$). Write $a_n \preceq b_n$ for two sequences a_n, b_n if there exists a constant C independent of n such that $a_n \leq Cb_n$ for all $n \in \mathbb{N}$*

Proof of Theorem 1. By Parseval's identity, I have

$$|g_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y, x, w^*, h) - g_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y, x, w^*)|$$

$$
\begin{aligned}
&= \left| \frac{1}{2\pi} \int (-\mathbf{i}t)^{\lambda_1} e^{-\mathbf{i}tw^*} \left(\phi_K(h_1 t) - 1\right) \frac{\phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(y,x,\cdot)}(t)}{\phi_{W_2}(t)} \phi_{W^*}(t) dt \right| \\
&\leq \frac{1}{2\pi} \int |t|^{\lambda_1} \left|\phi_K(h_1 t) - 1\right| \left| \phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W^*}(y,x,\cdot)}(t) \right| dt \\
&= \frac{1}{2\pi} \int_{|h_1 t| > \bar{\zeta}_K} |t|^{\lambda_1} \left|\phi_K(h_1 t) - 1\right| \left| \phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W^*}(y,x,\cdot)}(t) \right| dt \\
&\preceq \int_{|h_1 t| > \bar{\zeta}_K} |t|^{\lambda_1} \left| \phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W^*}(y,x,\cdot)}(t) \right| dt
\end{aligned}
$$

where I have used Assumption 9 to ensure $\phi_K(t) = 1$ for $|t| \leq \bar{\zeta}_K$ and $\sup_\zeta |\phi_K(\zeta)| < \infty$. Then by Assumption 11 and Lemma 7 in Schennach (2004b),

$$
\begin{aligned}
&\left| g_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h) - g_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*) \right| \\
&\preceq \int_{|h_1 t| > \bar{\zeta}_K} |t|^{\lambda_1} (1+|t|)^{\gamma_\phi} \exp\left(\alpha_\phi |t|^{\beta_\phi}\right) dt \\
&= O\left(\left(\frac{\bar{\zeta}_K}{h_1}\right)^{1+\gamma_\phi+\lambda_1} \exp\left(\alpha_\phi \left(\frac{\bar{\zeta}_K}{h_1}\right)^{\beta_\phi} \right) \right) = O\left(\left(h_1^{-1}\right)^{1+\gamma_\phi+\lambda_1} \exp\left(\alpha_\phi \bar{\zeta}_K^{\beta_\phi} \left(h_1^{-1}\right)^{\beta_\phi} \right) \right).
\end{aligned}
$$

Next I show that the second term in $B_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h)$ is of a smaller order. To do this, I first state a useful representation of $E\left[\hat{\phi}_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W^*}(y,x,\cdot)}(\xi)\right]$

$$
\begin{aligned}
&E\left[\hat{\phi}_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(y,x,\cdot)}(\xi)\right] \\
&= E\left[e^{\mathbf{i}\xi W_2} \frac{1}{h_{2,1}^{1+\lambda_{2,1}} h_{2,2}^{1+\lambda_{2,2}}} G_Y^{(\lambda_{2,1})}\left(\frac{y-Y}{h_{2n,1}}\right) G_X^{(\lambda_{2,2})}\left(\frac{x-X}{h_{2n,2}}\right) \right] \\
&= \int\int e^{\mathbf{i}\xi\tilde{w}} \frac{1}{h_{2,1}^{1+\lambda_{2,1}} h_{2,2}^{1+\lambda_{2,2}}} G_Y^{(\lambda_{2,1})}\left(\frac{y-\tilde{y}}{h_{2n,1}}\right) G_X^{(\lambda_{2,2})}\left(\frac{x-\tilde{x}}{h_{2,2}}\right) f_{Y,X,W_2}(\tilde{y},\tilde{x},\tilde{w}) d\tilde{y} d\tilde{x} d\tilde{w} \\
&= \int \phi_{f^{(0,0)}_{Y,X,W_2}(\tilde{y},\tilde{x},\cdot)}(\xi) \frac{1}{h_{2,1}^{1+\lambda_{2,1}} h_{2,2}^{1+\lambda_{2,2}}} G_Y^{(\lambda_{2,1})}\left(\frac{y-\tilde{y}}{h_{2n,1}}\right) G_X^{(\lambda_{2,2})}\left(\frac{x-\tilde{x}}{h_{2,2}}\right) d\tilde{y} d\tilde{x}
\end{aligned}
$$

If $\lambda_{2,1} = 1, \lambda_{2,2} = 0$,

$$
\begin{aligned}
&E\left[\hat{\phi}_{f^{(1,0)}_{Y,X,W_2}(y,x,\cdot)}(\xi)\right] \\
&= -\int \left(\int \phi_{f^{(0,0)}_{Y,X,W_2}(\tilde{y},\tilde{x},\cdot)}(\xi) \frac{1}{h_{2,1}} dG_Y\left(\frac{y-\tilde{y}}{h_{2,1}}\right) \right) \frac{1}{h_{2,2}} G_X\left(\frac{x-\tilde{x}}{h_{2,2}}\right) d\tilde{x} \\
&= \int \left(-\phi_{f^{(0,0)}_{Y,X,W_2}(\tilde{y},\tilde{x},\cdot)}(\xi) \frac{1}{h_{2,1}} G_Y\left(\frac{\tilde{y}-y}{h_{2,1}}\right) \Big|_{-\infty}^{\infty} \right. \\
&\qquad \left. + \int \frac{1}{h_{2,1}} G_Y\left(\frac{y-\tilde{y}}{h_{2,1}}\right) \phi_{f^{(1,0)}_{Y,X,W_2}(\tilde{y},\tilde{x},\cdot)}(\xi) d\tilde{y} \right) \frac{1}{h_{2,2}} G_X\left(\frac{x-\tilde{x}}{h_{2,2}}\right) d\tilde{x}
\end{aligned}
$$

$$= \int \int \phi_{f^{(1,0)}_{Y,X,W_2}(\tilde{y},\tilde{x},\cdot)}(\xi) \frac{1}{h_{2,1}} G_Y \left(\frac{y-\tilde{y}}{h_{2,1}} \right) \frac{1}{h_{2,2}} G_X \left(\frac{x-\tilde{x}}{h_{2,2}} \right) d\tilde{y} d\tilde{x}$$

For $\lambda_{2,1} = 0, \lambda_{2,2} = 1$, I have similar results, so

$$E\left[\hat{\phi}_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(y,x,\cdot)}(\xi)\right] = \int \int \phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(\tilde{y},\tilde{x},\cdot)}(\xi) \frac{1}{h_{2,1}} G_Y \left(\frac{y-\tilde{y}}{h_{2,1}} \right) \frac{1}{h_{2,2}} G_X \left(\frac{x-\tilde{x}}{h_{2,2}} \right) d\tilde{y} d\tilde{x}.$$

Then note that

$$\begin{aligned}
&\left| E\left[\delta\hat{\phi}_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(y,x,\cdot)}(\xi)\right] \frac{\phi_{W^*}(\xi)}{\phi_{W_2}(\xi)} \right| \\
&= \left| \int \int \phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W^*}(\tilde{y},\tilde{x},\cdot)}(\xi) \frac{1}{h_{2,1}} G_Y \left(\frac{y-\tilde{y}}{h_{2,1}} \right) \frac{1}{h_{2,2}} G_X \left(\frac{x-\tilde{x}}{h_{2,2}} \right) d\tilde{y} d\tilde{x} - \phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W^*}(y,x,\cdot)}(\xi) \right| \\
&= \left| \int \int \phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W^*}(\tilde{y},\tilde{x},\cdot)}(\xi) \left(\frac{1}{(2\pi)^2} \int \int e^{-\mathbf{i}t_1(y-\tilde{y})-\mathbf{i}t_2(x-\tilde{x})} \phi_{G_Y}(t_1 h_{2,1}) \phi_{G_X}(t_2 h_{2,2}) dt_1 dt_2 \right) d\tilde{y} d\tilde{x} \right. \\
&\quad \left. - \frac{1}{(2\pi)^2} \int \int e^{-\mathbf{i}t_1 y - \mathbf{i}t_2 x} \left(\int \int e^{\mathbf{i}t_1 y + \mathbf{i}t_2 x} \phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W^*}(y,x,\cdot)}(\xi) dy dx \right) dt_1 dt_2 \right| \\
&= \left| \frac{1}{(2\pi)^2} \int \int e^{-\mathbf{i}t_1 y - \mathbf{i}t_2 x} \left(\phi_{G_Y}(t_1 h_{2,1}) \phi_{G_X}(t_2 h_{2,2}) - 1 \right) \left(\int \int e^{\mathbf{i}t_1 \tilde{y} + \mathbf{i}t_2 \tilde{x}} \phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W^*}(\tilde{y},\tilde{x},\cdot)}(\xi) d\tilde{y} d\tilde{x} \right) dt_1 dt_2 \right| \\
&= \left| \frac{1}{(2\pi)^2} \int_{\frac{\bar{\xi}_{G_X}}{h_{2,2}}} \int_{\frac{\bar{\xi}_{G_Y}}{h_{2,1}}} e^{-\mathbf{i}t_1 y - \mathbf{i}t_2 x} \phi_{f^{(\lambda_{2,1},\lambda_{2,2})}}(t_1, t_2, \xi) dt_1 dt_2 \right| \\
&\leq \frac{1}{(2\pi)^2} \int_{\frac{\bar{\xi}_{G_X}}{h_{2,2}}} \int_{\frac{\bar{\xi}_{G_Y}}{h_{2,1}}} \left| \phi_{f^{(\lambda_{2,1},\lambda_{2,2})}}(t_1, t_2, \xi) \right| dt_1 dt_2 \\
&\preceq |\phi_{W^*}(\xi)| \int_{\frac{\bar{\xi}_{G_X}}{h_{2,2}}} \int_{\frac{\bar{\xi}_{G_Y}}{h_{2,1}}} (1+|t_1|)^{\gamma_{f_1}} (1+|t_2|)^{\gamma_{f_2}} \exp\left(\alpha_{f_1} |t_1|^{\beta_{f_1}} + \alpha_{f_2} |t_2|^{\beta_{f_2}} \right) dt_1 dt_2 \\
&\preceq |\phi_{W^*}(\xi)| O\left(\left(\frac{\bar{\xi}_{G_Y}}{h_{2,1}} \right)^{1+\gamma_{f_1}} \left(\frac{\bar{\xi}_{G_X}}{h_{2,2}} \right)^{1+\gamma_{f_2}} \exp\left(\alpha_{f_1} \left(\frac{\bar{\xi}_{G_Y}}{h_{2,1}} \right)^{\beta_{f_1}} + \alpha_{f_2} \left(\frac{\bar{\xi}_{G_X}}{h_{2,2}} \right)^{\beta_{f_2}} \right) \right) \\
&\preceq |\phi_{W^*}(\xi)| o(1), && (34)
\end{aligned}$$

so

$$\begin{aligned}
&\left| \frac{1}{2\pi} \int (-\mathbf{i}t)^{\lambda_1} e^{-\mathbf{i}tw^*} \phi_K(h_{1n}t) \frac{E\left[\delta\hat{\phi}_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(v,\cdot)}(t)\right] \phi_{W^*}(t)}{\phi_{W_2}(t)} dt \right| \\
&\preceq o(1) \int_0^{h_{1n}^{-1}} |t|^{\lambda_1} \frac{|\phi_{W^*}(t)|}{|\phi_{W_2}(t)|} |\phi_{W_2}(t)| \, dt \\
&\preceq o(1) \int_0^{h_{1n}^{-1}} (1+|t|)^{\lambda_1+\gamma_\phi-\gamma_2} \exp\left(\alpha_\phi |t|^{\beta_\phi}\right) \exp\left(-\alpha_2 |t|^{\beta_2}\right) dt \\
&\preceq o(1) O\left(\left(h_1^{-1}\right)^{1+\gamma_\phi+\lambda_1-\gamma_2} \exp\left(\alpha_\phi \left(h_1^{-1}\right)^{\beta_\phi} \right) \exp\left(-\alpha_2 \left(h_1^{-1}\right)^{\beta_2} \right) \right) \\
&= o\left(\left(h_1^{-1}\right)^{1+\gamma_\phi+\lambda_1} \exp\left(\alpha_\phi \bar{\zeta}_K^{\beta_\phi} \left(h_1^{-1}\right)^{\beta_\phi} \right) \right).
\end{aligned}$$

This completes the proof. ■

Lemma 14. *Suppose the conditions of Lemma 5 hold. For each ξ and h_1, let*

$\Psi^{+}_{l,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,h_1) \equiv \sup_{(y,x,w^*)\in\mathbb{S}_{(Y,X,W^*)}} |\Psi_{l,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,y,x,w^*,h_1)|$ *for* $l=1,2,3$. *Define*

$$\Psi^{+}_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(h) \equiv \sum_{l=1}^{2}\int \Psi^{+}_{l,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,h_1)d\xi + (h_{2,1}^{-1})^{1+\lambda_{2,1}}(h_{2,2}^{-1})^{1+\lambda_{2,2}}\int \Psi^{+}_{3,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,h_1)d\xi.$$

If Assumption 11 and 13 also hold, then for $h_1>0$

$$\begin{aligned}&\Psi^{+}_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(h)\\ &= O\left(\max\left\{(h_1^{-1})^{1+\gamma_*},(h_{2,1}^{-1})^{1+\lambda_{2,1}}(h_{2,2}^{-1})^{1+\lambda_{2,2}}\right\}(h_1^{-1})^{1-\gamma_2+\gamma_\phi+\lambda_1}\exp\left(\left(\alpha_\phi\mathbf{1}\{\beta_\phi=\beta_2\}-\alpha_2\right)\left(h_1^{-1}\right)^{\beta_2}\right)\right).\end{aligned}$$

Proof.

$$\begin{aligned}&\Psi^{+}_{1,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,h_1)\\ &= \sup_{(y,x,w^*)\in\mathbb{S}_{(Y,X,W^*)}} |\Psi_{1,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,y,x,w^*,h_1)|\\ &\preceq \sup_{(y,x,w^*)\in\mathbb{S}_{(Y,X,W^*)}} \frac{1}{|\phi_{W_2}(\xi)|}\int_{\xi}^{\pm\infty}|t|^{\lambda_1}\left|e^{-\mathbf{i}tw^*}\right||\phi_K(h_1t)|\left(\sup_{(y,x)\in\mathbb{S}_{(Y,X)}}\left|\phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W^*}(y,x,\cdot)}(t)\right|\right)dt\\ &\preceq \frac{1}{|\phi_{W_2}(\xi)|}\int_{\xi}^{\pm\infty}|t|^{\lambda_1}|\phi_K(h_1t)|\left(\sup_{(y,x)\in\mathbb{S}_{(Y,X)}}\left|\phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W^*}(y,x,\cdot)}(t)\right|\right)dt\end{aligned}$$

so that

$$\begin{aligned}\int \Psi^{+}_{1,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,h_1)d\xi &\preceq \int\left[\frac{1}{|\phi_{W_2}(\xi)|}\mathbf{1}\{|\xi|\le h_1^{-1}\}\int_{\xi}^{h_1^{-1}}|t|^{\lambda_1}\left|\phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W^*}(y,x,\cdot)}(t)\right|dt\right]d\xi\\ &\preceq \int\left[\frac{1}{|\phi_{W_2}(\xi)|}\mathbf{1}\{|\xi|\le h_1^{-1}\}\int_{\xi}^{h_1^{-1}}|t|^{\lambda_1}(1+|t|)^{\gamma_\phi}\exp\left(\alpha_\phi|t|^{\beta_\phi}\right)dt\right]d\xi\\ &\preceq (h_1^{-1})^{2-\gamma_2+\gamma_\phi+\lambda_1}\exp\left(-\alpha_2\left(h_1^{-1}\right)^{\beta_2}\right)\exp\left(\alpha_\phi\left(h_1^{-1}\right)^{\beta_\phi}\right)\end{aligned}$$

where in the third $\preceq$, I used the assumption $1-\gamma_2+\gamma_\phi+\lambda_1>0$ when $\beta_2=0$ and invoked Lemma 7 and Lemma 8 in Schennach (2004b).

