



# Identification in nonseparable models with measurement errors and endogeneity<sup>☆</sup>



Yingyao Hu<sup>a</sup>, Ji-Liang Shiu<sup>b</sup>, Tiemen Woutersen<sup>c,\*</sup>

<sup>a</sup> Johns Hopkins University, United States

<sup>b</sup> Renmin University of China, China

<sup>c</sup> University of Arizona, United States

## ARTICLE INFO

### Article history:

Received 2 August 2015

Received in revised form

13 March 2016

Accepted 10 April 2016

Available online 20 April 2016

### JEL classification:

C20

J20

C27

### Keywords:

Nonclassical measurement error

Measurement error and endogeneity

Labor supply elasticity

## ABSTRACT

Economic variables are often measured with errors and may be endogenous. This paper extends Chesher (2003) and gives new identification results for the ratio of partial effects in a class of nonseparable index models with measurement error and endogeneity. The identification restrictions include a triangular system and the derivative of some conditional mean functions being nonzero. An example that motivates the paper is identification of the labor supply elasticity.

© 2016 Elsevier B.V. All rights reserved.

## 1. Introduction

ECONOMIC VARIABLES ARE OFTEN MEASURED WITH ERRORS and may be endogenous. For example, when estimating the labor supply elasticity, it is likely that the wage and the number of hours worked are measured with errors.<sup>1</sup> In this paper, we use the estimation of the labor supply elasticity as a leading example, but the estimation procedure is more general than that. We consider models with measurement error and endogeneity. In particular, let

$$y = m(\theta x^* + w, w', \eta), \quad (1)$$

$$x = x^* + \varepsilon, \quad (2)$$

$$x^* = g(z, w, w', u), \quad (3)$$

where we observe  $\{y, x, z, w, w'\}$ . The model involves a dependent variable  $y$ , a true regressor  $x^*$ , a mismeasured regressor  $x$ , cor-

rectly measured regressors  $w$  and  $w'$ , and an instrument  $z$ , while the unobservables are  $\eta$ ,  $\varepsilon$ , and  $u$ . We do not assume that  $u$  and  $(z, w)$  are independent. Moreover we do not impose a restriction on the dimension of the disturbance  $\eta$  in the regression function or on  $w'$ . The true regressor  $x^*$  is endogenous in the sense that it is determined by  $g(z, w, w', u)$  while  $u$  and  $\eta$  are generally correlated. The measurement error  $\varepsilon$  may be correlated with  $\eta$  so that Hu (2008) and Hu and Schennach (2008) do not apply. Three related papers are Chesher (2003), Abrevaya et al. (2010) and Hu et al. (2015). Our results differ from Chesher (2003) since we allow for measurement error but we have less to say about quantiles. Abrevaya et al. (2010) consider estimation up to scale in the transformation model in order to test the null hypothesis of no causal effect. Our work differs from theirs in that we allow the regressors to be measured with error and the model to be nonseparable and our identification result immediately implies a test in a more general framework. Hu et al. (2015) consider a separable model and provide a literature review. All the identification results in this paper are constructive in the sense that the results suggest an estimator, which we demonstrate by doing simulations.

The rest of the paper is organized as follows. In Section 2 we extend the model of Eqs. (1)–(3) with a triangular system and present the identification result. Section 3 contains results from Monte Carlo experiments and Section 4 concludes. The

<sup>☆</sup> We thank an editor and a referee for excellent comments and we thank Maddie Pickens for research assistance.

\* Correspondence to: Department of Economics, Eller College of Management, University of Arizona, P.O. Box 210108, Tucson, AZ 8572, United States.

E-mail address: [woutersen@email.arizona.edu](mailto:woutersen@email.arizona.edu) (T. Woutersen).

<sup>1</sup> Borjas (2009) gives an overview of empirical studies that estimate the labor supply elasticity and also discusses the problems caused by measurement error.

detailed mathematical proof of the main result is provided in the [Appendix](#).

