Nonparametric Estimation of Wages and Labor Force Participation

John Pepper
University of Virginia

Steven Stern
University of Virginia

May 15, 2000
Preliminary Draft - Comments Welcome

Abstract

1. Model

Let y_i be a binary indicator for labor force participation: $y_i = 1$ iff i works. Let w_i be the wage i would get if she worked, and let r_i be her reservation wage. It is assumed that $y_i = 1$ iff $w_i > r_i$. Let

$$w_i = g_w\left(X_i\beta_w, d_i, E_i\right) + u_i^w$$

where d_i is a binary indicator of i's disability ($d_i = 1$ iff i is disabled), E_i is a continuous measure of i's education, X_i is a vector of other observed variables affecting w_i, $g_w(\bullet)$ is an unspecified function with $g_{w1}(\bullet) \geq 0$, $g_{w2}(\bullet) \leq 0$, and $g_{w3}(\bullet) \geq 0$, and u_i^w is an error with finite mean and variance. Define $\overline{w}_i = g_w\left(X_i\beta_w, d_i, E_i\right)$.

Similarly, let

$$r_i = g_r\left(X_i\beta_r, d_i, E_i\right) + u_i^r$$

where $g_r(\bullet)$ is an unspecified function with $g_{r1}(\bullet) \geq 0$ and $g_{r2}(\bullet) \geq 0$ and u_i^r is an error with finite mean and variance. Define $\overline{r}_i = g_r\left(X_i\beta_r, d_i, E_i\right)$.

The joint density of $\left(u_i^w, u_i^r\right)$, $f\left(u_i^w, u_i^r\right)$, is unspecified beyond the moment restrictions already stated.

2. Identification

The goal is to estimate $\theta = [\beta_w, \beta_r, g_w(\bullet), g_r(\bullet), f(\bullet)]$ or that part of θ that is identified. The data consists of $\{y_i, w_i y_i, X_i, d_i, E_i\}_{i=1}^n$. Assume temporarily that there are no restrictions on $g_w(X_i\beta_w, d_i, E_i)$ or $g_r(X_i\beta_r, d_i, E_i)$. Then, for observations where $w_i \leq r_i$, we observe only $y_i = 0$, while, for observations where where $w_i > r_i$, we observe $y_i = 1$ and w_i. Define

$$F_1(u_i^w, u_i^r) = \int_{-\infty}^{u_i^r} f(u_i^w, u^r)\, du^r,$$

and

$$\begin{aligned} F^*(u^*) &= \Pr[u_i^w - u_i^r \leq u^*] \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{u^*+u^r} f(u^w, u^r)\, du^w du^r. \end{aligned}$$

Note that $\frac{\partial}{\partial u_i^r}F_1(u_i^w, u_i^r) > 0$ and $\frac{\partial}{\partial u^*}F^*(u^*) > 0$. Then the likelihood contribution for i is

$$\left\{ \begin{array}{ll} F_1(w_i - g_w(X_i\beta_w, d_i, E_i), w_i - g_r(X_i\beta_r, d_i, E_i)) & \text{if } y_i = 1 \\ F^*(g_r(X_i\beta_r, d_i, E_i) - g_w(X_i\beta_w, d_i, E_i)) & \text{if } y_i = 0 \end{array} \right. . \tag{2.1}$$

Now define

$$\Im_1[w, X\beta_w, X\beta_r, d, E] = F_1(w - g_w(X\beta_w, d, E), w - g_r(X\beta_r, d, E)) \tag{2.2}$$

and

$$\Im^*[X\beta_w, X\beta_r, d, E] = F^*(g_r(X\beta_r, d, E) - g_w(X\beta_w, d, E)) \tag{2.3}$$

nonparametrically. Define

$$z_j = \begin{pmatrix} X_j\beta_w \\ X_j\beta_r \\ d_j \\ E_j \end{pmatrix}. \tag{2.4}$$

Then $\Im_1$ and $\Im^*$ can be estimated as

$$\widehat{\Im}_1[w, z] = \frac{\frac{1}{\sigma}\sum_j y_j H\left[\frac{w_j - w}{\sigma}\right] K\left[(z_j - z)'\,\Omega^{-1}(z_j - z)\right]}{\sum_j K\left[(z_j - z)'\,\Omega^{-1}(z_j - z)\right]} \tag{2.5}$$

and

$$\widehat{\Im}^*[z] = \frac{\sum_j (1-y_j) K\left[(z_j - z)' \Omega^{-1} (z_j - z)\right]}{\sum_j K\left[(z_j - z)' \Omega^{-1} (z_j - z)\right]} \tag{2.6}$$

where $K[\bullet]$ and $H[\bullet]$ are kernel functions, Ω is a bandwidth matrix, and σ is a bandwidth.