$$\begin{aligned}&\Psi^{+}_{2,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,h_1)\\ &= \sup_{(y,x,w^*)\in\mathbb{S}_{(Y,X,W^*)}}\left|\Psi_{2,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,y,x,w^*,h_1)\right|\\ &\preceq \sup_{(y,x,w^*)\in\mathbb{S}_{(Y,X,W^*)}}\frac{|\theta(\xi)|}{|\phi_{W_2}(\xi)|^2}\int_{\xi}^{\pm\infty}|t|^{\lambda_1}|e^{-\mathbf{i}tw^*}||\phi_K(h_1t)|\left|\phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W^*}(y,x,\cdot)}(t)\right|dt\\ &\quad+\sup_{(y,x,w^*)\in\mathbb{S}_{(Y,X,W^*)}}|\xi|^{\lambda_1}|e^{-\mathbf{i}\xi w^*}||\phi_K(h_1\xi)|\frac{\left|\phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W^*}(y,x,\cdot)}(\xi)\right|}{|\phi_{W_2}(\xi)|}\\ &\preceq \sup_{(y,x)\in\mathbb{S}_{(Y,X)}}\left\{\frac{|\theta(\xi)|}{|\phi_{W_2}(\xi)|^2}\int_{\xi}^{\pm\infty}|t|^{\lambda_1}|\phi_K(h_1t)|\left|\phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W^*}(y,x,\cdot)}(t)\right|dt\right.\end{aligned}$$

$$+|\xi|^{\lambda_1}|\phi_K(h_1\xi)|\frac{\left|\phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W^*}(y,x,\cdot)}(t)\right|}{|\phi_{W_2}(\xi)|}\Bigg\}$$

$$\begin{aligned}
&\preceq \frac{1}{|\phi_{W_2}(\xi)|}\left(\frac{|\theta(\xi)|}{|\phi_{W_2}(\xi)|}\int_{\xi}^{\pm\infty}|t|^{\lambda_1}|\phi_K(h_1t)|\left(\sup_{(y,x)\in\mathbb{S}_{(Y,X)}}\left|\phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W^*}(y,x,\cdot)}(t)\right|\right)dt\right.\\
&\quad\left.+|\xi|^{\lambda_1}|\phi_K(h_1\xi)|\left(\sup_{(y,x)\in\mathbb{S}_{(Y,X)}}\left|\phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W^*}(y,x,\cdot)}(\xi)\right|\right)\right)\\
&\preceq \frac{1}{|\phi_{W_2}(\xi)|}\left(\frac{|\phi'_{W^*}(\xi)|}{|\phi_{W^*}(\xi)|}\int_{\xi}^{\pm\infty}|t|^{\lambda_1}|\phi_K(h_1t)|\left(\sup_{(y,x)\in\mathbb{S}_{(Y,X)}}\left|\phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W^*}(y,x,\cdot)}(t)\right|\right)dt\right.\\
&\quad\left.+|\xi|^{\lambda_1}|\phi_K(h_1\xi)|\left(\sup_{(y,x)\in\mathbb{S}_{(Y,X)}}\left|\phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W^*}(y,x,\cdot)}(\xi)\right|\right)\right)
\end{aligned}$$

where I used the fact that

$$\begin{aligned}
\frac{\theta(\xi)}{\phi_{W_2}(\xi)} &= \frac{E\left[W_1e^{\mathrm{i}\xi W_2}\right]}{E\left[e^{\mathrm{i}\xi W_2}\right]} = \frac{E\left[\left(W^*+\Delta W_1\right)e^{\mathrm{i}\xi(W^*+\Delta W_2)}\right]}{E\left[e^{\mathrm{i}\xi(W^*+\Delta W_2)}\right]}\\
&= \frac{E\left[W^*e^{\mathrm{i}\xi(W^*+\Delta W_2)}\right]+E\left[E\left[\Delta W_1 \mid W^*,\Delta W_2\right]e^{\mathrm{i}\xi(W^*+\Delta W_2)}\right]}{E\left[e^{\mathrm{i}\xi(W^*+\Delta_2)}\right]}\\
&= \frac{E\left[W^*e^{\mathrm{i}\xi(W^*+\Delta W_2)}\right]}{E\left[e^{\mathrm{i}\xi(W^*+\Delta W_2)}\right]} = \frac{E\left[W^*e^{\mathrm{i}\xi W^*}\right]}{E\left[e^{\mathrm{i}\xi W^*}\right]}\\
&= \frac{-\mathbf{i}(d/d\xi)E\left[e^{\mathrm{i}\xi W^*}\right]}{E\left[e^{\mathrm{i}\xi W^*}\right]} = -\mathbf{i}\frac{(d/d\xi)\phi_{W^*}(\xi)}{\phi_{W^*}(\xi)}.
\end{aligned}$$

Integrating $\Psi^+_{2,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,h_1)$ with respect to ξ and using Assumption 11 and 13 gives

$$\begin{aligned}
&\int\Psi^+_{2,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,h_1)d\xi\\
&\preceq \int\frac{1}{|\phi_{W_2}(\xi)|}\left(\frac{|\phi'_{W^*}(\xi)|}{|\phi_{W^*}(\xi)|}\int_{\xi}^{\pm\infty}\mathbf{1}\left\{|t|\le h_1^{-1}\right\}|t|^{\lambda_1}\left(\sup_{(y,x)\in\mathbb{S}_{(Y,X)}}\left|\phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W^*}(y,x,\cdot)}(t)\right|\right)dt\right.\\
&\quad\left.+\mathbf{1}\left\{|\xi|\le h_1^{-1}\right\}|\xi|^{\lambda_1}\left(\sup_{(y,x)\in\mathbb{S}_{(Y,X)}}\left|\phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W^*}(y,x,\cdot)}(\xi)\right|\right)\right)d\xi\\
&\preceq \int\frac{1}{|\phi_{W_2}(\xi)|}\left(\frac{|\phi'_{W^*}(\xi)|}{|\phi_{W^*}(\xi)|}\mathbf{1}\left\{|\xi|\le h_1^{-1}\right\}\int_{|\xi|}^{h_1^{-1}}|t|^{\lambda_1}\left(\sup_{(y,x)\in\mathbb{S}_{(Y,X)}}\left|\phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W^*}(y,x,\cdot)}(t)\right|\right)dt\right.\\
&\quad\left.+\mathbf{1}\left\{|\xi|\le h_1^{-1}\right\}|\xi|^{\lambda_1}\left(\sup_{(y,x)\in\mathbb{S}_{(Y,X)}}\left|\phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W^*}(y,x,\cdot)}(\xi)\right|\right)\right)d\xi\\
&\preceq \int(1+|\xi|)^{-\gamma_2}\exp(-\alpha_2|\xi|^{\beta_2})\mathbf{1}\left\{|\xi|\le h_1^{-1}\right\}\\
&\quad\times\left((1+|\xi|)^{\gamma_*}\int_0^{h_1^{-1}}|t|^{\lambda_1}(1+|t|)^{\gamma_\phi}\exp\left(\alpha_\phi|t|^{\beta_\phi}\right)dt+|\xi|^{\lambda_1}(1+|\xi|)^{\gamma_\phi}\exp\left(\alpha_\phi|\xi|^{\beta_\phi}\right)\right)d\xi
\end{aligned}$$

$$
\begin{aligned}
&\preceq (1+h_1^{-1})^{1-\gamma_2}\exp\left(-\alpha_2\left(h_1^{-1}\right)^{\beta_2}\right)\\
&\quad\times\left((1+h_1^{-1})^{\gamma_*}(1+h_1^{-1})^{1+\gamma_\phi+\lambda_1}\exp\left(\alpha_\phi\left(h_1^{-1}\right)^{\beta_\phi}\right)+(1+h_1^{-1})^{\gamma_\phi+\lambda_1}\exp\left(\alpha_\phi\left(h_1^{-1}\right)^{\beta_\phi}\right)\right)\\
&\preceq (h_1^{-1})^{2-\gamma_2+\gamma_\phi+\gamma_*+\lambda_1}\exp\left(-\alpha_2\left(h_1^{-1}\right)^{\beta_2}\right)\exp\left(\alpha_\phi\left(h_1^{-1}\right)^{\beta_\phi}\right)
\end{aligned}
$$

$$
\begin{aligned}
\Psi^+_{3,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,h_1) &= \sup_{(y,x,w^*)\in\mathbb{S}_{(Y,X,W^*)}}\left|\Psi_{3,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,y,x,w^*,h_1)\right|\\
&\preceq \sup_{(y,x,w^*)\in\mathbb{S}_{(Y,X,W^*)}}\frac{|\phi_{W^*}(\xi)|}{|\phi_{W_2}(\xi)|}|\xi|^{\lambda_1}|e^{-\mathbf{i}\xi w^*}||\phi_K(h_1\xi)|\\
&= \frac{|\phi_{W^*}(\xi)|}{|\phi_{W_2}(\xi)|}|\xi|^{\lambda_1}|\phi_K(h_1\xi)|.
\end{aligned}
$$

so that

$$
\begin{aligned}
&(h_{2,1}^{-1})^{1+\lambda_{2,1}}(h_{2,2}^{-1})^{1+\lambda_{2,2}}\int\Psi^+_{3,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,h_1)d\xi\\
&\preceq (h_{2,1}^{-1})^{1+\lambda_{2,1}}(h_{2,2}^{-1})^{1+\lambda_{2,2}}\int\frac{|\phi_{W^*}(\xi)|}{|\phi_{W_2}(\xi)|}|\xi|^{\lambda_1}\mathbf{1}\left\{|\xi|\le h_1^{-1}\right\}d\xi\\
&\preceq (h_{2,1}^{-1})^{1+\lambda_{2,1}}(h_{2,2}^{-1})^{1+\lambda_{2,2}}\int_0^{h_1^{-1}}\frac{|\phi_{W^*}(\xi)|}{|\phi_{W_2}(\xi)|}|\xi|^{\lambda_1}d\xi\\
&\preceq (h_{2,1}^{-1})^{1+\lambda_{2,1}}(h_{2,2}^{-1})^{1+\lambda_{2,2}}(h_1^{-1})^{1-\gamma_2+\gamma_\phi+\lambda_1}\exp\left(-\alpha_2\left(h_1^{-1}\right)^{\beta_2}\right)\exp\left(\alpha_\phi\left(h_1^{-1}\right)^{\beta_\phi}\right)
\end{aligned}
$$

Collecting together these rates delivers the conclusion of the lemma. ■

Lemma 15. *Suppose the conditions for Lemma 5 hold, then*

$$
\begin{aligned}
\sup_{(y,x)\in\mathbb{S}_{(Y,X)}}\Bigg|\hat{E}\left[e^{\mathbf{i}\xi W_2}\frac{1}{h_{2,1}^{1+\lambda_{2,1}}h_{2,2}^{1+\lambda_{2,2}}}G_Y^{(\lambda_{2,1})}\left(\frac{y-Y}{h_{2,1}}\right)G_X^{(\lambda_{2,2})}\left(\frac{x-X}{h_{2,2}}\right)\right]&\\
-E\left[e^{\mathbf{i}\xi W_2}\frac{1}{h_{2,1}^{1+\lambda_{2,1}}h_{2,2}^{1+\lambda_{2,2}}}G_Y^{(\lambda_{2,1})}\left(\frac{y-Y}{h_{2,1}}\right)G_X^{(\lambda_{2,2})}\left(\frac{x-X}{h_{2,2}}\right)\right]\Bigg| &\preceq \frac{1}{h_{2,1}^{1+\lambda_{2,1}}h_{2,2}^{1+\lambda_{2,2}}\sqrt{n}}.
\end{aligned}
$$

Proof. Denote the Fourier transform of $f_{Y,X,W_2}(y,x,w^*)$ as $\phi_{f_2}(t_1,t_2,\xi)$.

$$
\begin{aligned}
&E\left[e^{\mathbf{i}\xi W_2}\frac{1}{h_{2,1}^{1+\lambda_{2,1}}h_{2,2}^{1+\lambda_{2,2}}}G_Y^{(\lambda_{2,1})}\left(\frac{y-Y}{h_{2,1}}\right)G_X^{(\lambda_{2,2})}\left(\frac{x-X}{h_{2,2}}\right)\right]\\
&=\int\int\int e^{\mathbf{i}\xi\tilde{w}}\frac{1}{h_{2,1}^{1+\lambda_{2,1}}h_{2,2}^{1+\lambda_{2,2}}}G_Y^{(\lambda_{2,1})}\left(\frac{y-\tilde{y}}{h_{2,1}}\right)G_X^{(\lambda_{2,2})}\left(\frac{x-\tilde{x}}{h_{2,2}}\right)f_{Y,X,W_2}(\tilde{y},\tilde{x},\tilde{w})d\tilde{w}d\tilde{y}d\tilde{x}\\
&=\int\int\int e^{\mathbf{i}\xi\tilde{w}}\left(\frac{1}{(2\pi)^2}\int\int e^{-\mathbf{i}t_1(x-\tilde{x})-\mathbf{i}t_2(y-\tilde{y})}(\mathbf{i}t_1)^{\lambda_{2,1}}(\mathbf{i}t_2)^{\lambda_{2,2}}\phi_{G_X}(t_1h_{2,1})\phi_{G_Y}(t_2h_{2,2})dt_1dt_2\right)\times\\
&\qquad\qquad f_{Y,X,W_2}(\tilde{y},\tilde{x},\tilde{w})d\tilde{w}d\tilde{y}d\tilde{x}
\end{aligned}
$$

$$= \frac{1}{(2\pi)^2}\int\int (\mathbf{i}t_1)^{\lambda_{2,1}}(\mathbf{i}t_2)^{\lambda_{2,2}}e^{-\mathbf{i}t_1y-\mathbf{i}t_2x}\phi_{G_Y}(t_1h_{2,1})\phi_{G_X}(t_2h_{2,2})\phi_{f_2}(t_1,t_2,\xi)dt_1dt_2$$

so

$$\sup_{(y,x)\in\mathbb{S}_{(Y,X)}}\left|\hat{E}\left[e^{\mathbf{i}\xi W_2}\frac{1}{h_{2,1}^{1+\lambda_{2,1}}h_{2,2}^{1+\lambda_{2,2}}}G_Y^{(\lambda_{2,1})}\left(\frac{y-Y}{h_{2,1}}\right)G_X^{(\lambda_{2,2})}\left(\frac{x-X}{h_{2,2}}\right)\right]\right.$$
$$\left.-E\left[e^{\mathbf{i}\xi W_2}\frac{1}{h_{2,1}^{1+\lambda_{2,1}}h_{2,2}^{1+\lambda_{2,2}}}G_Y^{(\lambda_{2,1})}\left(\frac{y-Y}{h_{2,1}}\right)G_X^{(\lambda_{2,2})}\left(\frac{x-X}{h_{2,2}}\right)\right]\right|$$
$$= \sup_{(y,x)\in\mathbb{S}_{(Y,X)}}\left|\frac{1}{(2\pi)^2}\int\int (\mathbf{i}t_1)^{\lambda_{2,1}}(\mathbf{i}t_2)^{\lambda_{2,2}}e^{-\mathbf{i}t_1y-\mathbf{i}t_2x}\phi_{G_Y}(t_1h_{2,1})\phi_{G_X}(t_2h_{2,2})\times\right.$$
$$\left.\left(\frac{1}{n}\sum_{j=1}^n e^{\mathbf{i}t_1Y_j+\mathbf{i}t_2X_j+\mathbf{i}\xi W_{2,j}}-\phi_{f_2}(t_1,t_2,\xi)\right)dt_1dt_2\right|$$
$$\leq \frac{1}{(2\pi)^2}\int\int |t_1|^{\lambda_{2,1}}|t_2|^{\lambda_{2,2}}\left|\phi_{G_Y}(t_1h_{2,1})\phi_{G_X}(t_2h_{2,2})\right|\left|\frac{1}{n}\sum_{j=1}^n e^{\mathbf{i}t_1Y_j+\mathbf{i}t_2X_j+\mathbf{i}\xi W_{2,j}}-\phi_{f_2}(t_1,t_2,\xi)\right|dt_2.$$

Then since

$$E\left[\left(\frac{1}{n}\sum_{j=1}^n e^{\mathbf{i}t_1Y_j+\mathbf{i}t_2X_j+\mathbf{i}\xi W_{2,j}}-\phi_{f_2}(t_1,t_2,\xi)\right)^2\right]$$
$$= \frac{1}{n}E\left[\left(e^{\mathbf{i}t_1Y_j+\mathbf{i}t_2X_j+\mathbf{i}\xi W_{2,j}}-\phi_{f_2}(t_1,t_2,\xi)\right)^2\right]$$
$$\leq \frac{1}{n}E\left[\left(e^{\mathbf{i}t_1Y_j+\mathbf{i}t_2X_j+\mathbf{i}\xi W_{2,j}}\right)^2\right] = O\left(\frac{1}{n}\right),$$