**2. Identification in nonseparable models containing a triangular system**

We are interested in the parameter  $\theta$  of the nonseparable regression model,

$$y = m(\theta x^* + w, w', \eta), \tag{4}$$

$$x = x^* + \varepsilon, \tag{5}$$

$$x^* = g(z, w, w', u), \tag{6}$$

where we observe  $\{y, x, z, w, w'\}$ . As discussed in the introduction, [Hu et al. \(2015\)](#) assume that the observable variables are independent of all error terms, i.e.  $(z, w) \perp (\eta, \varepsilon, u)$  in Eqs. (4)–(6). However, in many cases in economics, establishing the marginal independence in Eq. (6) such as  $(z, w) \perp u$  may require undesirable strong assumptions.<sup>2</sup> Worker’s labor supply function can be a motivating example of the nonseparable model to illustrate this case. Suppose worker’s labor supply function can be written as Eq. (4). Also, the dependent variable  $y$  could denote hours worked,  $x^*$  the unobserved true wage rate, and  $\varepsilon$  the measurement error. In addition,  $w$  could be a measure of worker productivity,  $z$  education, and  $w'$  a list of covariates such as demographics, housing, etc. While the error term  $\eta$  can be interpreted as unobserved health status, the error term  $u$  can represent job training possibilities at firms. Under the empirical setting, the parameter of interest  $\theta$  is the ratio of partial effects corresponding to the true wage rate and a measure of worker productivity on hours worked. While Eq. (4) is the hours worked equation, Eq. (6) is the wage equation. In this case, a potential correlation between education and job training possibilities or worker productivity and job training possibilities would violate the marginal independence assumption  $(z, w) \perp u$ . To allow  $(z, w)$  and  $u$  to be correlated we extend the nonseparable model (4)–(6) with the following marginal independence and triangular system:

**Assumption 2.1.** The variables  $(z, w, w')$  are jointly independent of the error terms  $(\eta, \varepsilon)$ , i.e.

$$(z, w, w') \perp (\eta, \varepsilon).$$

Each  $(z, w)$  contains a continuous element and the functions  $m(\cdot, \cdot, \cdot)$ , and  $g(\cdot, \dots, \cdot)$ , are differentiable. In addition, the function  $m(\cdot, \cdot, \cdot)$  is strictly monotone in  $\eta$  and the derivative of its first argument is nonzero, i.e.  $\frac{\partial m(\cdot, \cdot, \cdot)}{\partial \eta} \neq 0$  where  $\psi$  denotes the first argument of  $m(\cdot, \cdot, \cdot)$ . The function  $g(\cdot, \dots, \cdot)$  is strictly monotone in  $u$ .

These monotonicity conditions can be illustrated by the motivating example of labor supply and the wage equation. In particular, holding other factors fixed, good health may increase the hours worked and job training possibilities may increase the unobserved true wage rate.

**Assumption 2.2.** Let  $v$  denote an unobservable error term. Let

$$w = h_1(z, w', u, v), \tag{7}$$

$$z = h_2(w', v), \tag{8}$$

where  $h_1$ , and  $h_2$  are differentiable,  $h_1$  is strictly monotone in its argument  $u$ , and  $h_2$  is strictly monotone in its argument  $v$ . No restrictions are imposed on  $v$ .

<sup>2</sup> For example [Newey \(2001\)](#) argues that independence between  $z, w$  and  $u$  is a strong assumption.

If we interpret the unobservable error term  $v$  as ability then the motivating example satisfies these monotonicity assumptions because ability increases the observable wage rate and education ceteris paribus. The modeling in Eq. (7) allows for correlation between  $w$  and  $u$ .