Then, note that

$$\begin{aligned} \Im_{11} &= F_{11} + F_{12}, \\ \Im_{12} &= -F_{11} g_{w1}, \\ \Im_{13} &= -F_{12} g_{r1}, \\ \Im_1^* &= -F^{*\prime} g_{w1}, \text{ and} \\ \Im_2^* &= F^{*\prime} g_{r1}. \end{aligned} \tag{2.7}$$

Thus, $(F_{11}, F_{12}, F^{*\prime}, g_{w1}, g_{r1})$ is identified by $(\Im_{11}, \Im_{12}, \Im_{13}, \Im_1^*, \Im_2^*)$. Next, note that

$$\begin{aligned} \Im_{14} &= F_{12} g_{r2} - F_{11} g_{w2} \text{ and} \\ \Im_{15} &= F_{12} g_{r3} - F_{11} g_{w3}. \end{aligned} \tag{2.8}$$

Assuming that F_{11} and F_{12} (already identified) are not colinear as w changes, (g_{r2}, g_{w2}) is identified by $\Im_{14}$ and variation in (F_{11}, F_{12}) with w, and (g_{r3}, g_{w3}) is identified by $\Im_{15}$ and variation in (F_{11}, F_{12}) with w.

Note that there are also many overidentifying restrictions. Most obviously, because

$$\begin{aligned} \Im_3^* &= F^{*\prime}[g_{r3} - g_{w3}], \text{ and} \\ \Im_4^* &= F^{*\prime}[g_{r4} - g_{w4}], \end{aligned}$$

there are restrictions on $(\Im_3^*, \Im_4^*)$. Also, any two values of w identify (g_{r3}, g_{w3}) and (g_{r4}, g_{w4}); all other values of w provide overidentifying restrictions. Finally, I think that there might be some restrictions placed on variation in structural functionals as $(w, X\beta_r, X\beta_w, d, E)$ varies, but I haven't convinced myself I am right, much less what they would be.

One needs to anchor the relevant functionals. For example, define $g_r(0,0,12) = a_r$ and $g_w(0,0,12) = a_w$.[1] Once g_r and g_w are anchored, $F_1(w - a_w, w - a_r)$ and $F^*(a_r - a_w)$ are identified.

---

[1] This is equivalent to assuming the mean of the errors or not including a constant in g_r and g_w.

Note that one can estimate β_r and β_w without making any functional form assumptions. A nonparametric estimator of $\beta = (\beta_w, \beta_r)$ is

$$\widehat{\beta} = \arg\max_{\beta} \sum_i y_i \log \widehat{\Im}_1 \left[w_i, z_i(\beta)\right] + (1-y_i) \log \widehat{\Im}^* \left[z_i(\beta)\right] \tag{2.9}$$

where $\widehat{\Im}_1[w_i, z_i]$ and $\widehat{\Im}^*[z_i]$ are defined in equations (2.5) and (2.6) (with sums over $j \neq i$) and $z_i(\beta)$ is defined in equation (2.4). What are the statistical properties of $\widehat{\beta}$? Once β is estimated, we can infer the unspecified functionals (F_1, F^*, g_w, g_r) (see appendix for suggestion).

3. Estimation Strategy

3.1. Adding Restrictions

We know that equation (2.2) holds. The restrictions we want to add are $\frac{\partial F_1}{\partial u^r} > 0$, $\frac{\partial g_r}{\partial X\beta_r} \geq 0$, $\frac{\partial g_r}{\partial d} \geq 0$, $\frac{\partial g_w}{\partial X\beta_w} \geq 0$, $\frac{\partial g_w}{\partial d} \leq 0$, and $\frac{\partial g_w}{\partial E} \geq 0$. Equation (2.2) implies that

$$\begin{aligned} \frac{\partial \Im_1}{\partial X\beta_r} &= \frac{\partial F_1}{\partial X\beta_r} = -\frac{\partial F_1}{\partial u^r}\frac{\partial g_r}{\partial X\beta_r} \leq 0; \\ \frac{\partial \Im_1}{\partial d} &= \frac{\partial F_1}{\partial d} = -\frac{\partial F_1}{\partial u^w}\frac{\partial g_w}{\partial d} - \frac{\partial F_1}{\partial u^r}\frac{\partial g_r}{\partial d}; \text{ and} \\ \frac{\partial \Im_1}{\partial w} &= \frac{\partial F_1}{\partial w} = \frac{\partial F_1}{\partial u^w} + \frac{\partial F_1}{\partial u^r}. \end{aligned} \tag{3.1}$$

$\frac{\partial \Im_1}{\partial d}$ and $\frac{\partial \Im_1}{\partial w}$ impose no restrictions because we can not sign $\frac{\partial F_1}{\partial u^w}$. But, in regions where the estimate of $\frac{\partial F_1}{\partial u^w}$ is nonpositive, $\frac{\partial F_1}{\partial d} \leq 0$, and, in regions where the estimate of $\frac{\partial F_1}{\partial u^w}$ is nonnegative, $\frac{\partial F_1}{\partial w} \geq 0$. Thus imposing restrictions on the estimates of structural functions, F_1 and g_r, implies restrictions on the nonparametric function $\Im_1$.

Similarly, equation (2.3) implies that

$$\begin{aligned} \frac{\partial \Im^*}{\partial X_i\beta_r} &= \frac{\partial F^*}{\partial u^*}\frac{\partial g_r}{\partial X_i\beta_r} \geq 0; \\ \frac{\partial \Im^*}{\partial X_i\beta_w} &= -\frac{\partial F^*}{\partial u^*}\frac{\partial g_w}{\partial X_i\beta_w} \leq 0; \text{ and} \\ \frac{\partial \Im^*}{\partial d_i} &= \frac{\partial F^*}{\partial u^*}\left[\frac{\partial g_r}{\partial d_i} - \frac{\partial g_w}{\partial d_i}\right] \geq 0. \end{aligned} \tag{3.2}$$

Thus imposing restrictions on the estimates of structural functions, F^*, g_w, and g_r, implies restrictions on the nonparametric function .