I have

$$\sup_{(y,x)\in\mathbb{S}_{(Y,X)}}\left|\hat{E}\left[e^{\mathbf{i}\xi W_2}\frac{1}{h_{2,1}^{1+\lambda_{2,1}}h_{2,2}^{1+\lambda_{2,2}}}G_Y^{(\lambda_{2,1})}\left(\frac{y-Y}{h_{2,1}}\right)G_X^{(\lambda_{2,2})}\left(\frac{x-X}{h_{2,2}}\right)\right]\right.$$
$$\left.-E\left[e^{\mathbf{i}\xi W_2}\frac{1}{h_{2,1}^{1+\lambda_{2,1}}h_{2,2}^{1+\lambda_{2,2}}}G_Y^{(\lambda_{2,1})}\left(\frac{y-Y}{h_{2,1}}\right)G_X^{(\lambda_{2,2})}\left(\frac{x-X}{h_{2,2}}\right)\right]\right|$$
$$\preceq \frac{1}{\sqrt{n}}\int\int |t_1|^{\lambda_{2,1}}|t_2|^{\lambda_{2,2}}\left|\phi_{G_Y}(t_1h_{2,1})\phi_{G_X}(t_2h_{2,2})\right|dt_1dt_2$$
$$\preceq \frac{1}{\sqrt{n}}\int_0^{h_{2,2}^{-1}}\int_0^{h_{2,1}^{-1}}|t_1|^{\lambda_{2,1}}|t_2|^{\lambda_{2,2}}dt_1dt_2 \preceq \frac{1}{h_{2,1}^{1+\lambda_{2,1}}h_{2,2}^{1+\lambda_{2,2}}\sqrt{n}}.$$

■

Lemma 16. *For a finite integer J, let $\{P_{n,j}(a)\}$ be a sequence of nonrandom real-valued continuously differentiable functions of random vectors a, for $j = 1,...,J$ respectively. Let A be a random vector. If for each n, $\sigma_n = \sqrt{var\left(\sum_{j=1}^J P_{n,j}(A)\right)}$ exist, and $\sigma_n > 0$ for n sufficiently large, then*

$$\sigma_n^{-1}n^{1/2}\left(\hat{E}\left[\sum_{j=1}^J P_{n,j}(A)\right]\right) \xrightarrow{d} N(0,1).$$

Proof. Denote the centered version of the summands as $Z_{n,j}(A) \equiv P_{n,j}(A) - E\left[P_{n,j}(A)\right]$. Then $\sigma_n^2 \equiv E\left[\left(\sum_{j=1}^{J} Z_{n,j}(A)\right)^2\right]$ It's sufficient to verify that the Lindeberg condition holds:

$$\sum_{i=1}^{n} E\left[\mathbf{1}\left\{\left|\frac{\sum_{j=1}^{J} Z_{n,j}(A)}{\sqrt{n}\sigma_n}\right| > \epsilon\right\}\left|\frac{\sum_{j=1}^{J} Z_{n,j}(A)}{\sqrt{n}\sigma_n}\right|^2\right] \to 0, \text{ as } n \to \infty,$$

for all $\epsilon > 0$. Note that as $n \to \infty$

$$E\left[\left|\frac{\sum_{j=1}^{J} Z_{n,j}(A)}{\sqrt{n}\sigma_n}\right|^2\right] = \frac{1}{n} E\left[\frac{\left|\sum_{j=1}^{J} Z_{n,j}(A)\right|^2}{E\left[\left(\sum_{j=1}^{J} Z_{n,j}(A)\right)^2\right]}\right] \to 0$$

By Markov Inequality, for any $\epsilon > 0$, as $n \to \infty$, $Pr\left(\left|\frac{\sum_{j=1}^{J} Z_{n,j}(A)}{\sqrt{n}\sigma_n}\right| > \epsilon\right) \to 0$, which implies $\mathbf{1}\left\{\left|\frac{\sum_{j=1}^{J} Z_{n,j}(A)}{\sqrt{n}\sigma_n}\right| > \epsilon\right\} = o_p(1)$. Note that $\mathbf{1}\left\{\left|\frac{\sum_{j=1}^{J} Z_{n,j}(A)}{\sqrt{n}\sigma_n}\right| > \epsilon\right\}\left|\frac{\sum_{j=1}^{J} Z_{n,j}(A)}{\sigma_n}\right|^2 \leq \left|\frac{\sum_{j=1}^{J} Z_{n,j}(A)}{\sigma_n}\right|^2$ and that $E\left[\left|\frac{\sum_{j=1}^{J} Z_{n,j}(A)}{\sigma_n}\right|^2\right] = 1 < \infty$. By Dominated Convergence Theorem

$$E\left[\mathbf{1}\left\{\left|\frac{\sum_{j=1}^{J} Z_{n,j}(A)}{\sqrt{n}\sigma_n}\right| > \epsilon\right\}\left|\frac{\sum_{j=1}^{J} Z_{n,j}(A)}{\sigma_n}\right|^2\right] \to 0, \text{ as } n \to \infty.$$

The conclusion follows. ■

Proof of Theorem 2. (i) The fact that $E[L_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y, x, w^*, h)] = 0$ follows directly from Eq. (33). Next, with fixed value of h, Assumption 9 and 12 ensures the existence and finiteness of

$$\begin{aligned}
E[L^2_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y, x, w^*, h)] &= E\left[\left(\hat{E}\left[l_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y, x, w^*, h; Y, X, W_1, W_2)\right]\right)^2\right] \\
&= n^{-1} E\left[\left(l_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y, x, w^*, h; Y, X, W_1, W_2)\right)^2\right] \\
&= n^{-1} \Omega_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y, x, w^*, h).
\end{aligned}$$

From Eq. (33), I have

$$\begin{aligned}
&\Omega_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y, x, w^*, h) \\
&\equiv E\left[n\left(\bar{g}_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y, x, w^*, h) - g_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y, x, w^*, h)\right)^2\right]
\end{aligned}$$

$$
\begin{aligned}
&= E\Bigg[\Bigg(\int \Psi_{1,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,y,x,w^*,h_1) n^{1/2}\delta\hat{\theta}(\xi)d\xi \\
&\qquad + \int \Psi_{2,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,y,x,w^*,h_1) n^{1/2}\delta\hat{\phi}_{W_2}(\xi)d\xi \\
&\qquad + \int \Psi_{3,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,y,x,w^*,h_1) n^{1/2}\delta\hat{\phi}_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(y,x,\cdot)}(\xi)d\xi\Bigg)^2\Bigg] \\
&= \int\int \Psi_{1,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,y,x,w^*,h_1) E\left[n\delta\hat{\theta}(\xi)\delta\hat{\theta}^\dagger(\zeta)\right]\left(\Psi_{1,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\zeta,y,x,w^*,h_1)\right)^\dagger d\xi d\zeta \\
&\quad + \int\int \Psi_{2,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,y,x,w^*,h_1) E\left[n\delta\hat{\phi}_{W_2}(\xi)\delta\hat{\phi}^\dagger_{W_2}(\zeta)\right]\left(\Psi_{2,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\zeta,y,x,w^*,h_1)\right)^\dagger d\xi d\zeta \\
&\quad + \int\int (h_{2,1}^{-2})^{1+\lambda_{2,1}}(h_{2,2}^{-2})^{1+\lambda_{2,2}}\Psi_{3,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,y,x,w^*,h_1)\times \\
&\qquad\qquad E\left[n h_{2,1}^{2+2\lambda_{2,1}} h_{2,2}^{2+2\lambda_{2,2}}\delta\hat{\phi}_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(y,x,\cdot)}(\xi)\delta\hat{\phi}^\dagger_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(y,x,\cdot)}(\zeta)\right]\times \\
&\qquad\qquad \left(\Psi_{3,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\zeta,y,x,w^*,h_1)\right)^\dagger d\xi d\zeta \\
&\quad + \int\int \Psi_{1,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,y,x,w^*,h_1) E\left[n\delta\hat{\theta}(\xi)\delta\hat{\phi}^\dagger_{W_2}(\zeta)\right]\left(\Psi_{2,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\zeta,y,x,w^*,h_1)\right)^\dagger d\xi d\zeta \\
&\quad + \int\int \Psi_{1,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,y,x,w^*,h_1) E\left[n\delta\hat{\theta}(\xi) h_{2,1}^{1+\lambda_{2,1}} h_{2,2}^{1+\lambda_{2,2}}\delta\hat{\phi}^\dagger_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(y,x,\cdot)}(\zeta)\right]\times \\
&\qquad\qquad\qquad (h_{2,1}^{-1})^{1+\lambda_{2,1}}(h_{2,2}^{-1})^{1+\lambda_{2,2}}\left(\Psi_{3,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\zeta,y,x,w^*,h_1)\right)^\dagger d\xi d\zeta \\
&\quad + \int\int \Psi_{2,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,y,x,w^*,h_1) E\left[n\delta\hat{\phi}_{W_2}(\xi) h_{2,1}^{1+\lambda_{2,1}} h_{2,2}^{1+\lambda_{2,2}}\delta\hat{\phi}^\dagger_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(y,x,\cdot)}(\zeta)\right]\times \\
&\qquad\qquad\qquad (h_{2,1}^{-1})^{1+\lambda_{2,1}}(h_{2,2}^{-1})^{1+\lambda_{2,2}}\left(\Psi_{3,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\zeta,y,x,w^*,h_1)\right)^\dagger d\xi d\zeta \\
&\quad + \int\int \Psi_{2,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,y,x,w^*,h_1) E\left[n\delta\hat{\phi}_{W_2}(\zeta)\delta\hat{\theta}^\dagger(\xi)\right]\left(\Psi_{1,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\zeta,y,x,w^*,h_1)\right)^\dagger d\xi d\zeta \\
&\quad + \int\int (h_{2,1}^{-1})^{1+\lambda_{2,1}}(h_{2,2}^{-1})^{1+\lambda_{2,2}}\Psi_{3,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\zeta,y,x,w^*,h_1)\times \\
&\qquad E\left[n h_{2,1}^{1+\lambda_{2,1}} h_{2,2}^{1+\lambda_{2,2}}\delta\hat{\phi}_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(y,x,\cdot)}(\zeta)\delta\hat{\theta}^\dagger(\xi)\right]\times\left(\Psi_{1,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,y,x,w^*,h_1)\right)^\dagger d\xi d\zeta \\
&\quad + \int\int (h_{2,1}^{-1})^{1+\lambda_{2,1}}(h_{2,2}^{-1})^{1+\lambda_{2,2}}\Psi_{3,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\zeta,y,x,w^*,h_1)\times \\
&\qquad E\left[n h_{2,1}^{1+\lambda_{2,1}} h_{2,2}^{1+\lambda_{2,2}}\delta\phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(y,x,\cdot)}(\zeta)\delta\hat{\phi}^\dagger_{W_2}(\xi)\right]\times\left(\Psi_{2,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,y,x,w^*,h_1)\right)^\dagger d\xi d\zeta \qquad (35)
\end{aligned}
$$

Note that I have

$$
\begin{aligned}
E\left[n\delta\hat{\theta}(\xi)\delta\hat{\theta}^\dagger(\zeta)\right] &= E\left[n\left(\hat{\theta}(\xi)-\theta(\xi)\right)\left(\hat{\theta}^\dagger(\zeta)-\theta^\dagger(\zeta)\right)\right] \\
&= E\left[\left(W_1 e^{\mathbf{i}\xi W_2}-\theta(\xi)\right)\left(W_1 e^{-\mathbf{i}\zeta W_2}-\theta(-\zeta)\right)\right] \\
&= E\left[W_1^2 e^{\mathbf{i}(\xi-\zeta)W_2}\right]-\theta(\xi)E\left[W_1 e^{-\mathbf{i}\zeta W_2}\right]-E\left[W_1 e^{\mathbf{i}\xi W_2}\right]\theta^\dagger(\zeta)+\theta(\xi)\theta^\dagger(\zeta) \\
&= E\left[W_1^2 e^{\mathbf{i}(\xi-\zeta)W_2}\right]-\theta(\xi)\theta(-\zeta),
\end{aligned}
$$

$$
\begin{aligned}
E\left[n\delta\hat{\theta}(\xi)\delta\hat{\phi}^\dagger_{W_2}(\zeta)\right] &= E\left[n\left(\hat{\theta}(\xi)-\theta(\xi)\right)\left(\hat{\phi}^\dagger_{W_2}(\zeta)-\phi^\dagger_{W_2}(\zeta)\right)\right] \\
&= E\left[\left(W_1 e^{\mathbf{i}\xi W_2}-\theta(\xi)\right)\left(e^{-\mathbf{i}\zeta W_2}-\phi^\dagger_{W_2}(\zeta)\right)\right]
\end{aligned}
$$

$$= E\left[W_1 e^{\mathbf{i}\xi W_2} e^{-\mathbf{i}\zeta W_2}\right] - \theta(\xi) E\left[e^{-\mathbf{i}\zeta W_2}\right] - E\left[W_1 e^{\mathbf{i}\xi W_2}\right] \phi^{\dagger}_{W_2}(\zeta) + \theta(\xi)\phi^{\dagger}_{W_2}(\zeta)$$
$$= E\left[W_1 e^{\mathbf{i}(\xi-\zeta) W_2}\right] - \theta(\xi)\phi_{W_2}(-\zeta),$$

$$E\left[n\delta\hat{\theta}(\xi) h_{2,1}^{1+\lambda_{2,1}} h_{2,2}^{1+\lambda_{2,2}} \delta\hat{\phi}^{\dagger}_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(y,x,\cdot)}(\zeta)\right]$$
$$= E\Big[\left(W_1 e^{\mathbf{i}\xi W_2} - \theta(\xi)\right) \times$$
$$\left(e^{-\mathbf{i}\zeta W_2} G_Y\left(\frac{y-Y}{h_{2,1}}\right) G_X^{(\lambda_{2,1},\lambda_{2,2})}\left(\frac{x-X}{h_{2,2}}\right) - h_{2,1}^{1+\lambda_{2,1}} h_{2,2}^{1+\lambda_{2,2}} \phi^{\dagger}_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(y,x,\cdot)}(\zeta)\right)\Big]$$
$$= E\left[W_1 e^{\mathbf{i}(\xi-\zeta) W_2} G_Y\left(\frac{y-Y}{h_{2,1}}\right) G_X^{(\lambda_{2,1},\lambda_{2,2})}\left(\frac{x-X}{h_{2,2}}\right)\right] - \theta(\xi) h_{2,1}^{1+\lambda_{2,1}} h_{2,2}^{1+\lambda_{2,2}} \phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(y,x,\cdot)}(-\zeta),$$

$$E\left[n\delta\hat{\phi}_{W_2}(\xi)\delta\hat{\phi}^{\dagger}_{W_2}(\zeta)\right] = E\left[\left(e^{\mathbf{i}\xi W_2} - \phi_{W_2}(\xi)\right)\left(e^{-\mathbf{i}\zeta W_2} - \phi^{\dagger}_{W_2}(\zeta)\right)\right]$$
$$= E\left[e^{\mathbf{i}(\xi-\zeta) W_2}\right] - \phi_{W_2}(\xi)\phi_{W_2}(-\zeta),$$

$$E\left[n\delta\hat{\phi}_{W_2}(\xi) h_{2,1}^{1+\lambda_{2,1}} h_{2,2}^{1+\lambda_{2,2}} \delta\hat{\phi}^{\dagger}_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(y,x,\cdot)}(\zeta)\right]$$
$$= E\Big[\left(e^{\mathbf{i}\xi W_2} - \phi_{W_2}(\xi)\right) \times$$
$$\left(e^{-\mathbf{i}\zeta W_2} G_Y\left(\frac{y-Y}{h_{2,1}}\right) G_X^{(\lambda_{2,1},\lambda_{2,2})}\left(\frac{x-X}{h_{2,2}}\right) - h_{2,1}^{1+\lambda_{2,1}} h_{2,2}^{1+\lambda_{2,2}} \phi^{\dagger}_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(y,x,\cdot)}(\zeta)\right)\Big]$$
$$= E\left[e^{\mathbf{i}(\xi-\zeta) W_2} G_Y\left(\frac{y-Y}{h_{2,1}}\right) G_X^{(\lambda_{2,1},\lambda_{2,2})}\left(\frac{x-X}{h_{2,2}}\right)\right] - \phi_{W_2}(\xi) h_{2,1}^{1+\lambda_{2,1}} h_{2,2}^{1+\lambda_{2,2}} \phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(y,x,\cdot)}(-\zeta),$$

$$E\left[n h_{2,1}^{1+\lambda_{2,1}} h_{2,2}^{1+\lambda_{2,2}} \delta\hat{\phi}_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(y,x,\cdot)}(\xi) h_{2,1}^{1+\lambda_{2,1}} h_{2,2}^{1+\lambda_{2,2}} \delta\hat{\phi}^{\dagger}_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(y,x,\cdot)}(\zeta)\right]$$
$$= E\Big[\left(e^{\mathbf{i}\xi W_2} G_Y\left(\frac{y-Y}{h_{2,1}}\right) G_X^{(\lambda_{2,1},\lambda_{2,2})}\left(\frac{x-X}{h_{2,2}}\right) - h_{2,1}^{1+\lambda_{2,1}} h_{2,2}^{1+\lambda_{2,2}} \phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(y,x,\cdot)}(\xi)\right) \times$$
$$\left(e^{-\mathbf{i}\zeta W_2} G_Y\left(\frac{y-Y}{h_{2,1}}\right) G_X^{(\lambda_{2,1},\lambda_{2,2})}\left(\frac{x-X}{h_{2,2}}\right) - h_{2,1}^{1+\lambda_{2,1}} h_{2,2}^{1+\lambda_{2,2}} \phi^{\dagger}_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(y,x,\cdot)}(\zeta)\right)\Big]$$
$$= E\left[e^{\mathbf{i}(\xi-\zeta) W_2}\left(G_Y\left(\frac{y-Y}{h_{2,1}}\right) G_X^{(\lambda_{2,1},\lambda_{2,2})}\left(\frac{x-X}{h_{2,2}}\right)\right)^2\right]$$
$$- h_{2,1}^{2+2\lambda_{2,1}} h_{2,2}^{2+2\lambda_{2,2}} \phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(y,x,\cdot)}(\xi) \phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(y,x,\cdot)}(-\zeta).$$