**Assumption 2.3.** Assume the cumulative distribution functions  $F_{Y|z, w, w'}(y|z, w, w')$  and  $F_{X|z, w, w'}(x|z, w, w')$  are differentiable with respect to  $z$  and  $w$ . For some  $(\bar{y}, \bar{x})$  and  $(z, w, w')$  let

$$\frac{\partial F_{Y|z, w, w'}(\bar{y})}{\partial z} \frac{\partial F_{X|z, w, w'}(\bar{x})}{\partial w} - \frac{\partial F_{Y|z, w, w'}(\bar{y})}{\partial w} \frac{\partial F_{X|z, w, w'}(\bar{x})}{\partial z} \neq 0.$$

Also, at the point at which these derivatives are evaluated,  $(z, w, w')$ , let (i) the density of  $(z, w, w')$  be strictly positive or (ii) the conditional density of  $(z, w)$  given  $w'$  be strictly positive and  $P(W' = w') > 0$ .

Our main result can be stated as follows and we leave the proof to the [Appendix](#):

**Theorem 2.1.** Suppose that Assumptions 2.1, 2.2, and 2.3 hold. Then the parameter of interest  $\theta$  is identified. In particular, any  $(\bar{y}, \bar{x})$  and  $(z, w, w')$

$$\theta = \frac{\frac{\partial F_{Y|z, w, w'}(\bar{y})}{\partial z} \frac{\partial F_{X|z, w, w'}(\bar{x})}{\partial x}}{\frac{\partial F_{Y|z, w, w'}(\bar{y})}{\partial z} \frac{\partial F_{X|z, w, w'}(\bar{x})}{\partial w} - \frac{\partial F_{Y|z, w, w'}(\bar{y})}{\partial w} \frac{\partial F_{X|z, w, w'}(\bar{x})}{\partial z}}. \tag{9}$$

In order to estimate  $\theta$  we have to use estimators for the derivatives of the cumulative distribution functions. This would cause two types of bias to  $\theta$ , one is the approximation bias for these derivatives and the other one is ratio bias since the denominator in Eq. (7) may be close to zero. To mitigate these biases, in the estimation section, we propose to use a weighted average over sample points. Also, the condition in [Assumption 2.3](#) is expressed in terms of distribution functions so the restriction can be tested. In particular, if [Assumption 2.3](#) holds for more than one value  $(\bar{y}, \bar{x})$  then we can compute an empirical estimator for the term  $\theta$  for different values of  $(\bar{y}, \bar{x})$  to test the model.

**3. Estimation and simulation**

Under the assumptions of [Theorem 2.1](#), the parameter of interest  $\theta$  can be identified and written in terms of the first derivatives of the distribution functions  $F_{X|z, w, w'}$  and  $F_{Y|z, w, w'}$ . From an i.i.d. random sample  $\{y_i, x_i, z_i, w_i, w'_i\}_{i=1}^n$ , the empirical estimator for  $F_{Y|z, w, w'}$  is

$$\widehat{F}(y|z, w, w') = \frac{1}{n} \sum_{i=1}^n \left[ \text{number of } y_1, \dots, y_n \text{ less than } y \text{ such that } |z_i - z| < \sigma \cdot \text{std}(z), |w_i - w| < \sigma \cdot \text{std}(w), |w'_i - w'| < \sigma \cdot \text{std}(w') \right], \tag{10}$$

where  $\text{std}(z)$ ,  $\text{std}(w)$ , and  $\text{std}(w')$  are standard deviations of  $z_i$ ,  $w_i$ , and  $w'_i$  in the sample respectively and  $\sigma$  is a bandwidth that converges to zero as the sample size increases. This implies that a consistent estimator for the derivative of  $F_{Y|z, w, w'}$  with respect to  $z$  can be obtained by

$$\frac{\partial \widehat{F}_{Y|z, w, w'}(y, z, w, w')}{\partial z} = \frac{1}{2\delta} [\widehat{F}(y|z + \delta, w, w') - \widehat{F}(y|z - \delta, w, w')]$$

using a small value of  $\delta$ . A similar formula can be applied to estimate  $F_{X|z, w, w'}$  and its derivatives.