The next issue is whether restrictions on $\Im_1$ and $\Im^*$ implied in equations (3.1) and (3.2) imply restrictions on F_1, F^*, g_w, and g_r. Note that the third restriction in equation (3.1) restricts $\frac{\partial F_1}{\partial u^r}$. Given that restriction, the first two restrictions in equation (3.1) restrict $\frac{\partial g_r}{\partial X\beta_r}$ and $\frac{\partial g_r}{\partial d}$. These, together with the first restriction in equation (3.2), restrict $\frac{\partial F^*}{\partial u^*}$. Then, the second restriction in equation (3.2) restricts $\frac{\partial g_w}{\partial X\beta_w}$. Finally, the last restriction in equation (3.2) restricts $\frac{\partial g_r}{\partial d} - \frac{\partial g_w}{\partial d} \geq 0$. But this implies only that $\frac{\partial g_w}{\partial d} \leq \frac{\partial g_r}{\partial d}$ which is a much weaker restriction than $\frac{\partial g_w}{\partial d} \leq 0$. In fact, there is no obvious way to impose restrictions on the nonparametric functions, $\Im_1$ and $\Im^*$, such that they will imply that $\frac{\partial g_w}{\partial d} \leq 0$. Note, however, that, even though we can't restrict $\frac{\partial g_w}{\partial d} \leq 0$, we can still identify $\frac{\partial g_w}{\partial d}$.

3.2. General Estimation Problem: Nonparametric Estimation of a Conditional Density

The general estimation problem consists of estimating two conditional density functions. Consider the generic conditional density function $f(\varepsilon \mid \xi)$, and consider estimating $f(\varepsilon \mid \xi)$ using local regression methods. In our problem, when $y = 0$,

$$f(\varepsilon \mid \xi) = \Im^*[z]$$

with $\varepsilon = y$ and $\xi = z$, and, when $y = 1$,

$$f(\varepsilon \mid \xi) = \Im_1[w, z]$$

with $\varepsilon = y, w$ and $\xi = z$. Define

$$f(u \mid t) = \exp\{p(u, \varepsilon, t, \xi)\} \tag{3.3}$$

at (u, t) near (ε, ξ) with

$$p(u, \varepsilon, t, \xi) = a_{\varepsilon\xi} + b_{\varepsilon}^{'}(u-\varepsilon) + b_{\xi}^{'}(t-\xi) + (u-\varepsilon)' C_{\varepsilon}(u-\varepsilon) + (t-\xi)' C_{\xi}(t-\xi)$$

and

$$\frac{\partial}{\partial a_{\varepsilon\xi}} p(u, \varepsilon, t, \xi) = 1;$$

$$\begin{aligned}
\frac{\partial}{\partial b_\varepsilon} p(u,\varepsilon,t,\xi) &= (u-\varepsilon)\,; \\
\frac{\partial}{\partial b_\xi} p(u,\varepsilon,t,\xi) &= (t-\xi)\,; \\
\frac{\partial}{\partial C_\varepsilon} p(u,\varepsilon,t,\xi) &= (u-\varepsilon)(u-\varepsilon)'\,; \\
\frac{\partial}{\partial C_\xi} p(u,\varepsilon,t,\xi) &= (t-\xi)(t-\xi)'\,.
\end{aligned}$$

Note that

$$\begin{aligned}
\log f(\varepsilon \mid \xi) &= a_{\varepsilon\xi}; \\
\frac{\partial}{\partial \varepsilon} \log f(\varepsilon \mid \xi) &= b_\varepsilon; \\
\frac{\partial}{\partial \xi} \log f(\varepsilon \mid \xi) &= b_\xi.
\end{aligned} \tag{3.4}$$

Define

$$\log \widehat{f}(\varepsilon \mid \xi) = \arg\min_{a_{\varepsilon\xi}} \left\{ \min_{b,C} L\left(\varepsilon, \xi, a_{\varepsilon\xi}, b, C\right) \right\} \tag{3.5}$$

where

$$\begin{aligned}
& L(\varepsilon, \xi, a_{\varepsilon\xi}, b, C) \\
= \; & \frac{\frac{1}{n} |\Omega_\xi|^{-\frac{1}{2}} |\Omega_\varepsilon|^{-\frac{1}{2}} \sum_i K_\xi(\xi_i - \xi; \Omega_\xi) K_\varepsilon(\varepsilon_i - \varepsilon; \Omega_\varepsilon) p(\varepsilon_i, \varepsilon, \xi_i, \xi)}{\frac{1}{n} |\Omega_\xi|^{-\frac{1}{2}} \sum_i K_\xi(\xi_i - \xi; \Omega_\xi)} \\
& - \frac{|\Omega_\xi|^{-\frac{1}{2}} |\Omega_\varepsilon|^{-\frac{1}{2}} \int\int K_\xi(t - \xi; \Omega_\xi) K_\varepsilon(u - \varepsilon; \Omega_\varepsilon) \exp\{p(u,\varepsilon,t,\xi)\}\, dudt}{|\Omega_\xi|^{-\frac{1}{2}} \int K_\xi(t-\xi; \Omega_\xi)\, dt}.
\end{aligned} \tag{3.6}$$