By Assumption 12,

$$\left|E\left[n\delta\hat{\theta}(\xi)\delta\hat{\theta}^{\dagger}(\zeta)\right]\right| = \left|E\left[W_1^2 e^{\mathbf{i}(\xi-\zeta) W_2}\right] - \theta(\xi)\theta(-\zeta)\right|$$
$$\leq E\left[W_1^2\left|e^{\mathbf{i}(\xi-\zeta) W_2}\right|\right] + E\left[|W_1|\left|e^{\mathbf{i}\xi W_2}\right|\right] E\left[|W_1|\left|e^{-\mathbf{i}\zeta W_2}\right|\right]$$
$$\leq E\left[W_1^2\right] + E\left[|W_1|\right] E\left[|W_1|\right] \preceq 1,$$

$$\left|E\left[n\delta\hat{\theta}(\xi)\delta\hat{\phi}^{\dagger}_{W_2}(\zeta)\right]\right| = \left|\theta(\xi-\zeta) - \theta(\xi)\phi_{W_2}(-\zeta)\right|$$

$$
\begin{aligned}
&\leq E\left[|W_1|\left|e^{\mathbf{i}(\xi-\zeta)W_2}\right|\right]+E\left[|W_1|\left|e^{\mathbf{i}\xi W_2}\right|\right]E\left[\left|e^{-\mathbf{i}\zeta W_2}\right|\right]\\
&\leq 2E[|W_1|]\preceq 1,
\end{aligned}
$$

$$
\begin{aligned}
&\sup_{x\in\mathbb{S}_X}\left|E\left[n\delta\hat{\theta}(\xi)h_{2,1}^{1+\lambda_{2,1}}h_{2,2}^{1+\lambda_{2,2}}\delta\hat{\phi}^{\dagger}_{f_{Y,X,W_2}^{(\lambda_{2,1},\lambda_{2,2})}(y,x,\cdot)}(\zeta)\right]\right|\\
&=\sup_{x\in\mathbb{S}_X}\left|E\left[W_1e^{\mathbf{i}(\xi-\zeta)W_2}G_Y\left(\frac{y-Y}{h_{2,1}}\right)G_X^{(\lambda_{2,1},\lambda_{2,2})}\left(\frac{x-X}{h_{2,2}}\right)\right]-\theta(\xi)h_{2,1}^{1+\lambda_{2,1}}h_{2,2}^{1+\lambda_{2,2}}\phi_{f_{Y,X,W_2}^{(\lambda_{2,1},\lambda_{2,2})}(y,x,\cdot)}(-\zeta)\right|\\
&\leq E\left[|W_1|\left|e^{\mathbf{i}(\xi-\zeta)W_2}\right|\sup_{x\in\mathbb{S}_X}\left|G_Y\left(\frac{y-Y}{h_{2,1}}\right)G_X^{(\lambda_{2,1},\lambda_{2,2})}\left(\frac{x-X}{h_{2,2}}\right)\right|\right]\\
&\qquad+E\left[|W_1|\left|e^{\mathbf{i}\xi W_2}\right|\right]E\left[\left|e^{-\mathbf{i}\zeta W_2}\right|\sup_{x\in\mathbb{S}_X}\left|G_Y\left(\frac{y-Y}{h_{2,1}}\right)G_X^{(\lambda_{2,1},\lambda_{2,2})}\left(\frac{x-X}{h_{2,2}}\right)\right|\right]\\
&\leq 2E[|W_1|]\preceq 1,
\end{aligned}
$$

where the last line follows by Assumption 9. Following the same logic, I have that

$$
\begin{aligned}
&\left|E\left[n\delta\hat{\phi}_{W_2}(\xi)\delta\hat{\phi}^{\dagger}_{W_2}(\zeta)\right]\right|\preceq 1,\\
&\sup_{x\in\mathbb{S}_X}\left|E\left[n\delta\hat{\phi}_{W_2}(\xi)h_{2,1}^{1+\lambda_{2,1}}h_{2,2}^{1+\lambda_{2,2}}\delta\hat{\phi}^{\dagger}_{f_{Y,X,W_2}^{(\lambda_{2,1},\lambda_{2,2})}(y,x,\cdot)}(\zeta)\right]\right|\preceq 1,\\
&\sup_{x\in\mathbb{S}_X}\left|E\left[nh_{2,1}^{2+2\lambda_{2,1}}h_{2,2}^{2+2\lambda_{2,2}}\delta\hat{\phi}_{f_{Y,X,W_2}^{(\lambda_{2,1},\lambda_{2,2})}(y,x,\cdot)}(\xi)\delta\hat{\phi}^{\dagger}_{f_{Y,X,W_2}^{(\lambda_{2,1},\lambda_{2,2})}(y,x,\cdot)}(\zeta)\right]\right|\preceq 1,\\
&\left|E\left[n\delta\hat{\phi}_{W_2}(\zeta)\delta\hat{\theta}^{\dagger}(\xi)\right]\right|\preceq 1,\\
&\sup_{x\in\mathbb{S}_X}\left|E\left[nh_{2,1}^{1+\lambda_{2,1}}h_{2,2}^{1+\lambda_{2,2}}\delta\hat{\phi}_{f_{Y,X,W_2}^{(\lambda_{2,1},\lambda_{2,2})}(y,x,\cdot)}(\zeta)\delta\hat{\theta}^{\dagger}(\xi)\right]\right|\preceq 1,\\
&\sup_{x\in\mathbb{S}_X}\left|E\left[nh_{2,1}^{1+\lambda_{2,1}}h_{2,2}^{1+\lambda_{2,2}}\delta\hat{\phi}_{f_{Y,X,W_2}^{(\lambda_{2,1},\lambda_{2,2})}(y,x,\cdot)}(\zeta)\delta\hat{\phi}^{\dagger}_{W_2}(\xi)\right]\right|\preceq 1.
\end{aligned}
$$

It follows from Equation (35) that

$$
\begin{aligned}
&\Omega_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h)\\
&\leq\int\int\left|\Psi_{1,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,y,x,w^*,h_1)\right|\left|\left(\Psi_{1,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\zeta,y,x,w^*,h_1)\right)^{\dagger}\right|d\xi d\zeta\\
&\quad+\int\int\left|\Psi_{2,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,y,x,w^*,h_1)\right|\left|\left(\Psi_{2,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\zeta,y,x,w^*,h_1)\right)^{\dagger}\right|d\xi d\zeta\\
&\quad+\int\int\left|(h_{2,1}^{-1})^{1+\lambda_{2,1}}(h_{2,2}^{-1})^{1+\lambda_{2,2}}\Psi_{3,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,y,x,w^*,h_1)\right|\times\\
&\qquad\qquad\left|\left((h_{2,1}^{-1})^{1+\lambda_{2,1}}(h_{2,2}^{-1})^{1+\lambda_{2,2}}\Psi_{3,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\zeta,y,x,w^*,h_1)\right)^{\dagger}\right|d\xi d\zeta\\
&\quad+\int\int\left|\Psi_{1,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,y,x,w^*,h_1)\right|\left|\left(\Psi_{2,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\zeta,y,x,w^*,h_1)\right)^{\dagger}\right|d\xi d\zeta\\
&\quad+\int\int\left|\Psi_{1,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,y,x,w^*,h_1)\right|\left|\left((h_{2,1}^{-1})^{1+\lambda_{2,1}}(h_{2,2}^{-1})^{1+\lambda_{2,2}}\Psi_{3,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\zeta,y,x,w^*,h_1)\right)^{\dagger}\right|d\xi d\zeta
\end{aligned}
$$

$$
\begin{aligned}
&+\int\int\left|\Psi_{2,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,y,x,w^*,h_1)\right|\left|\left((h_{2,1}^{-1})^{1+\lambda_{2,1}}(h_{2,2}^{-1})^{1+\lambda_{2,2}}\Psi_{3,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\zeta,y,x,w^*,h_1)\right)^{\dagger}\right|d\xi d\zeta\\
&+\int\int\left|\Psi_{2,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,y,x,w^*,h_1)\right|\left|\left(\Psi_{1,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\zeta,y,x,w^*,h_1)\right)^{\dagger}\right|d\xi d\zeta\\
&+\int\int\left|(h_{2,1}^{-1})^{1+\lambda_{2,1}}(h_{2,2}^{-1})^{1+\lambda_{2,2}}\Psi_{3,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\zeta,y,x,w^*,h_1)\right|\left|\left(\Psi_{1,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,y,x,w^*,h_1)\right)^{\dagger}\right|d\xi d\zeta\\
&+\int\int\left|(h_{2,1}^{-1})^{1+\lambda_{2,1}}(h_{2,2}^{-1})^{1+\lambda_{2,2}}\Psi_{3,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\zeta,y,x,w^*,h_1)\right|\left|\left(\Psi_{2,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,y,x,w^*,h_1)\right)^{\dagger}\right|d\xi d\zeta\\
=&\left(\int\left|\Psi_{1,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,y,x,w^*,h_1)\right|d\xi+\int\left|\Psi_{2,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,y,x,w^*,h_1)\right|d\xi\right.\\
&\left.\qquad+\int\left|(h_{2,1}^{-1})^{1+\lambda_{2,1}}(h_{2,2}^{-1})^{1+\lambda_{2,2}}\Psi_{3,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,y,x,w^*,h_1)\right|d\xi\right)^2\\
\le&\left(\int\Psi^+_{1,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,h_1)d\xi+\int\Psi^+_{2,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,h_1)d\xi+\int(h_{2,1}^{-1})^{1+\lambda_{2,1}}(h_{2,2}^{-1})^{1+\lambda_{2,2}}\Psi^+_{3,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,h)d\xi\right)^2\\
=&\left(\Psi^+_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(h)\right)^2,
\end{aligned}
$$

where $\Psi^+_{1,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,h_1)$, $\Psi^+_{2,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,h_1)$, $\Psi^+_{3,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,h)$, and $\Psi^+_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(h)$ are as defined in Lemma 14 and

$$
\begin{aligned}
&\Psi^+_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(h)\\
=&O\left(\max\left\{(h_1^{-1})^{1+\gamma_*},(h_{2,1}^{-1})^{1+\lambda_{2,1}}(h_{2,2}^{-1})^{1+\lambda_{2,2}}\right\}(h_1^{-1})^{1-\gamma_2+\gamma_\phi+\lambda_1}\exp\left(\left(\alpha_\phi\mathbf{1}\{\beta_\phi=\beta_2\}-\alpha_2\right)\left(h_1^{-1}\right)^{\beta_2}\right)\right).
\end{aligned}
$$

Hence I have proved Eq. (20)

Next I show Eq. (21). From Eq. (33) I have

$$
\begin{aligned}
&\sup_{(y,x,w^*)\in\mathbb{S}_{(Y,X,W^*)}}\left|L_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h)\right|\\
=&\sup_{(y,x,w^*)\in\mathbb{S}_{(Y,X,W^*)}}\left|\int\Psi_{1,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,y,x,w^*,h_1)\left(\hat{E}[W_1e^{\mathbf{i}\xi W_2}]-E[W_1e^{\mathbf{i}\xi W_2}]\right)d\xi\right.\\
&\qquad+\int\Psi_{2,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,y,x,w^*,h_1)\left(\hat{E}[e^{\mathbf{i}\xi W_2}]-E[e^{\mathbf{i}\xi W_2}]\right)d\xi\\
&\qquad+\int\Psi_{3,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,y,x,w^*,h_1)\times\\
&\qquad\qquad\left(\hat{E}\left[e^{\mathbf{i}\xi W_2}\frac{1}{h_{2,1}^{1+\lambda_{2,1}}h_{2,2}^{1+\lambda_{2,2}}}G_Y\left(\frac{y-Y}{h_{2,1}}\right)G_X^{(\lambda_{2,1},\lambda_{2,2})}\left(\frac{x-X}{h_{2,2}}\right)\right]\right.\\
&\qquad\qquad\quad\left.\left.-E\left[e^{\mathbf{i}\xi W_2}\frac{1}{h_{2,1}^{1+\lambda_{2,1}}h_{2,2}^{1+\lambda_{2,2}}}G_Y\left(\frac{y-Y}{h_{2,1}}\right)G_X^{(\lambda_{2,1},\lambda_{2,2})}\left(\frac{x-X}{h_{2,2}}\right)\right]\right)d\xi\right|\\
\le&\int\left(\sup_{(y,x,w^*)\in\mathbb{S}_{(Y,X,W^*)}}\Psi_{1,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,y,x,w^*,h_1)\right)\left|\hat{E}[W_1e^{\mathbf{i}\xi W_2}]-E[W_1e^{\mathbf{i}\xi W_2}]\right|d\xi\\
&+\int\left(\sup_{(y,x,w^*)\in\mathbb{S}_{(Y,X,W^*)}}\Psi_{2,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,y,x,w^*,h_1)\right)\left|\hat{E}[e^{\mathbf{i}\xi W_2}]-E[e^{\mathbf{i}\xi W_2}]\right|d\xi\\
&+\int\left(\sup_{(y,x,w^*)\in\mathbb{S}_{(Y,X,W^*)}}(h_{2,1}^{-1})^{1+\lambda_{2,1}}(h_{2,2}^{-1})^{1+\lambda_{2,2}}\Psi_{3,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,y,x,w^*,h_1)\right)\times\\
&\sup_{x\in\mathbb{S}_X}\left|\hat{E}\left[e^{\mathbf{i}\xi W_2}G_Y\left(\frac{y-Y}{h_{2,1}}\right)G_X^{(\lambda_{2,1},\lambda_{2,2})}\left(\frac{x-X}{h_{2,2}}\right)\right]-E\left[e^{\mathbf{i}\xi W_2}G_Y\left(\frac{y-Y}{h_{2,1}}\right)G_X^{(\lambda_{2,1},\lambda_{2,2})}\left(\frac{x-X}{h_{2,2}}\right)\right]\right|d\xi\\
=&\int\Psi^+_{1,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,h_1)\left|\hat{E}[W_1e^{\mathbf{i}\xi W_2}]-E[W_1e^{\mathbf{i}\xi W_2}]\right|d\xi
\end{aligned}
$$