**Table 1**  
Estimations of single linear index models ( $n = 1000$ ).

		Separable estimator $\theta = -1$	Nonseparable estimator $\theta = -1$
Simulation I	Mean	-1.029	-0.987
	Median	-1.080	-0.960
	Std. dev.	0.623	0.351
Simulation II	Mean	-0.225	-0.979
	Median	-0.216	-0.932
	Std. dev.	0.075	0.183
Simulation III	Mean	0.282	-1.062
	Median	-0.076	-1.060
	Std. dev.	4.730	0.433
Simulation IV	Mean	-0.303	-1.069
	Median	-0.301	-1.018
	Std. dev.	0.195	0.180

Note: Standard deviations of the parameters are computed by the standard deviation of the estimates across 100 simulations and called (simulation) standard deviations. While the Separable Estimator indicates the local polynomial regression method in Hu et al. (2015) the Nonseparable Estimator is the proposed estimator in Section 3. The abbreviated names reflects whether the function  $g$  in Eq. (3) is nonseparable or not.

Because the formula for  $\theta$  in Eq. (9) works for the domain of  $(y, x, z, w, w')$ , we can use

$$\theta = \int \dots \int \frac{\frac{\partial F_{Y|Z,W,W'}(y)}{\partial z} \frac{\partial F_{X|Z,W,W'}(x)}{\partial x}}{\frac{\partial F_{Y|Z,W,W'}(y)}{\partial z} \frac{\partial F_{X|Z,W,W'}(x)}{\partial w} - \frac{\partial F_{Y|Z,W,W'}(y)}{\partial w} \frac{\partial F_{X|Z,W,W'}(x)}{\partial z}} \times \omega(y, x, z, w, w') dy dx dz dw dw'$$

where  $\omega(y, x, z, w, w')$  is a weighting function that is zero if the denominator in the last equation is zero. If we use the empirical distribution of  $(y, x, z, w, w')$  as our weighting function and the denominator is nonzero, a sample counterpart estimator of  $\theta$  is given in Box I. Because the estimator is constructed using empirical estimators for the distribution functions and their derivatives, the proposed estimator should be consistent under general conditions such as those stated by Hu et al. (2015). We illustrate that the identification results are constructive and investigate the finite sample properties of an estimator through a Monte Carlo simulation. In particular, we conducted four sets of simulations. In these experiments the data generating process (DGP) has that  $w$  and  $u$  are correlated. We assume that the variables  $x, w', u, v$ , and  $\eta$  are generated using a Gaussian process with mean zero and unit variance. The four sets of simulations for the functions  $m, g, h_1$ , and  $h_2$  are given in Box II. These simulations are intended to capture the different degrees of separability in the functions  $m$ , and  $g$ . In Simulation I,  $m$ , and  $g$  are both separable,  $g$  in Simulation II and III is nonseparable and  $m$ , and  $g$  are both nonseparable in Simulation IV. We use 100 replications for each experiment with a sample size 1000, and report mean, median, and standard deviation for the Separable Estimator in Hu et al. (2015) and the Nonseparable Estimator (the estimator suggested by the identification results).

The estimation results are summarized in Table 1 and clearly show that the Nonseparable Estimator performs well in all experiments. As expected, the Separable Estimator in Hu et al. (2015) causes larger bias in Simulations II, III and IV. When  $m$  and  $g$  are both separable in Simulation I, both estimators have small bias. However the Nonseparable Estimator suggested in this paper has a lower standard deviation.

#### 4. Conclusion

Many empirical problems in economics have some variables that are measured with error and/or are endogenous. This paper gives new identification results for the ratio of partial effects in a class of nonseparable index models with measurement error and

endogeneity. The identification restrictions include a triangular system and the derivative of some conditional mean function being nonzero. The results are more general than Chesher (2003) in some aspects since we allow for measurement error but the results are more restrictive in other aspects. Measurement error and endogeneity are features of many empirical problems and we discussed the identification of the labor supply elasticity as an example.