Define

$$\psi_\varsigma(\varsigma_i, \varsigma) = |\Omega_\varsigma|^{-\frac{1}{2}} K_\varsigma(\varsigma_i - \varsigma; \Omega_\varsigma) \exp\left\{ b_\varsigma^{'}(\varsigma_i - \varsigma) + (\varsigma_i - \varsigma)' C_\varsigma (\varsigma_i - \varsigma) \right\}$$

for $\varsigma = \varepsilon, \xi$ and

$$\Psi_\varsigma(\varsigma) = \int \psi_\varsigma(v, \varsigma)\, dv$$

with

$$\begin{aligned}
\frac{\partial}{\partial b_\varsigma} \Psi_\varsigma(\varsigma) &= \int (v - \varsigma)\, \psi_\varsigma(v, \varsigma)\, dv; \\
\frac{\partial}{\partial C_\varsigma} \Psi_\varsigma(\varsigma) &= \int (v - \varsigma)(v - \varsigma)^{'}\, \psi_\varsigma(v, \varsigma)\, dv.
\end{aligned}$$

Then equation (3.6) can be written as

$$\begin{aligned} L\left(\varepsilon,\xi,a,b,C\right) \;=\;& \frac{\frac{1}{n}\left|\Omega_\xi\right|^{-\frac{1}{2}}\left|\Omega_\varepsilon\right|^{-\frac{1}{2}}\sum_i K_\xi\left(\xi_i-\xi;\Omega_\xi\right)K_\varepsilon\left(\varepsilon_i-\varepsilon;\Omega_\varepsilon\right)p(\varepsilon_i,\varepsilon,\xi_i,\xi)}{\frac{1}{n}\left|\Omega_\xi\right|^{-\frac{1}{2}}\sum_i K_\xi\left(\xi_i-\xi;\Omega_\xi\right)} \\ & -e^{a_{\varepsilon\xi}}\Psi_\xi\left(\xi\right)\Psi_\varepsilon\left(\varepsilon\right). \end{aligned} \tag{3.7}$$

The partial derivatives are

$$\frac{\partial}{\partial a_{\varepsilon\xi}}L=\frac{\frac{1}{n}\left|\Omega_\xi\right|^{-\frac{1}{2}}\left|\Omega_\varepsilon\right|^{-\frac{1}{2}}\sum_i K_\xi\left(\xi_i-\xi;\Omega_\xi\right)K_\varepsilon\left(\varepsilon_i-\varepsilon;\Omega_\varepsilon\right)}{\frac{1}{n}\left|\Omega_\xi\right|^{-\frac{1}{2}}\sum_i K_\xi\left(\xi_i-\xi;\Omega_\xi\right)}-e^{a_{\varepsilon\xi}}\Psi_\xi\left(\xi\right)\Psi_\varepsilon\left(\varepsilon\right); \tag{3.8}$$

$$\begin{aligned} \frac{\partial}{\partial b_\varepsilon}L \;=\;& \frac{\frac{1}{n}\left|\Omega_\xi\right|^{-\frac{1}{2}}\left|\Omega_\varepsilon\right|^{-\frac{1}{2}}\sum_i K_\xi\left(\xi_i-\xi;\Omega_\xi\right)K_\varepsilon\left(\varepsilon_i-\varepsilon;\Omega_\varepsilon\right)\left(\varepsilon_i-\varepsilon\right)}{\frac{1}{n}\left|\Omega_\xi\right|^{-\frac{1}{2}}\sum_i K_\xi\left(\xi_i-\xi;\Omega_\xi\right)} \\ & -e^{a_{\varepsilon\xi}}\Psi_\xi\left(\xi\right)\int\left(u-\varepsilon\right)\psi_\varepsilon\left(u,\varepsilon\right)du; \end{aligned} \tag{3.9}$$

$$\begin{aligned} \frac{\partial}{\partial b_\xi}L \;=\;& \frac{\left|\Omega_\xi\right|^{-\frac{1}{2}}\sum_i \Psi_\varepsilon\left(\varepsilon_i,\varepsilon\right)K_\xi\left(\xi_i-\xi;\Omega_\varepsilon\right)\left(\xi_i-\xi\right)}{\sum_i \Psi_\xi\left(\xi_i,\xi\right)} \\ & -e^{a_{\varepsilon\xi}}\Psi_\varepsilon\left(\varepsilon\right)\int\left(t-\xi\right)\psi_\xi\left(t,\xi\right)dt; \end{aligned} \tag{3.10}$$

$$\begin{aligned} \frac{\partial}{\partial C_\varepsilon}L \;=\;& \frac{\frac{1}{n}\left|\Omega_\varepsilon\right|^{-\frac{1}{2}}\sum_i \Psi_\xi\left(\xi_i,\xi\right)K_\varepsilon\left(\varepsilon_i-\varepsilon;\Omega_\varepsilon\right)\left(\varepsilon_i-\varepsilon\right)(\varepsilon_i-\varepsilon)'}{\frac{1}{n}\sum_i \Psi_\xi\left(\xi_i,\xi\right)} \\ & -e^{a_{\varepsilon\xi}}\Psi_\xi\left(\xi\right)\int\left(u-\varepsilon\right)\left(u-\varepsilon\right)'\psi_\varepsilon\left(u,\varepsilon\right)du; \end{aligned} \tag{3.11}$$