$$+\int \Psi^{+}_{2,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,h_1)\left|\hat{E}[e^{\mathbf{i}\xi W_2}]-E[e^{\mathbf{i}\xi W_2}]\right|d\xi$$
$$+\int (h_{2,1}^{-1})^{1+\lambda_{2,1}}(h_{2,2}^{-1})^{1+\lambda_{2,2}}\Psi^{+}_{3,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,h_1)\times$$
$$\sup_{x\in\mathbb{S}_X}\left|\hat{E}\left[e^{\mathbf{i}\xi W_2}G_Y\left(\frac{y-Y}{h_{2,1}}\right)G_X^{(\lambda_{2,1},\lambda_{2,2})}\left(\frac{x-X}{h_{2,2}}\right)\right]-E\left[e^{\mathbf{i}\xi W_2}G_Y\left(\frac{y-Y}{h_{2,1}}\right)G_X^{(\lambda_{2,1},\lambda_{2,2})}\left(\frac{x-X}{h_{2,2}}\right)\right]\right|d\xi$$

where $\Psi^{+}_{1,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,h_1)$, $\Psi^{+}_{2,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,h_1)$ and $\Psi^{+}_{3,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,h_1)$ are defined in Lemma 14. The integrals are finite for two reasons. First, with probability approaching 1, I have $\left|\hat{E}[W_1e^{\mathbf{i}\xi W_2}]-E[W_1e^{\mathbf{i}\xi W_2}]\right| \le 2E[|W_1|]+\epsilon \le \infty$, for some positive real number ϵ. This follows from Assumption 12, the fact that

$$\begin{aligned}\left|\hat{E}[W_1e^{\mathbf{i}\xi W_2}]-E[W_1e^{\mathbf{i}\xi W_2}]\right| &\le \left|\hat{E}[W_1e^{\mathbf{i}\xi W_2}]\right|+\left|E[W_1e^{\mathbf{i}\xi W_2}]\right|\\ &\le \hat{E}[\left|W_1e^{\mathbf{i}\xi W_2}\right|]+E[\left|W_1e^{\mathbf{i}\xi W_2}\right|]\\ &\le \hat{E}[|W_1|]+E[|W_1|],\end{aligned}$$

and that $\hat{E}[|W_1|] \xrightarrow{p} E[|W_1|]$. Also, it's easy to show that $\left|\hat{E}[e^{\mathbf{i}\xi W_2}]-E[e^{\mathbf{i}\xi W_2}]\right| \le 2+\epsilon < \infty$ and that $\sup_{x\in\mathbb{S}_X}\left|\hat{E}\left[e^{\mathbf{i}\xi W_2}G_Y\left(\frac{y-Y}{h_{2,1}}\right)G_X^{(\lambda_{2,1},\lambda_{2,2})}\left(\frac{x-X}{h_{2,2}}\right)\right]-E\left[e^{\mathbf{i}\xi W_2}G_Y\left(\frac{y-Y}{h_{2,1}}\right)G_X^{(\lambda_{2,1},\lambda_{2,2})}\left(\frac{x-X}{h_{2,2}}\right)\right]\right| \preceq 1$. Second, it follows by Lemma 14 that $\int \Psi^{+}_{j,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,h_1)d\xi < \infty$ for $j=1,2$ and that $\int (h_{2,1}^{-1})^{1+\lambda_{2,1}}(h_{2,2}^{-1})^{1+\lambda_{2,2}}\Psi^{+}_{3,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,h_1)d\xi<\infty$.

By Assumption 12,

$$\begin{aligned}E\left[\left(\hat{E}[W_1e^{\mathbf{i}\xi W_2}]-E[W_1e^{\mathbf{i}\xi W_2}]\right)^2\right] &\le \frac{1}{n}E\left[\left(W_1e^{\mathbf{i}\xi W_2}-E[W_1e^{\mathbf{i}\xi W_2}]\right)^2\right]\le\frac{1}{n}E\left[\left(W_1e^{\mathbf{i}\xi W_2}\right)^2\right]\\ &\le \frac{1}{n}E\left[W_1^2\right]=O\left(\frac{1}{n}\right),\end{aligned}$$

which implies $\left|\hat{E}[W_1e^{\mathbf{i}\xi W_2}]-E[W_1e^{\mathbf{i}\xi W_2}]\right| = O_p(n^{-1/2})$. Similarly, I also have $\left|\hat{E}[e^{\mathbf{i}\xi W_2}]-E[e^{\mathbf{i}\xi W_2}]\right|=O_p(n^{-1/2})$.

By the conclusions above and Lemma 15, I have

$$\sup_{(y,x,w^*)\in\mathbb{S}_{(Y,X,W^*)}}\left|L_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h)\right|$$
$$\le \int \Psi^{+}_{1,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,h_1)\left|\hat{E}[W_1e^{\mathbf{i}\xi W_2}]-E[W_1e^{\mathbf{i}\xi W_2}]\right|d\xi$$
$$+\int \Psi^{+}_{2,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,h_1)\left|\hat{E}[e^{\mathbf{i}\xi W_2}]-E[e^{\mathbf{i}\xi W_2}]\right|d\xi$$
$$+\int \Psi^{+}_{3,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,h_1)\times$$
$$\sup_{(y,x)\in\mathbb{S}_{(Y,X)}}\left|\hat{E}\left[e^{\mathbf{i}\xi W_2}\frac{1}{h_{2,1}^{1+\lambda_{2,1}}h_{2,2}^{1+\lambda_{2,2}}}G_Y^{(\lambda_{2,1})}\left(\frac{y-Y}{h_{2,1}}\right)G_X^{(\lambda_{2,2})}\left(\frac{x-X}{h_{2,2}}\right)\right]\right.$$
$$\left.-E\left[e^{\mathbf{i}\xi W_2}\frac{1}{h_{2,1}^{1+\lambda_{2,1}}h_{2,2}^{1+\lambda_{2,2}}}G_Y^{(\lambda_{2,1})}\left(\frac{y-Y}{h_{2,1}}\right)G_X^{(\lambda_{2,2})}\left(\frac{x-X}{h_{2,2}}\right)\right]\right|d\xi$$
$$\preceq n^{-1/2}\times$$

$$\left(\int \Psi^{+}_{1,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,h_1)d\xi + \int \Psi^{+}_{2,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,h_1)d\xi + \int (h_{2,1}^{-1})^{1+\lambda_{2,1}}(h_{2,2}^{-1})^{1+\lambda_{2,2}}\Psi^{+}_{3,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,h_1)d\xi\right)$$
$$= n^{-1/2}\Psi^{+}_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(h),$$

where

$$\Psi^{+}_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(h)$$
$$= O\left(\max\left\{(h_1^{-1})^{1+\gamma_*},(h_{2,1}^{-1})^{1+\lambda_{2,1}}(h_{2,2}^{-1})^{1+\lambda_{2,2}}\right\}(h_1^{-1})^{1-\gamma_2+\gamma_\phi+\lambda_1}\exp\left((\alpha_\phi \mathbf{1}\{\beta_\phi=\beta_2\}-\alpha_2)\left(h_1^{-1}\right)^{\beta_2}\right)\right),$$

as shown in Lemma 14.

(ii) Next I show asymptotic normality. I apply Lemma 16 to

$$\begin{aligned}
&l_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h_n;Y,X,W_1,W_2)\\
&=\int \Psi_{1,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,y,x,w^*,h_1)\left(W_1e^{\mathbf{i}\xi W_2}-E[W_1e^{\mathbf{i}\xi W_2}]\right)d\xi\\
&\quad+\int \Psi_{2,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,y,x,w^*,h_1)\left(e^{\mathbf{i}\xi W_2}-E[e^{\mathbf{i}\xi W_2}]\right)d\xi\\
&\quad+\int \Psi_{3,\lambda_1,\lambda_{2,1},\lambda_{2,2}}(\xi,y,x,w^*,h_1)\times\\
&\qquad\left(\frac{1}{h_{2,1}^{1+\lambda_{2,1}}h_{2,2}^{1+\lambda_{2,2}}}G_Y\left(\frac{y-Y}{h_{2,1}}\right)G_X^{(\lambda_{2,1},\lambda_{2,2})}\left(\frac{x-X}{h_{2,2}}\right)e^{\mathbf{i}\xi W_2}\right.\\
&\qquad\quad\left.-E\left[\frac{1}{h_{2,1}^{1+\lambda_{2,1}}h_{2,2}^{1+\lambda_{2,2}}}G_Y\left(\frac{y-Y}{h_{2,1}}\right)G_X^{(\lambda_{2,1},\lambda_{2,2})}\left(\frac{x-X}{h_{2,2}}\right)e^{\mathbf{i}\xi W_2}\right]\right)d\xi,
\end{aligned}$$

where $A=Y,X,W_1,W_2$ Previous argument ensures that for fixed h, $\Omega_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h)<\infty$. The desired conclusion follows. ■

Lemma 17. *(Schennach (2004a) Lemma 6) Let A and W_2 be random variables satisfying $E[|A|^2]<\infty$ and $E[|A||W_2|]<\infty$ and let $(A_j,W_{2,j})_{j=1,\dots,n}$ be a corresponding IID sample. Then for any $u,U\geq 0$ and $\epsilon>0$,*

$$\sup_{\zeta\in[-Un^u,Un^u]}\left|\hat{E}[A\exp(\mathbf{i}\zeta W_2)]-E[A\exp(\mathbf{i}\zeta W_2)]\right|=o_p\left(n^{-1/2+\epsilon}\right).$$

Proof of Theorem 3. Plug (27) (28) and (29) into

$$\begin{aligned}
&\hat{g}_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h)-g_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h_1)\\
&=\frac{1}{2\pi}\int(-\mathbf{i}t)^{\lambda_2}e^{-\mathbf{i}tw^*}\phi_K(h_1t)\\
&\quad\times\left[\frac{\hat{\phi}_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(y,x,\cdot)}(t)}{\hat{\phi}_{W_2}(t)}\exp\left(\int_0^t\frac{\mathbf{i}\hat{\theta}(\xi)}{\hat{\phi}_{W_2}(\xi)}d\xi\right)-\frac{\phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(y,x,\cdot)}(t)}{\phi_{W_2}(t)}\exp\left(\int_0^t\frac{\mathbf{i}\theta(\xi)}{\phi_{W_2}(\xi)}d\xi\right)\right]dt.
\end{aligned}$$

and remove terms linear in $\delta\hat{\theta}(t)$, $\delta\hat{\phi}_{W_2}(\xi)$, and $\delta\hat{\phi}_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W^*}(y,x,\cdot)}$. For notation simplicity, I write h instead of h_n here. I can then find that $\left|\hat{g}_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h) - g_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h_1)\right| \preceq \frac{1}{2\pi}\sum_{l=1}^{7} R_{l,\lambda_1,\lambda_{2,1},\lambda_{2,2}}$, where

$$
\begin{aligned}
R_{1,\lambda_1,\lambda_{2,1},\lambda_{2,2}} &= \int_0^\infty |t|^{\lambda_1}\,|\phi_K(h_1 t)|\,|\delta_1\hat{q}_{\lambda_2}(t)|\,|\phi_{W^*}(t)| \left(\int_0^t |\delta_1\hat{q}_{W_1}(\xi)|\,d\xi\right) dt \\
R_{2,\lambda_1,\lambda_{2,1},\lambda_{2,2}} &= \int_0^\infty |t|^{\lambda_1}\,|\phi_K(h_1 t)|\,|\delta_2\hat{q}_{\lambda_2}(t)|\,|\phi_{W^*}(t)|\,dt \\
R_{3,\lambda_1,\lambda_{2,1},\lambda_{2,2}} &= \int_0^\infty |t|^{\lambda_1}\,|\phi_K(h_1 t)|\,|\delta_2\hat{q}_{\lambda_2}(t)|\,|\phi_{W^*}(t)| \left(\int_0^t |\delta_1\hat{q}_{W_1}(\xi)|\,d\xi\right) dt \\
R_{4,\lambda_1,\lambda_{2,1},\lambda_{2,2}} &= \int_0^\infty |t|^{\lambda_1}\,|\phi_K(h_1 t)|\,|q_{\lambda_2}(t)|\,|\phi_{W^*}(t)| \left(\int_0^t |\delta_2\hat{q}_{W_1}(\xi)|\,d\xi\right) dt \\
R_{5,\lambda_1,\lambda_{2,1},\lambda_{2,2}} &= \int_0^\infty |t|^{\lambda_1}\,|\phi_K(h_1 t)|\,|\delta\hat{q}_{\lambda_2}(t)|\,|\phi_{W^*}(t)| \left(\int_0^t |\delta_2\hat{q}_{W_1}(\xi)|\,d\xi\right) dt \\
R_{6,\lambda_1,\lambda_{2,1},\lambda_{2,2}} &= \int_0^\infty |t|^{\lambda_1}\,|\phi_K(h_1 t)|\,|q_{\lambda_2}(t)|\,|\phi_{W^*}(t)|\,\frac{1}{2}\exp\left(\left|\delta\bar{Q}(t)\right|\right) \left(\int_0^t |\delta\hat{q}_{W_1}(\xi)|\,d\xi\right)^2 dt \\
R_{7,\lambda_1,\lambda_{2,1},\lambda_{2,2}} &= \int_0^\infty |t|^{\lambda_1}\,|\phi_K(h_1 t)|\,|\delta\hat{q}_{\lambda_2}(t)|\,|\phi_{W^*}(t)|\,\frac{1}{2}\exp\left(\left|\delta\bar{Q}(t)\right|\right) \left(\int_0^t |\delta\hat{q}_{W_1}(\xi)|\,d\xi\right)^2 dt,
\end{aligned}
$$

where $q_{\lambda_2}(t)$, $\delta\hat{q}_{\lambda_2}(t)$, $\delta_1\hat{q}_{\lambda_2}(t)$, $\delta_2\hat{q}_{\lambda_2}(t)$, $q_{W_1}(t)$, $\delta\hat{q}_{W_1}(t)$, $\delta_1\hat{q}_{W_1}(t)$, $\delta_2\hat{q}_{W_1}(t)$, and $\delta\bar{Q}(t)$ were defined in the proof of Lemma 5.

Lemma 17 gives that for any $\epsilon > 0$,

$$
\begin{aligned}
&\sup_{(y,x)\in\mathbb{S}_{(Y,X)}}\ \sup_{\xi\in[-h_{1n}^{-1},h_{1n}^{-1}]} \Bigg|\hat{E}\left[e^{\mathbf{i}\xi W_2}\frac{1}{h_{2,1}^{1+\lambda_{2,1}}h_{2,2}^{1+\lambda_{2,2}}}G_Y^{(\lambda_{2,1})}\left(\frac{y-Y}{h_{2n,1}}\right)G_X^{(\lambda_{2,2})}\left(\frac{x-X}{h_{2n,2}}\right)\right] \\
&\qquad -E\left[e^{\mathbf{i}\xi W_2}\frac{1}{h_{2,1}^{1+\lambda_{2,1}}h_{2,2}^{1+\lambda_{2,2}}}G_Y^{(\lambda_{2,1})}\left(\frac{y-Y}{h_{2n,1}}\right)G_X^{(\lambda_{2,2})}\left(\frac{x-X}{h_{2n,2}}\right)\right]\Bigg| \\
&= \left(h_{2,1}^{-1}\right)^{1+\lambda_{2,1}}(h_{2,2}^{-1})^{1+\lambda_{2,2}} \sup_{x\in\mathbb{S}_X}\left|G_Y^{(\lambda_{2,1})}\left(\frac{y-Y}{h_{2n,1}}\right)G_X^{(\lambda_{2,2})}\left(\frac{x-X}{h_{2n,2}}\right)\right| \sup_{\xi\in[-h_{1n}^{-1},h_{1n}^{-1}]}\left|\hat{E}\left[e^{\mathbf{i}\xi W_2}\right]-E\left[e^{\mathbf{i}\xi W_2}\right]\right| \\
&= o_p\left(\left(h_{2,1}^{-1}\right)^{1+\lambda_{2,1}}(h_{2,2}^{-1})^{1+\lambda_{2,2}}n^{-1/2+\epsilon}\right).
\end{aligned}
$$