#### Appendix. Proof of Theorem 2.1

With the help of Eqs. (7)–(8) from Assumption 2.2, the model (4)–(6) can be divided into two triangular systems, (A) and (B):

$$\begin{aligned} y &= m(\theta g(z, w, w', u) + w, w', \eta) \\ w &= h_1(z, w', u, v) \\ z &= h_2(w', v), \end{aligned} \tag{A}$$

and

$$\begin{aligned} x &= g(z, w, w', u) + \varepsilon \\ w &= h_1(z, w', u, v) \\ z &= h_2(w', v). \end{aligned} \tag{B}$$

We apply the identification technique developed in Chesher (2003) to these triangular systems. Since  $h_2$  is strictly monotone in  $v$  from Assumption 2.2, for a given  $w'$  the inverse of  $h_2$  exists and is denoted by  $\tilde{h}_2(w', z) = v$  for a fixed  $w'$  and we call this function the inverse of  $h_2$  conditional on  $w'$ . Substituting it in  $h_1$ , we can have  $w = h_1(z, w', u, \tilde{h}_2(w', z))$  conditional on  $z, w'$ . That  $h_1$  is strictly monotone in  $u$  implies that the inverse function  $\tilde{h}_1(z, w', w) = u$  exists. Plugging this equation into  $m$ , we can have that  $y = m(\theta g(z, w, w', \tilde{h}_1(z, w', w)) + w, w', \eta)$  conditional on  $z, w$ , and  $w'$ . Define

$$n_1(z, w, w', \eta) = m(\theta g(z, w, w', \tilde{h}_1(z, w', w)) + w, w', \eta).$$

Taking derivatives with respect to  $z, w$  on both sides of the above equation, and using the independence assumption in Assumption 2.1 ( $z, w, w' \perp \eta$ ), we obtain

$$\frac{\partial n_1(z, w, w', \eta)}{\partial w} = \frac{\partial m(\theta g(z, w, w', \tilde{h}_1(z, w', w)) + w, w', \eta)}{\partial \psi} \times \left[ \theta \left( g_w + g_u \frac{\partial \tilde{h}_1}{\partial w} \right) + 1 \right], \tag{11}$$

$$\frac{\partial n_1(z, w, w', \eta)}{\partial z} = \frac{\partial m(\theta g(z, w, w', \tilde{h}_1(z, w', w)) + w, w', \eta)}{\partial \psi} \times \left[ \theta \left( g_z + g_u \frac{\partial \tilde{h}_1}{\partial z} \right) \right]. \tag{12}$$

Because  $\frac{\partial m(\dots)}{\partial \psi} \neq 0$  from Assumption 2.1, we can cross multiply the two equations above to obtain

$$\begin{aligned} \frac{\partial n_1(z, w, w', \eta)}{\partial w} \left[ \theta \left( g_z + g_u \frac{\partial \tilde{h}_1}{\partial z} \right) \right] \\ = \frac{\partial n_1(z, w, w', \eta)}{\partial z} \left[ \theta \left( g_w + g_u \frac{\partial \tilde{h}_1}{\partial w} \right) + 1 \right]. \end{aligned} \tag{13}$$

Using the independence assumption,  $(z, w, w') \perp \eta$ , again, and using the strictly monotone assumption of  $n_1(z, w, w', \eta)$  in  $\eta$  from the monotone property of  $m$  in  $\eta$  in Assumption 2.1 yields

$$\begin{aligned} F_\eta(\bar{\eta}) &= F_{\eta|Z,W,W'}(\bar{\eta}) \\ &= Pr(\eta < \bar{\eta} | Z, W, W') \\ &= Pr(n_1(z, w, w', \eta) < n_1(z, w, w', \bar{\eta}) | Z, W, W') \end{aligned}$$