and

$$\begin{aligned} \frac{\partial}{\partial C_\xi}L \;=\;& \frac{\left|\Omega_\xi\right|^{-\frac{1}{2}}\sum_i \Psi_\varepsilon\left(\varepsilon_i,\varepsilon\right)K_\xi\left(\xi_i-\xi;\Omega_\xi\right)\left(\xi_i-\xi\right)(\xi_i-\xi)'}{\sum_i \Psi_\xi\left(\xi_i,\xi\right)} \\ & -e^{a_{\varepsilon\xi}}\Psi_\varepsilon\left(\varepsilon\right)\int\left(t-\xi\right)\left(t-\xi\right)'\psi_\xi\left(t,\xi\right)dt. \end{aligned} \tag{3.12}$$

Also, note that, from equation (3.4), we can get a consistent estimate of $\frac{\partial\log f(\varepsilon|\xi)}{\partial\varepsilon}$ as $\widehat{b}_\varepsilon$ and of $\frac{\partial\log f(\varepsilon|\xi)}{\partial\xi}$ as $\widehat{b}_\xi$.

Now, let

$$K_\varsigma(\widetilde{v}) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{\widetilde{v}'\widetilde{v}}{2}\right\}.$$

with

$$\widetilde{v} = R_\varsigma^{-1}(v - \varsigma)$$

and

$$R_\varsigma R_\varsigma^{'} = \Omega_\varsigma.$$

Then,

$$\begin{aligned}
\psi_\varsigma(v,\varsigma) &= \frac{1}{\sqrt{2\pi}} |\Omega_\varsigma|^{-\frac{1}{2}} \exp\left\{-\frac{\widetilde{v}'\widetilde{v}}{2}\right\} \exp\left\{b_\varsigma^{'} R_\varsigma \widetilde{v} + \widetilde{v}' R_\varsigma^{'} C_\varsigma R_\varsigma \widetilde{v}\right\} \qquad (3.13)\\
&= \frac{1}{\sqrt{2\pi}} |\Omega_\varsigma|^{-\frac{1}{2}} \exp\left\{-\frac{\widetilde{v}'\widetilde{v}}{2} + b_\varsigma^{'} R_\varsigma \widetilde{v} + \widetilde{v}' R_\varsigma^{'} C_\varsigma R_\varsigma \widetilde{v}\right\}\\
&= \frac{1}{\sqrt{2\pi}} |\Omega_\varsigma|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}\widetilde{v}' \left[I - 2R_\varsigma^{'} C_\varsigma R_\varsigma\right] \widetilde{v} + b_\varsigma^{'} R_\varsigma \widetilde{v}\right\}.
\end{aligned}$$

We assume that $I - 2R_\varsigma^{'} C_\varsigma R_\varsigma$ is positive definite. Note that, asymptotically, $R_\varsigma \to 0$, and it is likely that C_ς is small. Then equation (3.13) can be written as

$$\psi_\varsigma(\varsigma_i,\varsigma) = \frac{1}{\sqrt{2\pi}} |\Omega_\varsigma|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\widetilde{v} - \alpha_\varsigma)' \left[I - 2R_\varsigma^{'} C_\varsigma R_\varsigma\right] (\widetilde{v} - \alpha_\varsigma) + \gamma_\varsigma\right\} \qquad (3.14)$$

where

$$\begin{aligned}
\alpha_\varsigma &= \left[I - 2R_\varsigma^{'} C_\varsigma R_\varsigma\right]^{-1} R_\varsigma^{'} b_\varsigma;\\
\gamma_\varsigma &= \frac{1}{2} b_\varsigma^{'} R_\varsigma \left[I - 2R_\varsigma^{'} C_\varsigma R_\varsigma\right]^{-1} R_\varsigma^{'} b_\varsigma.
\end{aligned}$$

Then equation (3.14) can be written as

$$\psi_\varsigma(v,\varsigma) = |\Omega_\varsigma|^{-\frac{1}{2}} \exp\{\gamma_\varsigma\}\, \phi\left(\left[I - 2R_\varsigma^{'} C_\varsigma R_\varsigma\right]^{\frac{1}{2}} (\widetilde{v} - \alpha_\varsigma)\right)$$

with $\phi(\cdot)$ being the standard multivariate normal density function. Note that

$$\Psi_\varsigma(\varsigma) = \int \psi_\varsigma(v,\varsigma)\, dv = \exp\{\gamma_\varsigma\};$$

$$\int \psi_\varsigma(v,\varsigma)(v - \varsigma)\, dv = \exp\{\gamma_\varsigma\} \left[I - 2R_\varsigma^{'} C_\varsigma R_\varsigma\right]^{-\frac{1}{2}} \alpha_\varsigma;$$

and

$$\int \psi_{\varsigma}\left(v,\varsigma\right)\left(v-\varsigma\right)\left(v-\varsigma\right)' dv = \exp\left\{\gamma_{\varsigma}\right\}\left[I-2R_{\varsigma}^{'}C_{\varsigma}R_{\varsigma}\right]^{-1}.$$