By (34) I have that

$$
\begin{aligned}
&\sup_{(y,x)\in\mathbb{S}_{(Y,X)}}\ \sup_{\xi\in[-h_{1n}^{-1},h_{1n}^{-1}]}\left|E\left[e^{\mathbf{i}\xi W_2}\frac{1}{h_{2,1}^{1+\lambda_{2,1}}h_{2,2}^{1+\lambda_{2,2}}}G_Y^{(\lambda_{2,1})}\left(\frac{y-Y}{h_{2n,1}}\right)G_X^{(\lambda_{2,2})}\left(\frac{x-X}{h_{2n,2}}\right)\right]-\phi_{f^{(\lambda_2)}_{Y,X,W_2}(y,x,\cdot)}(\xi)\right| \\
&\preceq \sup_{\xi\in[-h_{1n}^{-1},h_{1n}^{-1}]}|\phi_{W_2}(\xi)|\,O\left(\left(\frac{\bar{\xi}_{G_Y}}{h_{2,1}}\right)^{1+\gamma_{f_1}}\left(\frac{\bar{\xi}_{G_X}}{h_{2,2}}\right)^{1+\gamma_{f_2}}\exp\left(\alpha_{f_1}\left(\frac{\bar{\xi}_{G_Y}}{h_{2,1}}\right)^{\beta_{f_1}}+\alpha_{f_2}\left(\frac{\bar{\xi}_{G_X}}{h_{2,2}}\right)^{\beta_{f_2}}\right)\right) \\
&\preceq O\left(\left(\frac{\bar{\xi}_{G_Y}}{h_{2,1}}\right)^{1+\gamma_{f_1}}\left(\frac{\bar{\xi}_{G_X}}{h_{2,2}}\right)^{1+\gamma_{f_2}}\exp\left(\alpha_{f_1}\left(\frac{\bar{\xi}_{G_Y}}{h_{2,1}}\right)^{\beta_{f_1}}+\alpha_{f_2}\left(\frac{\bar{\xi}_{G_X}}{h_{2,2}}\right)^{\beta_{f_2}}\right)\right) \\
&= o\left(\left(h_{2,1}^{-1}\right)^{1+\lambda_{2,1}}(h_{2,2}^{-1})^{1+\lambda_{2,2}}n^{-1/2+\epsilon}\right),
\end{aligned}
$$

where the last inequality holds under Assumption 15. Combining results from above, I have that for any $\epsilon > 0$,

$$
\begin{aligned}
&\sup_{(y,x)\in\mathbb{S}(Y,X)} \sup_{\xi\in[-h_{1n}^{-1},h_{1n}^{-1}]} \left|\hat{\phi}_{f^{(\lambda_2)}_{Y,X,W_2}(y,x,\cdot)}(\xi) - \phi_{f^{(\lambda_2)}_{Y,X,W_2}(y,x,\cdot)}(\xi)\right| \\
&= \sup_{(y,x)\in\mathbb{S}(Y,X)} \sup_{\xi\in[-h_{1n}^{-1},h_{1n}^{-1}]} \left|\hat{E}\left[e^{\mathbf{i}\xi W_2} \frac{1}{h_{2,1}^{1+\lambda_{2,1}} h_{2,2}^{1+\lambda_{2,2}}} G_Y^{(\lambda_{2,1})}\left(\frac{y-Y}{h_{2n,1}}\right) G_X^{(\lambda_{2,2})}\left(\frac{x-X}{h_{2n,2}}\right)\right]\right. \\
&\qquad\qquad \left. -E\left[e^{\mathbf{i}\xi W_2} \frac{1}{h_{2,1}^{1+\lambda_{2,1}} h_{2,2}^{1+\lambda_{2,2}}} G_Y^{(\lambda_{2,1})}\left(\frac{y-Y}{h_{2n,1}}\right) G_X^{(\lambda_{2,2})}\left(\frac{x-X}{h_{2n,2}}\right)\right]\right| \\
&\quad + \sup_{(y,x)\in\mathbb{S}(Y,X)} \sup_{\xi\in[-h_{1n}^{-1},h_{1n}^{-1}]} \left|E\left[e^{\mathbf{i}\xi W_2} \frac{1}{h_{2,1}^{1+\lambda_{2,1}} h_{2,2}^{1+\lambda_{2,2}}} G_Y^{(\lambda_{2,1})}\left(\frac{y-Y}{h_{2n,1}}\right) G_X^{(\lambda_{2,2})}\left(\frac{x-X}{h_{2n,2}}\right)\right]\right. \\
&\qquad\qquad \left. -\phi_{f^{(\lambda_2)}_{Y,X,W_2}(y,x,\cdot)}(\xi)\right| \\
&= o_p\left(\left(h_{2,1}^{-1}\right)^{1+\lambda_{2,1}} (h_{2,2}^{-1})^{1+\lambda_{2,2}} n^{-1/2+\epsilon}\right).
\end{aligned}
$$

I define $\Upsilon(h_1)$ and $\hat{\Phi}_n$ as

$$
\begin{aligned}
\Upsilon(h_{1n}) &\equiv (1+h_{1n}^{-1}) \left(\sup_{\xi\in[-h^{-1},h^{-1}]} \frac{|\phi'_{W^*}(\xi)|}{|\phi_{W^*}(\xi)|}\right)\left(\sup_{\xi\in[-h_{1n}^{-1},h_{1n}^{-1}]} |\phi_{W_2}(\xi)|^{-1}\right) \\
&= O\left((1+h_{1n}^{-1})^{1+\gamma_*-\gamma_2} \exp\left(-\alpha_2 \left(h_{1n}^{-1}\right)^{\beta_2}\right)\right) \\
\hat{\Phi}_{n,\lambda_2} &\equiv \max\left\{\sup_{\xi\in[-h_{1n}^{-1},h_{1n}^{-1}]} \left|\hat{\theta}(\xi) - \theta(\xi)\right|, \sup_{\xi\in[-h_{1n}^{-1},h_{1n}^{-1}]} \left|\hat{\phi}_{W_2}(\xi) - \phi_{W_2}(\xi)\right|,\right. \\
&\qquad \left. \sup_{(y,x)\in\mathbb{S}(Y,X)} \sup_{\xi\in[-h_{1n}^{-1},h_{1n}^{-1}]} \left|\hat{\phi}_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(y,x,\cdot)}(\xi) - \phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(y,x,\cdot)}(\xi)\right|\right\} \\
&= o_p\left((h_{2,1}^{-1})^{1+\lambda_{2,1}} (h_{2,2}^{-1})^{1+\lambda_{2,2}} n^{-1/2+\epsilon}\right)
\end{aligned}
$$

for any $\epsilon > 0$. The latter order of magnitude follows from Lemma 17 and Assumption 15 and 14. Then $R_{1,\lambda_1,\lambda_{2,1},\lambda_{2,2}} - R_{7,\lambda_1,\lambda_{2,1},\lambda_{2,2}}$ can be bounded in terms of $\Psi^+_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(h_{1n})$, $\Upsilon(h_{1n})$, and $\hat{\Phi}_{n,\lambda_2}$. Note that under Assumption 15,

$$
\begin{aligned}
&\sup_{\xi\in[-h_{1n}^{-1},h_{1n}^{-1}]} \frac{\hat{\Phi}_{n,\lambda_2}}{|\phi_{W_2}(\xi)|} \\
&\preceq \hat{\Phi}_{n,\lambda_2} \Upsilon(h_{1n}) \\
&= o_p\left((h_{2,1}^{-1})^{1+\lambda_{2,1}} (h_{2,2}^{-1})^{1+\lambda_{2,2}} n^{-1/2+\epsilon}\right) O\left((1+h_{1n}^{-1})^{1+\gamma_*-\gamma_2} \exp\left(-\alpha_2 \left(h_{1n}^{-1}\right)^{\beta_2}\right)\right) \\
&= o_p(1).
\end{aligned}
$$

Now I have

$$
R_1 \le 2\int_0^\infty |t|^{\lambda_1} |\phi_K(h_1 t)| \left(\frac{1}{|\phi_{W_2}(t)|} + \frac{\left|\phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(y,x,\cdot)}(t)\right|}{|\phi_{W_2}(t)|^2}\right) \hat{\Phi}_{n,\lambda_2} |\phi_{W^*}(t)| \left(\int_0^t |\delta_1 \hat{q}_{W_1}(\xi)|\, d\xi\right) dt
$$

$$
\begin{aligned}
&\preceq \Upsilon(h_1)\hat{\Phi}_{n,\lambda_2}\int_0^\infty |t|^{\lambda_1}\left|\phi_K(h_1t)\right|\left(1+\frac{\left|\phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(y,x,\cdot)}(t)\right|}{\left|\phi_{W_2}(t)\right|}\right)\left|\phi_{W^*}(t)\right|\left(\int_0^t\left|\delta_1\hat{q}_{W_1}(\xi)\right|d\xi\right)d\tau dt\\
&=\Upsilon(h_1)\hat{\Phi}_{n,\lambda_2}\int_0^\infty\left[\int_\xi^\infty |t|^{\lambda_1}\left|\phi_K(h_1t)\right|\left(1+\frac{\left|\phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(y,x,\cdot)}(t)\right|}{\left|\phi_{W_2}(t)\right|}\right)\left|\phi_{W^*}(t)\right|dt\right]\left|\delta_1\hat{q}_{W_1}(\xi)\right|d\xi\\
&=\Upsilon(h_1)\hat{\Phi}_{n,\lambda_2}\int_0^\infty\left[\int_\xi^\infty |t|^{\lambda_1}\left|\phi_K(h_1t)\right|\left(\left|\phi_{W^*}(t)\right|+\left|\phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W^*}(y,x,\cdot)}(t)\right|\right)dt\right]\left|\delta_1\hat{q}_{W_1}(\xi)\right|d\xi\\
&\preceq \Upsilon(h_1)\hat{\Phi}^2_{n,\lambda_2}\int_0^\infty\left[\int_\xi^\infty |t|^{\lambda_1}\left|\phi_K(h_1t)\right|\left(\left|\phi_{W^*}(t)\right|+\left|\phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W^*}(y,x,\cdot)}(t)\right|\right)dt\right]\\
&\qquad\qquad\times\left(1+\frac{|\theta(\xi)|}{\left|\phi_{W_2}(\xi)\right|}\right)\frac{1}{\left|\phi_{W_2}(\xi)\right|}d\xi\\
&\preceq \Upsilon(h_1)\hat{\Phi}^2_{n,\lambda_2}\Psi^+_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(h)\\
&=o_p\left(h_{2,1}^{-2}(h_{2,2}^{-1})^{2+2\lambda_2}n^{-1+2\epsilon}(h_{1n}^{-1})^{1+\gamma_*-\gamma_2}\exp\left(-\alpha_2\left(h_{1n}^{-1}\right)^{\beta_2}\right)\right)\times\\
&O\left(\max\left\{(h_1^{-1})^{1+\gamma_*},(h_{2,1}^{-1})^{1+\lambda_{2,1}}(h_{2,2}^{-1})^{1+\lambda_{2,2}}\right\}(h_1^{-1})^{1-\gamma_2+\gamma_\phi+\lambda_1}\exp\left(\left(\alpha_\phi\mathbf{1}\{\beta_\phi=\beta_2\}-\alpha_2\right)\left(h_1^{-1}\right)^{\beta_2}\right)\right).
\end{aligned}
$$

If further Assumption 15 holds, I just need to show that

$$
o_p\left(h_{2,1}^{-2}(h_{2,2}^{-1})^{2+2\lambda_2}n^{-1/2+2\epsilon}(h_{1n}^{-1})^{1+\gamma_*-\gamma_2}\exp\left(-\alpha_2\left(h_{1n}^{-1}\right)^{\beta_2}\right)\right)=o_p(1).
$$

If $\beta_2\neq 0$, $h_{1n}^{-1}=O\left((\ln n)^{1/\beta_2-\eta}\right)$, $h_{2n,2}^{-1}=O\left(n^{(8+4\lambda_2)^{-1}-\eta}\right)$, so that

$$
\begin{aligned}
&o_p\left(n^{-1/2+2\epsilon}h_{2,1}^{-2}(h_{2,2}^{-1})^{2+2\lambda_2}(h_{1n}^{-1})^{1+\gamma_*-\gamma_2}\exp\left(-\alpha_2\left(h_{1n}^{-1}\right)^{\beta_2}\right)\right)\\
&=o_p\left(n^{-1/2+2\epsilon}n^{1/2-2\eta(2+\lambda_2)}(h_{1n}^{-1})^{1+\gamma_*-\gamma_2}\exp\left(-\alpha_2\left(h_{1n}^{-1}\right)^{\beta_2}\right)\right)\\
&=o_p\left(n^{2\epsilon-2\eta(2+\lambda_2)}((\ln n)^{1/\beta_2-\eta})^{1+\gamma_*-\gamma_2}\exp\left(-\alpha_2\left(\ln n\right)^{1-\eta\beta_2}\right)\right)\\
&=o_p\left(\exp\left[-\alpha_2\left(\ln n\right)^{1-\eta\beta_2}+(2\epsilon-2\eta(2+\lambda_2))\ln n+(1+\gamma_*-\gamma_2)\left(1/\beta_2-\eta\right)\ln\left(\ln n\right)\right]\right)\\
&=o_p(1),
\end{aligned}
$$

where the equality follows since $(\ln n)^{1-\eta\beta_2}$ and $\ln(\ln n)$ are dominated by $\ln n$ and by picking $\epsilon<\eta(2+\lambda_2)$.

If $\beta_2=0$, $h_{1n}^{-1}=O\left(n^{(4+4\gamma_*-4\gamma_2)^{-1}-\eta}\right)$, $h_{2n,2}^{-1}=O\left(n^{(16+8\lambda_2)^{-1}-\eta}\right)$ and with $\epsilon<\eta$,

$$
\begin{aligned}
&o_p\left(n^{-1/2+2\epsilon}(h_{1n}^{-1})^{1+\gamma_*-\gamma_2}(h_{2,1}^{-2}(h_{2,2}^{-1})^{2+2\lambda_2}\right)\\
&=o_p\left(n^{-1/2+2\epsilon}(n^{(4+4\gamma_*-4\gamma_2)^{-1}-\eta})^{1+\gamma_*-\gamma_2}n^{1/4-2\eta(2+\lambda_2)}\right)\\
&=o_p\left(n^{2\epsilon-\eta(5+\gamma_*-\gamma_2+2\lambda_2)}\right)
\end{aligned}
$$

$$= o_p(1).$$

The remaining terms are similarly bounded

$$\begin{aligned}
&R_{2,\lambda_1,\lambda_{2,1},\lambda_{2,2}} \\
&\leq \int_0^\infty |t|^{\lambda_1} |\phi_K(h_1 t)| \left| \frac{\left|\phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(y,x,\cdot)}(t)\right|}{|\phi_{W_2}(t)|^2} \frac{1}{|\phi_{W_2}(t)|} \hat{\Phi}^2_{n,\lambda_2} |1+o_p(1)|^{-1} \right. \\
&\qquad \left. + \frac{1}{|\phi_{W_2}(t)|^2} \hat{\Phi}^2_{n,\lambda_2} |1+o_p(1)|^{-1} \right| |\phi_{W^*}(t)| \, dt \\
&\preceq \Upsilon(h_1) \hat{\Phi}^2_{n,\lambda_2} |1+o_p(1)|^{-1} \int_0^\infty |t|^{\lambda_1} |\phi_K(h_1 t)| \frac{1}{|\phi_{W_2}(t)|} \left| \frac{\left|\phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(y,x,\cdot)}(t)\right|}{|\phi_{W_2}(t)|} + 1 \right| |\phi_{W^*}(t)| \, dt \\
&\preceq \Upsilon(h_1) \hat{\Phi}^2_{n,\lambda_2} |1+o_p(1)|^{-1} \left(\int_0^\infty |t|^{\lambda_1} |\phi_K(h_1 t)| \frac{\left|\phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W^*}(y,x,\cdot)}(t)\right|}{|\phi_{W_2}(t)|} dt \right. \\
&\qquad\qquad \left. + \int_0^\infty |t|^{\lambda_1} |\phi_K(h_1 t)| \frac{|\phi_{W^*}(t)|}{|\phi_{W_2}(t)|} dt \right) \\
&\preceq \Upsilon(h_1) \hat{\Phi}^2_{n,,\lambda_2} \Psi^+_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(h) \left(1+o_p(1)\right)^{-1}; \\
&R_{3,\lambda_1,\lambda_{2,1},\lambda_{2,2}} \\
&\preceq \Upsilon(h_1) \hat{\Phi}_{n,\lambda_2} R_{2,\lambda_1,\lambda_{2,1},\lambda_{2,2}} = o_p(1) R_{2,\lambda_1,\lambda_{2,1},\lambda_{2,2}}; \\
&R_{4,\lambda_1,\lambda_{2,1},\lambda_{2,2}} \\
&= \int_0^\infty |t|^{\lambda_1} |\phi_K(h_1 t)| \left|\phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W^*}(y,x,\cdot)}(t)\right| \left(\int_0^t |\delta_2 \hat{q}_{W_1}(\xi)| \, d\xi \right) dt \\
&\preceq \Upsilon(h_1) \hat{\Phi}^2_{n,\lambda_2} |1+o_p(1)|^{-1} \int_0^\infty \frac{\int_\xi^\infty |t|^{\lambda_1} |\phi_K(h_1 t)| \left|\phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W^*}(y,x,\cdot)}(t)\right| dt}{|\phi_{W_2}(\xi)|} d\xi \\
&= \Upsilon(h_1) \hat{\Phi}^2_{n,\lambda_2} \Psi^+_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(h)(1+o_p(1)); \\
&R_{5,\lambda_1,\lambda_{2,1},\lambda_{2,2}} \\
&\leq \int_0^\infty |t|^{\lambda_1} |\phi_K(h_1 t)| \left(\frac{1}{|\phi_{W_2}(t)|} + \frac{\left|\phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(y,x,\cdot)}(t)\right|}{|\phi_{W_2}(t)|^2} \right) \\
&\qquad \times \hat{\Phi}_{n,\lambda_2} |1+o_p(1)|^{-1} |\phi_{W^*}(t)| \left(\int_0^t |\delta_2 \hat{q}_{W_1}(\xi)| \, d\xi \right) dt \\
&\preceq \Upsilon(h_1) \hat{\Phi}_{n,\lambda_2} |1+o_p(1)|^{-1} \int_0^\infty |\phi_K(h_1 t)| \left(|\phi_{W^*}(t)| + \left|\phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W^*}(y,x,\cdot)}(t)\right| \right) \\
&\qquad \times \left(\int_0^t |\delta_2 \hat{q}_{W_1}(\xi)| \, d\xi \right) d\tau dt \\
&\preceq \Upsilon(h_1) \hat{\Phi}_{n,\lambda_2} \left(1+o_p(1)\right) R_{4,\lambda_1,\lambda_{2,1},\lambda_{2,2}} = o_p(1) R_{4,\lambda_1,\lambda_{2,1},\lambda_{2,2}};
\end{aligned}$$