$$\hat{\theta} = \frac{1}{n} \sum_i \frac{\frac{\partial F_{Y|Z,W,W'}(y_i, z_i, w_i, w'_i)}{\partial z} \frac{\partial F_{X|Z,W,W'}(x_i, z_i, w_i, w'_i)}{\partial x}}{\frac{\partial F_{Y|Z,W,W'}(y_i, z_i, w_i, w'_i)}{\partial z} \frac{\partial F_{X|Z,W,W'}(x_i, z_i, w_i, w'_i)}{\partial w} - \frac{\partial F_{Y|Z,W,W'}(y_i, z_i, w_i, w'_i)}{\partial w} \frac{\partial F_{X|Z,W,W'}(x_i, z_i, w_i, w'_i)}{\partial z}}$$

**Box I.**

Simulation I :  $m(\theta x^* + w, w', \eta) = 0.1(\theta x^* + w + w') + \eta, g(z, w, w', u) = 0.1(z + w + w' + u)$

$h_1(z, w', u, v) = z + w' + u + v, h_2(w', v) = 0.1(w' + v),$

Simulation II :  $m(\theta x^* + w, w', \eta) = 5 \exp(\theta x^* + w + w') + \eta, g(z, w, w', u) = 0.5(z + w + w')^2 \cdot u$

$h_1(z, w', u, v) = (z + w')^2 + u + v, h_2(w', v) = 0.5(w'^2 + v),$

Simulation III :  $m(\theta x^* + w, w', \eta) = 0.5(\theta x^* + w + w')^3 + \eta, g(z, w, w', u) = 0.5 \exp(z + w + w') \cdot u$

$h_1(z, w', u, v) = 0.5(z + w')^3 + u + v, h_2(w', v) = 0.1(w'^3 + v),$

Simulation IV :  $m(\theta x^* + w, w', \eta) = 0.1(\theta x^* + w + w')^3 \cdot \eta, g(z, w, w', u) = 0.5(z + w + w')^2 \cdot u$

$h_1(z, w', u, v) = (z + w')^2 + u + v, h_2(w', v) = 0.3(w'^2 + v)$

**Box II.**

$= Pr(y < n_1(z, w, w', \bar{\eta}) | Z, W, W')$

$= F_{Y|Z,W,W'}(n_1(z, w, w', \bar{\eta})).$

Differentiating the above expression with respect to  $w$  yields

$$0 = \frac{\partial F_{Y|Z,W,W'}(\bar{y})}{\partial w} + \frac{\partial F_{Y|Z,W,W'}(\bar{y})}{\partial y} \frac{\partial n_1(z, w, w', \bar{\eta})}{\partial w},$$

where  $\bar{y} = n_1(z, w, w', \bar{\eta})$ . It follows that

$$\frac{\partial n_1(z, w, w', \bar{\eta})}{\partial w} = - \left[ \frac{\partial F_{Y|Z,W,W'}(\bar{y})}{\partial y} \right]^{-1} \frac{\partial F_{Y|Z,W,W'}(\bar{y})}{\partial w}. \quad (14)$$

Similarly, we have

$$\frac{\partial n_1(z, w, w', \bar{\eta})}{\partial z} = - \left[ \frac{\partial F_{Y|Z,W,W'}(\bar{y})}{\partial y} \right]^{-1} \frac{\partial F_{Y|Z,W,W'}(\bar{y})}{\partial z}. \quad (15)$$

Hence we can express the derivative terms of  $n_1(z, w, w', \eta)$  in Eq. (11) in terms of the first derivatives of the cumulative distribution function  $F_{Y|Z,W,W'}$ . Next, we apply the above derivation to the triangular system (B). Denote

$$n_2(z, w, w', \varepsilon) = g(z, w, w', \tilde{h}_1(z, w', w)) + \varepsilon.$$