Define

$$\Upsilon = e^{a_{\varepsilon\xi}}\exp\left\{\gamma_{\xi}+\gamma_{\varepsilon}\right\}.$$

Then equations (**??**) through (3.12) become

$$\frac{\partial}{\partial a_{\varepsilon\xi}}L = \frac{\frac{1}{n}\left|\Omega_{\xi}\right|^{-\frac{1}{2}}\left|\Omega_{\varepsilon}\right|^{-\frac{1}{2}}\sum_{i}K_{\xi}\left(\xi_{i}-\xi;\Omega_{\xi}\right)K_{\varepsilon}\left(\varepsilon_{i}-\varepsilon;\Omega_{\varepsilon}\right)}{\frac{1}{n}\left|\Omega_{\xi}\right|^{-\frac{1}{2}}\sum_{i}K_{\xi}\left(\xi_{i}-\xi;\Omega_{\xi}\right)} - \Upsilon;$$

$$\begin{aligned}\frac{\partial}{\partial b_{\varepsilon}}L &= \frac{\frac{1}{n}\left|\Omega_{\xi}\right|^{-\frac{1}{2}}\left|\Omega_{\varepsilon}\right|^{-\frac{1}{2}}\sum_{i}K_{\xi}\left(\xi_{i}-\xi;\Omega_{\xi}\right)K_{\varepsilon}\left(\varepsilon_{i}-\varepsilon;\Omega_{\varepsilon}\right)\left(\varepsilon_{i}-\varepsilon\right)}{\frac{1}{n}\left|\Omega_{\xi}\right|^{-\frac{1}{2}}\sum_{i}K_{\xi}\left(\xi_{i}-\xi;\Omega_{\xi}\right)} \\ &\quad -\Upsilon\left[I-2R_{\varepsilon}^{'}C_{\varepsilon}R_{\varepsilon}\right]^{-\frac{1}{2}}\alpha_{\varepsilon};\end{aligned}$$

$$\begin{aligned}\frac{\partial}{\partial b_{\xi}}L &= \frac{\frac{1}{n}\left|\Omega_{\xi}\right|^{-\frac{1}{2}}\left|\Omega_{\varepsilon}\right|^{-\frac{1}{2}}\sum_{i}K_{\xi}\left(\xi_{i}-\xi;\Omega_{\xi}\right)K_{\varepsilon}\left(\varepsilon_{i}-\varepsilon;\Omega_{\varepsilon}\right)\left(\xi_{i}-\xi\right)}{\frac{1}{n}\left|\Omega_{\xi}\right|^{-\frac{1}{2}}\sum_{i}K_{\xi}\left(\xi_{i}-\xi;\Omega_{\xi}\right)} \\ &\quad -\Upsilon\left[I-2R_{\xi}^{'}C_{\xi}R_{\xi}\right]^{-\frac{1}{2}}\alpha_{\xi};\end{aligned}$$

$$\begin{aligned}\frac{\partial}{\partial C_{\varepsilon}}L &= \frac{\frac{1}{n}\left|\Omega_{\varepsilon}\right|^{-\frac{1}{2}}\sum_{i}\Psi_{\xi}\left(\xi_{i},\xi\right)K_{\varepsilon}\left(\varepsilon_{i}-\varepsilon;\Omega_{\varepsilon}\right)\left(\varepsilon_{i}-\varepsilon\right)(\varepsilon_{i}-\varepsilon)'}{\frac{1}{n}\sum_{i}\Psi_{\xi}\left(\xi_{i},\xi\right)} \\ &\quad -\Upsilon\left[I-2R_{\xi}^{'}C_{\xi}R_{\xi}\right]^{-1};\end{aligned}$$

and

$$\begin{aligned}\frac{\partial}{\partial C_{\xi}}L &= \frac{\left|\Omega_{\xi}\right|^{-\frac{1}{2}}\sum_{i}\Psi_{\varepsilon}\left(\varepsilon_{i},\varepsilon\right)K_{\xi}\left(\xi_{i}-\xi;\Omega_{\xi}\right)\left(\xi_{i}-\xi\right)(\xi_{i}-\xi)'}{\sum_{i}\Psi_{\xi}\left(\xi_{i},\xi\right)} \\ &\quad -\Upsilon\left[I-2R_{\varepsilon}^{'}C_{\varepsilon}R_{\varepsilon}\right]^{-1}.\end{aligned}$$

3.3. Estimation Procedure Without Restrictions

Estimation of $\theta = [\beta_w, \beta_r, g_w(\bullet), g_r(\bullet), f(\bullet)]$ follows from above. We consider the minimization problem analogous to equation (2.9):

$$\widehat{\beta} = \arg\max_{\beta} \sum_i y_i \log \widehat{\Im}_1 [w_i, z_i(\beta)] + (1 - y_i) \log \widehat{\Im}^* [z_i(\beta)] \tag{3.15}$$

where

$$\log \widehat{\Im}_1 [w_i, z_i(\beta)] = \arg\min_{a_\varepsilon} \left\{ \min_{a_\xi, b, C} L(\varepsilon, \xi, a, b, C) \right\} \tag{3.16}$$

from equation (3.5) with $\varepsilon = (y, w)$ and $\xi = z(\beta)$; and

$$\log \widehat{\Im}^* [z_i(\beta)] = \arg\min_{a_\varepsilon} \left\{ \min_{a_\xi, b, C} L(\varepsilon, \xi, a, b, C) \right\} \tag{3.17}$$

with $\varepsilon = 1 - y$ and $\xi = z(\beta)$. We can get consistent estimates of the partial derivatives of $\widehat{\Im}_1 [w_i, z_i(\beta)]$ and $\widehat{\Im}^* [z_i(\beta)]$ using our estimates of b from equations (3.16) and (3.17). We can stack first order conditions in equations (3.1) and (3.2) to estimate the structural functionals. Details are described in the appendix.