$$
\begin{aligned}
&R_{6,\lambda_1,\lambda_{2,1},\lambda_{2,2}}\\
&\leq \int_0^\infty |t|^{\lambda_1} |\phi_K(h_1 t)| \left|\phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W^*}(y,x,\cdot)}(t)\right| \frac{1}{2}\exp\left(\int_0^t |\delta \hat{q}_{W_1}(\xi)|\, d\xi\right)\left(\int_0^t |\delta \hat{q}_{W_1}(\xi)|\, d\xi\right)^2 dt\\
&\leq \frac{1}{2}\exp\left(o_p(1)\right)\int_0^\infty |t|^{\lambda_1} |\phi_K(h_1 t)| \left|\phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W^*}(y,x,\cdot)}(t)\right| \left(\int_0^t |\delta \hat{q}_{W_1}(\xi)|\, d\xi\right)^2 dt\\
&\preceq \frac{1}{2}\exp\left(o_p(1)\right)\Upsilon(h_1)\hat{\Phi}^2_{n,\lambda_2}\left|1+o_p(1)\right|^{-1}\int_0^\infty |t|^{\lambda_1} |\phi_K(h_1 t)| \left|\phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W^*}(y,x,\cdot)}(t)\right|\\
&\qquad\qquad\qquad\qquad \times\left(\int_0^t \frac{1}{|\phi_{W_2}(\xi)|}+\frac{|\theta(\xi)|}{|\phi_{W_2}(\xi)|^2}d\xi\right)dt\\
&\preceq \frac{1}{2}\exp\left(o_p(1)\right)\Upsilon(h_1)\hat{\Phi}^2_{n,\lambda_2}\left|1+o_p(1)\right|^{-1}\int_0^\infty \left(\int_\xi^\infty |t|^{\lambda_1} |\phi_K(h_1 t)| \left|\phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W^*}(y,x,\cdot)}(t)\right| dt\right)\\
&\qquad\qquad\qquad\qquad \times\left(\frac{1}{|\phi_{W_2}(\xi)|}+\frac{|\theta(\xi)|}{|\phi_{W_2}(\xi)|^2}\right)d\xi\\
&\preceq O_p(1)\Upsilon(h_1)\hat{\Phi}^2_{n,\lambda_2}\Psi^+_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(h)\\
&R_{7,\lambda_1,\lambda_{2,1},\lambda_{2,2}}\\
&\leq \int_0^\infty |t|^{\lambda_1} |\phi_K(h_1 t)| \left(1+\frac{\left|\phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W_2}(y,x,\cdot)}(t)\right|}{|\phi_{W_2}(t)|}\right)\Upsilon(h_1)\hat{\Phi}_{n,\lambda_2}\left|1+o_p(1)\right|^{-1}|\phi_{W^*}(t)|\\
&\quad\times\frac{1}{2}\exp\left(\int_0^t |\delta \hat{q}_{W_1}(\xi)|\, d\xi\right)\left(\int_0^t |\delta \hat{q}_{W_1}(\xi)|\, d\xi\right)^2 dt\\
&\preceq \Upsilon(h_1)\hat{\Phi}_{n,\lambda_2}\left|1+o_p(1)\right|^{-1}\int_0^\infty |t|^{\lambda_1} |\phi_K(h_1 t)| \left(|\phi_{W^*}(t)|+\left|\phi_{f^{(\lambda_{2,1},\lambda_{2,2})}_{Y,X,W^*}(y,x,\cdot)}(t)\right|\right)\\
&\quad\times\frac{1}{2}\exp\left(\int_0^t |\delta \hat{q}_{W_1}(\xi)|\, d\xi\right)\left(\int_0^t |\delta \hat{q}_{W_1}(\xi)|\, d\xi\right)^2 dt\\
&\preceq \Upsilon(h_1)\hat{\Phi}_{n,\lambda_2}\left|1+o_p(1)\right|^{-1}R_{6,\lambda_1,\lambda_{2,1},\lambda_{2,2}} = o_p(1)R_{6,\lambda_1,\lambda_{2,1},\lambda_{2,2}}.
\end{aligned}
$$

■

Proof of Theorem 4. In Lemma 12, let $Z=(X,W^*)$. By (9), for any value $(y,x,w^*)\in\mathbb{S}_{Y,X,W^*}$, and any value $\delta\in[0,1]$, I have that $\mathbf{WLAR}(x)=\tilde{\Gamma}(f)$ and that $\rho(y,x)=\tilde{\Xi}(f)$. (i) By Lemma 12,

$$
\begin{aligned}
&\sup_{(y,x,w^*)\in\mathbb{S}_\tau}\left|\widehat{\rho(y,x)}-\rho(y,x)\right|\\
&\leq \frac{O_p\left(\epsilon_{n,0}\right)+O_p\left(\tilde{\epsilon}_{n,0,0,0}\right)}{\tau^2}+O_p\left(\epsilon_{n,0}\right)+O_p\left(\tilde{\epsilon}_{n,0,0,1}\right)+\frac{O_p\left(\epsilon_{n,1}\right)+O_p\left(\tilde{\epsilon}_{n,1,0,0}\right)}{\tau^2}\\
&\leq O_p\left(\tilde{\epsilon}_{n,0,0,1}\right)+\frac{O_p\left(\epsilon_{n,1}\right)+O_p\left(\tilde{\epsilon}_{n,1,0,0}\right)}{\tau^2}.
\end{aligned}
$$

The last inequality follows since $\epsilon_{n,0}$ is of smaller order than $\epsilon_{n,1}$, $\tilde{\epsilon}_{n,0,0,0}$ is of smaller order than $\tilde{\epsilon}_{n,1,0,0}$. Choose τ_n such that $\tau_n > 0, \tau_n \to 0$ as $n \to \infty$, and that $\frac{\epsilon_{n,1}}{\tau^2} \to 0$ and $\frac{\tilde{\epsilon}_{n,1,0,0}}{\tau^2} \to 0$, I can get the desired result.

(ii) A direct application of Lemma 13 gives the desired result. ■

7.3 Appendix C: Asymptotic Normality

For asymptotic normality, I need to place a lower bound on $\Omega_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*.h_n)$ relative to $B_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*.h_n)$ and $R_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*.h_n)$. The following assumption is stated at a high level. More primitive sufficient conditions can be derived using techniques of Schennach (2004b). Combining this assumption with the results from Theorem 1,2 and 3 yields a corollary establishing the asymptotic normality of $\hat{g}_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h_n)$.

Assumption 16. *(High Level) For given $\lambda_1, \lambda_{2,1}, \lambda_{2,2} \in \{0,1\}$ and given $j \in \{1,2\}$, $h_n \to 0$ at a rate such that for each $(y,x,w^*) \in \mathbb{S}_{(Y,X,W^*)}$ such that $\Omega_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h_n) > 0$ for all n sufficiently large, I have $n^{1/2}\left(\Omega_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h_n)\right)^{-1/2}\left|B_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h_n)\right| \xrightarrow{p} 0$, and $n^{1/2}\left(\Omega_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h_n)\right)^{-1/2}\left|R_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h_n)\right| \xrightarrow{p} 0$.*

Corollary 2. *If the conditions of Theorem 2 and Assumption 16 hold, then for each $(y,x,w^*) \in \mathbb{S}_\tau$ such that $\Omega_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h_n) > 0$ for all n sufficiently large I have*

$$\begin{aligned}&n^{1/2}\left(\Omega_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h_n)\right)^{-1/2}\left(\hat{g}_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*,h_n) - g_{\lambda_1,\lambda_{2,1},\lambda_{2,2}}(y,x,w^*)\right)\\&\xrightarrow{d} N(0,1).\end{aligned}$$

In the next few lemmas, I show the asymptotic normality of $\widehat{\rho(y,x)}$ under some high-level assumptions. The asymptotic normality of $\widehat{\mathbf{WLAR}}(x)$ can be shown following the same logic. I first state a Lemma in a similar fashion as Lemma 5

Lemma 18. *For any value $(y,x,w^*) \in \mathbb{S}_\tau$, I have that*

$$\widehat{\rho(y,x)} - \rho(y,x) = BD\tilde{\Xi}_n + LD\tilde{\Xi}_n + RD\tilde{\Xi}_n + R\tilde{\Xi}_n,$$

where $R\tilde{\Xi}_n$ corresponds to the nonlinear part of functional $\tilde{\Xi}$; $BD\tilde{\Xi}_n$ and $RD\tilde{\Xi}_n$ are the bias *term contained in the linear part of functional* $\tilde{\Xi}$; $LD\tilde{\Xi}_n \equiv \hat{E}\left[lD\tilde{\Xi}_n\right]$ *is the linear term*[5] *contained in the linear part of functional* $\tilde{\Xi}$*, and is equal to* $\hat{E}\left[\sum_{k=1}^{11}\tilde{s}_k(y,x,w^*)lD\tilde{\Xi}_{n,k}\right]$ *in which*

$$\begin{aligned}
lD\tilde{\Xi}_{n,1} &= \int l_{0,0,0}(y,x,w^*,h_n)dy \\
lD\tilde{\Xi}_{n,2} &= \int_{-\infty}^{y} l_{0,0,0}(y,x,w^*,h_n)dy \\
lD\tilde{\Xi}_{n,3} &= \int l_{0,0,1}(y,x,w^*,h_n)dy \\
lD\tilde{\Xi}_{n,4} &= \int_{-\infty}^{y} l_{0,0,1}(y,x,w^*,h_n)dy \\
lD\tilde{\Xi}_{n,5} &= l_{0,0,0}(y,x,w^*,h_n) \\
lD\tilde{\Xi}_{n,6} &= \int l_{1,0,0}(y,x,w^*,h_n)dy \\
lD\tilde{\Xi}_{n,7} &= \int_{-\infty}^{y} l_{1,0,0}(y,x,w^*,h_n)dy \\
lD\tilde{\Xi}_{n,8} &= \int \tilde{l}_{0,0}(s,w^*,h_n)ds \\
lD\tilde{\Xi}_{n,9} &= \int_{-\infty}^{x} \tilde{l}_{0,0}(s,w^*,h_n)ds \\
lD\tilde{\Xi}_{n,10} &= \int \tilde{l}_{1,0}(s,w^*,h_n)ds \\
lD\tilde{\Xi}_{n,11} &= \int_{-\infty}^{x} \tilde{l}_{1,0}(s,w^*,h_n)ds.
\end{aligned}$$

$$\begin{aligned}
\tilde{s}_1(y,x,w^*) &= -\frac{F_{Y|X=x,W^*=w^*}(y)}{f_{X,W^*}(x,w^*)f_{Y,X,W^*}(y,x,w^*)}\left(\frac{\partial f_{X,W^*}(x,w^*)}{\partial x}+\frac{\partial f_{X,W^*}(x,w^*)}{\partial w^*}\frac{1}{\tilde{\Psi}_{(1)}(f)}\right) \\
&\quad -\frac{F_{X|W^*=w^*}(x)\frac{\partial f_{W^*}(w^*)}{\partial w^*}-\int_{-\infty}^{x}\frac{\partial f_{X,W^*}(s,w^*)}{\partial w^*}ds}{\tilde{f}^2(x,w^*)}\times\frac{\Psi_{(2)}(f)}{\tilde{\Psi}^2_{(1)}(f)} \\
\tilde{s}_2(y,x,w^*) &= \frac{1}{f_{X,W^*}(x,w^*)f_{Y,X,W^*}(y,x,w^*)}\left(\frac{\partial f_{X,W^*}(x,w^*)}{\partial x}+\frac{\partial f_{X,W^*}(x,w^*)}{\partial w^*}\frac{1}{\tilde{\Psi}_{(1)}(f)}\right) \\
\tilde{s}_3(y,x,w^*) &= \frac{F_{Y|X=x,W^*=w^*}(y)}{f_{Y,X,W^*}(y,x,w^*)} \\
\tilde{s}_4(y,x,w^*) &= -\frac{1}{f_{Y,X,W^*}(y,x,w^*)} \\
\tilde{s}_5(y,x,w^*) &= -\frac{1}{f^2(y,x,w^*)}\left[F_{Y|X=x,W^*=w^*}(y)\frac{\partial f_{X,W^*}(x,w^*)}{\partial x}-\int_{-\infty}^{y}\frac{\partial f_{Y,X,W^*}(y,x,w^*)}{\partial x}dy\right.
\end{aligned}$$

[5]Linear in terms of $\hat{E}\left[W_1e^{i\xi W_2}\right]$, $\hat{E}\left[e^{i\xi W_2}\right]$ and $\hat{E}\left[e^{i\xi W_2}\frac{1}{h_{2,1}^{1+\lambda_{2,1}}h_{2,2}^{1+\lambda_{2,2}}}G_Y\left(\frac{y-Y}{h_{2,1}}\right)G_X^{(\lambda_{2,2})}\left(\frac{x-X}{h_{2,2}}\right)\right]$

$$+ \left(F_{Y|X=x,W^*=w^*}(y) \frac{\partial f_{X,W^*}(x,w^*)}{\partial w^*} - \int_{-\infty}^{y} \frac{\partial f_{Y,X,W^*}(y,x,w^*)}{\partial w^*} dy \right) \frac{1}{\tilde{\Psi}_{(1)}(f)} \Bigg]$$

$$\tilde{s}_6(y,x,w^*) = \frac{F_{Y|X=x,W^*=w^*}(y)}{f_{Y,X,W^*}(y,x,w^*)} \frac{1}{\tilde{\Psi}_{(1)}(f)}$$

$$\tilde{s}_7(y,x,w^*) = -\frac{1}{f_{Y,X,W^*}(y,x,w^*)} \frac{1}{\tilde{\Psi}_{(1)}(f)}$$

$$\tilde{s}_8(y,x,w^*) = -\frac{F_{X|W^*=w^*}(x) \frac{\partial f_{W^*}(w^*)}{\partial w^*}}{f_{W^*}(w^*) f_{X,W^*}(x,w^*)} \times \frac{\Psi_{(2)}(f)}{\tilde{\Psi}^2_{(1)}(f)}$$

$$\tilde{s}_9(y,x,w^*) = \frac{\frac{\partial f_{W^*}(w^*)}{\partial w^*}}{f_{W^*}(w^*) f_{X,W^*}(x,w^*)} \times \frac{\Psi_{(2)}(f)}{\tilde{\Psi}^2_{(1)}(f)}$$

$$\tilde{s}_{10}(y,x,w^*) = \frac{F_{X|W^*=w^*}(x)}{f_{X,W^*}(x,w^*)} \times \frac{\Psi_{(2)}(f)}{\tilde{\Psi}^2_{(1)}(f)}$$

$$\tilde{s}_{11}(y,x,w^*) = -\frac{1}{f_{X,W^*}(x,w^*)} \times \frac{\Psi_{(2)}(f)}{\tilde{\Psi}^2_{(1)}(f)},$$

where $\alpha(f) \equiv F^{-1}_{Y|X=x,W^*=w^*}(\delta)$, $\Phi_{(2)}(f) \equiv \frac{\partial F^{-1}_{Y|X=x,W^*=w^*}(\delta)}{\partial w^*}$,
$\Psi_{(2)}(f) \equiv -\left. \frac{\partial F_{Y|X=x,W^*=w^*}(y)}{\partial w^*} \middle/ \frac{\partial F_{Y|X=x,W^*=w^*}(y)}{\partial y} \right.$, *and* $\tilde{\Psi}_{(1)}(f) = -\left. \frac{\partial F_{X|W^*=w^*}(x)}{\partial w^*} \middle/ \frac{\partial F_{X|W^*=w^*}(x)}{\partial x} \right.$.