The derivatives of  $n_2(z, w, w', \varepsilon)$  with respect to  $z, w$  are:

$$\frac{\partial n_2(z, w, w', \varepsilon)}{\partial w} = g_w + g_u \frac{\partial \tilde{h}_1}{\partial w}, \quad (16)$$

$$\frac{\partial n_2(z, w, w', \varepsilon)}{\partial z} = g_z + g_u \frac{\partial \tilde{h}_1}{\partial z}. \quad (17)$$

Applying the similar derivation of Eqs. (12) and (13) in the system (A) to the system (B), we have

$$\frac{\partial n_2(z, w, w', \bar{\varepsilon})}{\partial w} = - \left[ \frac{\partial F_{X|Z,W,W'}(\bar{x})}{\partial x} \right]^{-1} \frac{\partial F_{X|Z,W,W'}(\bar{x})}{\partial w}, \quad (18)$$

$$\frac{\partial n_2(z, w, w', \bar{\varepsilon})}{\partial z} = - \left[ \frac{\partial F_{X|Z,W,W'}(\bar{x})}{\partial x} \right]^{-1} \frac{\partial F_{X|Z,W,W'}(\bar{x})}{\partial z}, \quad (19)$$

where  $\bar{x} = n_2(z, w, w', \bar{\varepsilon})$ . Combining those results and plugging them into Eq. (11) results in

$$\frac{- \left[ \frac{\partial F_{Y|Z,W,W'}(\bar{y})}{\partial y} \right]^{-1} \frac{\partial F_{Y|Z,W,W'}(\bar{y})}{\partial w}}{- \left[ \frac{\partial F_{Y|Z,W,W'}(\bar{y})}{\partial y} \right]^{-1} \frac{\partial F_{Y|Z,W,W'}(\bar{y})}{\partial z}} = \frac{\theta \left( - \left[ \frac{\partial F_{X|Z,W,W'}(\bar{x})}{\partial x} \right]^{-1} \frac{\partial F_{X|Z,W,W'}(\bar{x})}{\partial w} \right) + 1}{\theta \left( - \left[ \frac{\partial F_{X|Z,W,W'}(\bar{x})}{\partial x} \right]^{-1} \frac{\partial F_{X|Z,W,W'}(\bar{x})}{\partial z} \right)}.$$

This gives our desired identification result,

$$\theta = \frac{\frac{\partial F_{Y|Z,W,W'}(\bar{y})}{\partial z} \frac{\partial F_{X|Z,W,W'}(\bar{x})}{\partial x}}{\frac{\partial F_{Y|Z,W,W'}(\bar{y})}{\partial z} \frac{\partial F_{X|Z,W,W'}(\bar{x})}{\partial w} - \frac{\partial F_{Y|Z,W,W'}(\bar{y})}{\partial w} \frac{\partial F_{X|Z,W,W'}(\bar{x})}{\partial z}},$$

because Assumption 2.3 guarantees that the denominator in the last expression is nonzero.

**References**

Abrevaya, J., Hausman, J.A., Khan, S., 2010. Testing for causal effects in a generalized regression model with endogenous regressors. *Econometrica* 78, 2043–2061.  
 Borjas, G.J., 2009. *Labor Economics*, fifth ed. McGraw-Hill/Irwin.  
 Chesher, A., 2003. Identification in nonseparable models. *Econometrica* 71 (5), 1405–1441.  
 Hu, Y., 2008. Identification and estimation of nonlinear models with misclassification error using instrumental variables: A general solution. *J. Econometrics* 144, 27–61.  
 Hu, Y., Schennach, S.M., 2008. Instrumental variable treatment of nonclassical measurement error models. *Econometrica* 76, 195–216.  
 Hu, Y., Shiu, J., Woutersen, T., 2015. Identification and estimation of single index models with measurement error and endogeneity. *Econom. J.* 118 (3), 347–362.  
 Newey, W., 2001. Flexible simulated moment estimation of nonlinear errors-in-variables models. *Rev. Econ. Stat.* 83 (4), 616–627.