3.4. Estimation Procedure With Restrictions

Let θ^k be a guess of θ. Using equations (3.1) and (3.2), we can evaluate $\widehat{\Im}_1 [w_i, z_i]$ and $\widehat{\Im}^* [z_i]$ and their partial derivatives at every point (w_i, z_i). This implies values of $\left(a^k_\varepsilon, a^k_\xi, b^k_\varepsilon, b^k_\xi\right)$ for both $\widehat{\Im}_1 [w_i, z_i]$ and $\widehat{\Im}^* [z_i]$ in equations (3.16) and (3.17). In particular, for example for $\widehat{\Im}_1 [w_i, z_i]$,

$$\begin{aligned} b^k_\varepsilon(\varepsilon) &= \partial \log \widehat{\Im}_1 [w_i, z_i] / \partial w_i; \\ b^k_\xi(\xi) &= \partial \log \widehat{\Im}_1 [w_i, z_i] / \partial z_i; \\ a^k_{\varepsilon\xi}(\varepsilon, \xi) &= \int^\varepsilon b^k_\varepsilon(u)\, du + \int^\xi b^k_\xi(t)\, dt. \end{aligned}$$

A similar set of equations apply for $\widehat{\Im}^* [z_i]$. The values of $\left(a^k_{\varepsilon\xi}, b^k_\varepsilon, b^k_\xi\right)$ for $\widehat{\Im}_1 [w_i, z_i]$ and $\widehat{\Im}^* [z_i]$ can be plugged into equation (3.15) and then maximized over β. This is equivalent to solving a constrained maximization problem that imposes the structure implied by θ on the nonparametric log likelihood function. It is also straightforward now to impose monotonicity constraints on the structural equations in that they correspond to imposing nonnegativity (nonpositivity) constraints on some elements of θ.

4. References

References

[1] Mukarjee, Hari and Steven Stern (1994). “Feasible Nonparametric Estimation of Multiargument Monotone Functions.” *Journal of the American Statistical Association.* 89, 77-80.

[2] Robertson, Tim, F. T. Wright, and R. L. Dykstra (1988). *Order Restricted Statistical Inference.* New York: John Wiley and Sons.

[3] Stern, Steven (1996). “Semiparametric Estimates of the Supply and Demand Effects of Disability on Labor Force Participation.” *Journal of Econometrics.* 49-70.

5. Appendix: Proposal to Solve for Structural Functionals Nonparametrically

The set of restrictions in equations (2.7) and (2.8) can be approximated by a Taylor series expansion as

$$\begin{aligned}
\Im_{11} &= F_{11} + F_{12}, \\
\Im_{12} &= -\underline{F_{11}}\,\underline{g_{w1}} - \underline{F_{11}}\left(g_{w1} - \underline{g_{w1}}\right) - \underline{g_{w1}}\left(F_{11} - \underline{F_{11}}\right), \\
\Im_{13} &= -\underline{F_{12}}\,\underline{g_{r1}} - \underline{F_{12}}\left(g_{r1} - \underline{g_{r1}}\right) - \underline{g_{r1}}\left(F_{12} - \underline{F_{12}}\right), \\
\Im_{14} &= \underline{F_{12}}\,\underline{g_{r2}} + \underline{F_{12}}\left(g_{r2} - \underline{g_{r2}}\right) + \underline{g_{r2}}\left(F_{12} - \underline{F_{12}}\right) \\
&\quad -\underline{F_{11}}\,\underline{g_{w2}} - \underline{F_{11}}\left(g_{w2} - \underline{g_{w2}}\right) - \underline{g_{w2}}\left(F_{11} - \underline{F_{11}}\right), \\
\Im_{15} &= \underline{F_{12}}\,\underline{g_{r3}} + \underline{F_{12}}\left(g_{r3} - \underline{g_{r3}}\right) + \underline{g_{r3}}\left(F_{12} - \underline{F_{12}}\right) \\
&\quad -\underline{F_{11}}\,\underline{g_{w3}} - \underline{F_{11}}\left(g_{w3} - \underline{g_{w3}}\right) - \underline{g_{w3}}\left(F_{11} - \underline{F_{11}}\right),
\end{aligned} \tag{5.1}$$