Proof of Lemma 18. By (11), for any value $(y,x,w^*) \in \mathbb{S}_\tau$, I have that

$$\widehat{\rho(y,x)} - \rho(y,x) = D\tilde{\Xi}_n + R\tilde{\Xi}_n.$$

By Lemma 12,

$$D\tilde{\Xi}_n = \sum_{k=1}^{11} \tilde{s}_k(x,w^*,\delta) D\tilde{\Xi}_{n,k},$$

where $\tilde{s}_k(x,w^*,\delta)$ is as stated in the Lemma. and

$$D\tilde{\Xi}_{n,1} = \int \left(\hat{g}_{0,0,0}(y,x,w^*,h_n) - g_{0,0,0}(y,x,w^*,h_n) \right) dy$$

$$D\tilde{\Xi}_{n,2} = \int_{-\infty}^{y} \left(\hat{g}_{0,0,0}(y,x,w^*,h_n) - g_{0,0,0}(y,x,w^*,h_n) \right) dy$$

$$D\tilde{\Xi}_{n,3} = \int \left(\hat{g}_{0,1,1}(y,x,w^*,h_n) - g_{0,1,1}(y,x,w^*,h_n) \right) dy$$

$$D\tilde{\Xi}_{n,4} = \int_{-\infty}^{y} \left(\hat{g}_{0,1,1}(y,x,w^*,h_n) - g_{0,1,1}(y,x,w^*,h_n) \right) dy$$

$$D\tilde{\Xi}_{n,5} = \hat{g}_{0,0,0}(y,x,w^*,h_n) - g_{0,0,0}(y,x,w^*,h_n)$$

$$D\tilde{\Xi}_{n,6} = \int \left(\hat{g}_{1,0,0}(y,x,w^*,h_n) - g_{1,0,0}(y,x,w^*,h_n) \right) dy$$

$$D\tilde{\Xi}_{n,7} = \int_{-\infty}^{y} \left(\hat{g}_{1,0,0}(y,x,w^*,h_n) - g_{1,0,0}(y,x,w^*,h_n) \right) dy$$

$$D\tilde{\Xi}_{n,8} = \int \left(\hat{\tilde{g}}_{0,0}(s,w^*,h_n) - \tilde{g}_{0,0}(s,w^*,h_n)\right) ds$$
$$D\tilde{\Xi}_{n,9} = \int_{-\infty}^{x} \left(\hat{\tilde{g}}_{0,0}(s,w^*,h_n) - \tilde{g}_{0,0}(s,w^*,h_n)\right) ds$$
$$D\tilde{\Xi}_{n,10} = \int \left(\hat{\tilde{g}}_{1,0}(s,w^*,h_n) - \tilde{g}_{1,0}(s,w^*,h_n)\right) ds$$
$$D\tilde{\Xi}_{n,11} = \int_{-\infty}^{x} \left(\hat{\tilde{g}}_{1,0}(s,w^*,h_n) - \tilde{g}_{1,0}(s,w^*,h_n)\right) ds.$$

By Lemma 5, a first order expansion of $D\tilde{\Xi}_n$ can be written as

$$\begin{aligned} D\tilde{\Xi}_n &= BD\tilde{\Xi}_n + LD\tilde{\Xi}_n + RD\tilde{\Xi}_n \\ &= BD\tilde{\Xi}_n + \sum_{k=1}^{11} \tilde{s}_k(x,w^*,\delta) LD\tilde{\Xi}_{n,k} + RD\tilde{\Xi}_n \end{aligned}$$

where

$$\begin{aligned}
LD\tilde{\Xi}_{n,1} &= \int L_{0,0,0}(y,x,w^*,h_n)dy = \int \hat{E}\left[l_{0,0,0}(y,x,w^*,h_n)\right]dy = \hat{E}\left[\int l_{0,0,0}(y,x,w^*,h_n)dy\right] \\
LD\tilde{\Xi}_{n,2} &= \int_{-\infty}^{y} L_{0,0,0}(y,x,w^*,h_n)dy = \int_{-\infty}^{y} \hat{E}\left[l_{0,0,0}(y,x,w^*,h_n)\right]dy \\
&= \hat{E}\left[\int_{-\infty}^{y} l_{0,0,0}(y,x,w^*,h_n)dy\right] \\
LD\tilde{\Xi}_{n,3} &= \int L_{0,0,1}(y,x,w^*,h_n)dy = \int \hat{E}\left[l_{0,0,1}(y,x,w^*,h_n)\right]dy \\
&= \hat{E}\left[\int l_{0,0,1}(y,x,w^*,h_n)dy\right] \\
LD\tilde{\Xi}_{n,4} &= \int_{-\infty}^{y} L_{0,0,1}(y,x,w^*,h_n)dy = \int_{-\infty}^{y} \hat{E}\left[l_{0,0,1}(y,x,w^*,h_n)\right]dy \\
&= \hat{E}\left[\int_{-\infty}^{y} l_{0,0,1}(y,x,w^*,h_n)dy\right] \\
LD\tilde{\Xi}_{n,5} &= L_{0,0,0}(y,x,w^*,h_n) = \hat{E}\left[l_{0,0,0}(y,x,w^*,h_n)\right] \\
LD\tilde{\Xi}_{n,6} &= \int L_{1,0,0}(y,x,w^*,h_n)dy = \int \hat{E}\left[l_{1,0,0}(y,x,w^*,h_n)\right]dy \\
&= \hat{E}\left[\int l_{1,0,0}(y,x,w^*,h_n)dy\right] \\
LD\tilde{\Xi}_{n,7} &= \int_{-\infty}^{y} L_{1,0,0}(y,x,w^*,h_n)dy = \int_{-\infty}^{y} \hat{E}\left[l_{1,0,0}(y,x,w^*,h_n)\right]dy \\
&= \hat{E}\left[\int_{-\infty}^{y} l_{1,0,0}(y,x,w^*,h_n)dy\right] \\
LD\tilde{\Xi}_{n,8} &= \int \tilde{L}_{0,0}(s,w^*,h_n)ds = \int \hat{E}\left[\tilde{l}_{0,0}(s,w^*,h_n)\right]ds = \hat{E}\left[\int \tilde{l}_{0}(s,w^*,h_n)ds\right] \\
LD\tilde{\Xi}_{n,9} &= \int_{-\infty}^{x} \tilde{L}_{0,0}(s,w^*,h_n)ds = \int_{-\infty}^{x} \hat{E}\left[\tilde{l}_{0,0}(s,w^*,h_n)\right]ds = \hat{E}\left[\int_{-\infty}^{x} \tilde{l}_{0,0}(s,w^*,h_n)ds\right]
\end{aligned}$$

$$LD\tilde{\Xi}_{n,10} = \int \tilde{L}_{1,0}(s,w^*,h_n)ds = \int \hat{E}\left[\tilde{l}_{1,0}(s,w^*,h_n)\right]ds = \hat{E}\left[\int \tilde{l}_{1,0}(s,w^*,h_n)ds\right]$$

$$LD\tilde{\Xi}_{n,11} = \int_{-\infty}^{x} \tilde{L}_{1,0}(s,w^*,h_n)ds = \int_{-\infty}^{x} \hat{E}\left[\tilde{l}_{1,0}(s,w^*,h_n)\right]ds = \hat{E}\left[\int_{-\infty}^{x} \tilde{l}_{1,0}(s,w^*,h_n)ds\right].$$

Then I can write

$$\begin{aligned} LD\tilde{\Xi}_n &= \hat{E}\left[lD\tilde{\Xi}_n\right] \\ &= \hat{E}\left[\sum_{k=1}^{11} \tilde{s}_k(x,w^*,\delta) lD\tilde{\Xi}_{n,k}\right], \end{aligned}$$

where $lD\tilde{\Xi}_n$ are as stated in the Lemma. ■

Lemma 19. *Suppose the conditions of Lemma 5 hold. (i) Then for each $(y,x,w^*) \in \mathbb{S}_\tau$, $E[LD\tilde{\Xi}_n] = 0$, and if Assumption 12 also holds, then $E[LD\tilde{\Xi}_n^2] = n^{-1}\tilde{\Omega}(y,x,w^*,h)$, where $\tilde{\Omega}(y,x,w^*,h) \equiv E\left[lD\tilde{\Xi}_n^2\right] < \infty$.*

(ii) If Assumption 12 also holds, and if for each $(y,x,w^*) \in \mathbb{S}_\tau$, $\tilde{\Omega}(y,x,w^*,h_n) > 0$ for all n sufficiently large, then for each $(y,x,w^*) \in \mathbb{S}_\tau$

$$n^{1/2}\left(\tilde{\Omega}(y,x,w^*,h_n)\right)^{-1/2} LD\tilde{\Xi}_n \xrightarrow{d} N(0,1).$$

Proof of Lemma 19. (i) The fact that $E\left[LD\tilde{\Xi}_n\right] = 0$ follows directly from Theorem 2. Next I show that for all $(y,x,w^*) \in \mathbb{S}_\tau$, $\tilde{\Omega}(y,x,w^*,h) < \infty$. It follows by Cauchy Schwartz that

$$\begin{aligned} \tilde{\Omega}(y,x,w^*,h_n) &\equiv E\left[lD\tilde{\Xi}_n^2\right] \\ &\preceq \sum_{k=1}^{11} \tilde{s}_k^2(y,x,w^*) E\left[lD\tilde{\Xi}_{n,k}^2\right]. \end{aligned}$$

Note that

$$\begin{aligned} E\left[lD\tilde{\Xi}_{n,1}^2\right] &= E\left[\left(\int l_{0,0,0}(y,x,w^*,h_n)dy\right)^2\right] \\ &\le E\left[\int \left(l_{0,0,0}(y,x,w^*,h_n)\right)^2 dy\right] \\ &= \int E\left[\left(l_{0,0,0}(y,x,w^*,h_n)\right)^2\right] dy \end{aligned}$$

where the second line follows from Jensen's Inequality and the third line follows from Tonelli's Theorem. By Theorem 2, $\int E\left[(l_{0,0,0}(y,x,w^*,h))^2\right]dy < \infty$ for each h. Conclusions for other values of $k \in \{1,2,3,...,11\}$ follows similarly. Then by Assumption 6, I have that $\tilde{\Omega}(y,x,w^*,h) < \infty$ for all values $(y,x,w^*) \in \mathbb{S}_\tau$.

(ii) To show asymptotic normality, I apply Lemma 16 to $lD\tilde{\Xi}_n = \sum_{k=1}^{11} \tilde{s}_k(x,w^*,\delta) lD\tilde{\Xi}_{n,k}$. By previous argument, $\tilde{\Omega}(y,x,w^*,h) < \infty$ for all values $(y,x,w^*) \in \mathbb{S}_\tau$. I've assumed that for n sufficiently large, $\tilde{\Omega}(y,x,w^*,h_n) > 0$. The conclusion follows. ∎

Finally, I state the asymptotic normality result for the estimator of $\rho(y,x)$ under a high-level assumption:

Assumption 17. *(High Level) For given $\lambda_1, \lambda_{2,1}, \lambda_{2,2} \in \{0,1\}$ and given $j \in \{1,2\}$, $h_n \to 0$ at a rate such that for each $(y,x,w^*) \in \mathbb{S}_\tau$ such that $\tilde{\Omega}(y,x,w^*,h_n) > 0$ for all n sufficiently large, I have $n^{1/2}\left(\tilde{\Omega}(y,x,w^*,h_n)\right)^{-1/2}\left|BD\tilde{\Xi}_n\right| \xrightarrow{p} 0$ and $n^{1/2}\left(\tilde{\Omega}(y,x,w^*,h_n)\right)^{-1/2}\left|RD\tilde{\Xi}_n\right| \xrightarrow{p} 0$.*

Theorem 5. *If the conditions of Theorem 2 and Assumption 16, 17 hold, then for each $(y,x,w^*) \in \mathbb{S}_\tau$ such that $\tilde{\Omega}(y,x,w^*,h_n) > 0$ for all n sufficiently large I have*

$$n^{1/2}\left(\tilde{\Omega}(y,x,w^*,h_n)\right)^{-1/2}\left(\widehat{\rho(y,x)} - \rho(y,x)\right) \xrightarrow{d} N(0,1).$$

References

Altonji, Joseph G and Matzkin, Rosa L (2005) "Cross section and panel data estimators for nonseparable models with endogenous regressors," *Econometrica*, Vol. 73, pp. 1053–1102.

Andrews, Donald WK (1995) "Nonparametric kernel estimation for semiparametric models," *Econometric Theory*, Vol. 11, pp. 560–586.

Aradillas-Lopez, Andres (2022) "Inference in models with partially identified control functions."

Blundell, Richard, Chen, Xiaohong, and Kristensen, Dennis (2007) "Semi-nonparametric IV estimation of shape-invariant Engel curves," *Econometrica*, Vol. 75, pp. 1613–1669.

Brancaccio, Giulia, Kalouptsidi, Myrto, and Papageorgiou, Theodore (2020) "Geography, transportation, and endogenous trade costs," *Econometrica*, Vol. 88, pp. 657–691.

Chesher, Andrew (2003) "Identification in nonseparable models," *Econometrica*, Vol. 71, pp. 1405–1441.

Fan, Jianqing (1991a) "Asymptotic normality for deconvolution kernel density estimators," *Sankhyā: The Indian Journal of Statistics, Series A*, pp. 97–110.

——— (1991b) "On the optimal rates of convergence for nonparametric deconvolution problems," *The Annals of Statistics*, pp. 1257–1272.

Fan, Jianqing and Truong, Young K (1993) "Nonparametric regression with errors in variables," *The Annals of Statistics*, pp. 1900–1925.

Hansen, Bruce E (2008) "Uniform convergence rates for kernel estimation with dependent data," *Econometric Theory*, Vol. 24, pp. 726–748.

Imbens, Guido W and Newey, Whitney K (2009) "Identification and estimation of triangular simultaneous equations models without additivity," *Econometrica*, Vol. 77, pp. 1481–1512.

Li, Tong (2002) "Robust and consistent estimation of nonlinear errors-in-variables models," *Journal of Econometrics*, Vol. 110, pp. 1–26.

Matzkin, Rosa L (2003) "Nonparametric estimation of nonadditive random functions," *Econometrica*, Vol. 71, pp. 1339–1375.

——— (2015) "Estimation of nonparametric models with simultaneity," *Econometrica*, Vol. 83, pp. 1–66.

——— (2016) "On independence conditions in nonseparable models: Observable and unobservable instruments," *Journal of Econometrics*, Vol. 191, pp. 302–311.

Politis, Dimitris N and Romano, Joseph P (1999) "Multivariate density estimation with general flat-top kernels of infinite order," *Journal of Multivariate Analysis*, Vol. 68, pp. 1–25.

Schennach, Susanne M (2004a) "Estimation of nonlinear models with measurement error," *Econometrica*, Vol. 72, pp. 33–75.

——— (2004b) "Nonparametric regression in the presence of measurement error," *Econometric Theory*, Vol. 20, pp. 1046–1093.

Schennach, Susanne, White, Halbert, and Chalak, Karim (2012) "Local indirect least squares and average marginal effects in nonseparable structural systems," *Journal of Econometrics*, Vol. 166, pp. 282–302.

Shaikh, Azeem M and Vytlacil, Edward J (2011) "Partial identification in triangular systems of equations with binary dependent variables," *Econometrica*, Vol. 79, pp. 949–955.

Song, Suyong, Schennach, Susanne M, and White, Halbert (2015) "Estimating nonseparable models with mismeasured endogenous variables," *Quantitative Economics*, Vol. 6, pp. 749–794.