$$\begin{aligned}
\Im_1^* &= \underline{F^{*\prime}}\,\underline{g_{r1}} + \underline{F^{*\prime}}\left(g_{r1} - \underline{g_{r1}}\right) + \underline{g_{r1}}\left(F^{*\prime} - \underline{F^{*\prime}}\right), \\
\Im_2^* &= -\underline{F^{*\prime}}\,\underline{g_{w1}} - \underline{F^{*\prime}}\left(g_{w1} - \underline{g_{w1}}\right) - \underline{g_{w1}}\left(F^{*\prime} - \underline{F^{*\prime}}\right), \\
\Im_3^* &= \underline{F^{*\prime}}\,\underline{g_{r2}} + \underline{F^{*\prime}}\left(g_{r2} - \underline{g_{r2}}\right) + \underline{g_{r2}}\left(F^{*\prime} - \underline{F^{*\prime}}\right) \\
&\quad -\underline{F^{*\prime}}\,\underline{g_{w2}} - \underline{F^{*\prime}}\left(g_{w2} - \underline{g_{w2}}\right) - \underline{g_{w2}}\left(F^{*\prime} - \underline{F^{*\prime}}\right), \text{ and}
\end{aligned} \tag{5.2}$$

$$\begin{aligned}\Im_4^* &= \underline{F^{*\prime}}\underline{g_{r3}} + \underline{F^{*\prime}}\left(g_{r3} - \underline{g_{r3}}\right) + \underline{g_{r3}}\left(F^{*\prime} - \underline{F^{*\prime}}\right) \\ &\quad -\underline{F^{*\prime}}\underline{g_{w3}} - \underline{F^{*\prime}}\left(g_{w3} - \underline{g_{w3}}\right) - \underline{g_{w3}}\left(F^{*\prime} - \underline{F^{*\prime}}\right)\end{aligned}$$

where underlined variables are evaluated at a fixed point. These equations can be written in matrix form. Let

$$W(w,z) = \begin{pmatrix} \Im_{11} - \underline{F_{11}}\underline{F_{12}} \\ \Im_{12} + \underline{F_{11}}\underline{g_{w1}} \\ \Im_{13} + \underline{F_{12}}\underline{g_{r1}} \\ \Im_{14} - \underline{F_{12}}\underline{g_{r2}} + \underline{F_{11}}\underline{g_{w2}} \\ \Im_{15} - \underline{F_{12}}\underline{g_{r3}} + \underline{F_{11}}\underline{g_{w3}} \\ \Im_1^* - \underline{F^{*\prime}}\underline{g_{r1}} \\ \Im_2^* + \underline{F^{*\prime}}\underline{g_{w1}} \\ \Im_3^* - \underline{F^{*\prime}}\underline{g_{r2}} + \underline{F^{*\prime}}\underline{g_{w3}} \\ \Im_4^* - \underline{F^{*\prime}}\underline{g_{r3}} + \underline{F^{*\prime}}\underline{g_{w3}} \end{pmatrix},$$

$$A(w,z) = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -\underline{g_{w1}} & 0 & 0 & -\underline{F_{11}} & 0 & 0 & 0 & 0 & 0 \\ 0 & -\underline{g_{r1}} & 0 & 0 & 0 & 0 & -\underline{F_{12}} & 0 & 0 \\ -\underline{g_{w2}} & \underline{g_{r2}} & 0 & 0 & -\underline{F_{11}} & 0 & 0 & \underline{F_{12}} & 0 \\ -\underline{g_{w3}} & \underline{g_{r3}} & 0 & 0 & 0 & -\underline{F_{11}} & 0 & 0 & \underline{F_{12}} \\ 0 & 0 & \underline{g_{r1}} & 0 & 0 & 0 & \underline{F^{*\prime}} & 0 & 0 \\ 0 & 0 & -\underline{g_{w1}} & -\underline{F^{*\prime}} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \underline{g_{r2}} - \underline{g_{w2}} & 0 & -\underline{F^{*\prime}} & 0 & 0 & \underline{F^{*\prime}} & 0 \\ 0 & 0 & \underline{g_{r3}} - \underline{g_{w3}} & 0 & 0 & -\underline{F^{*\prime}} & 0 & 0 & \underline{F^{*\prime}} \end{pmatrix},$$

and

$$Q(w,z) = \begin{pmatrix} F_{11} - \underline{F_{11}} \\ F_{12} - \underline{F_{12}} \\ F^{*\prime} - \underline{F^{*\prime}} \\ g_{w1} - \underline{g_{w1}} \\ g_{w2} - \underline{g_{w2}} \\ g_{w3} - \underline{g_{w3}} \\ g_{r1} - \underline{g_{r1}} \\ g_{r2} - \underline{g_{r2}} \\ g_{r3} - \underline{g_{r3}} \end{pmatrix}.$$

Then the Taylor series approximation in equations (5.1) and (5.2) can be written as

$$W(w,z) = A(w,z)\,Q(w,z) \tag{5.3}$$

for each combination of (w, z). One might approximate the structural functionals by

1) Make an initial guess of (F_1, F^*, g_w, g_r) satisfying initial conditions and set $k = 0$.

2) Evaluate $(F_1, F^*, g_w, g_r)^{(k)}$ and use it to evaluate $W^{(k)}(w, z)$ and $A^{(k)}(w, z)$.

3) Solve for $Q^{(k+1)}(w, z)$ using

$$Q^{(k+1)}(w, z) = \left[A^{(k)}(w, z)\right]^{-1} W^{(k)}(w, z).$$

4) Using the definition of $Q(w, z)$, solve for $(F_1, F^*, g_w, g_r)^{(k+1)}$ given $Q^{(k+1)}(w, z)$.

5) Check for convergence. If not, increment k by 1 and go to (2).

Note that this algorithm does not put restrictions on (F_1, F^*, g_w, g_r) that occur across different values of (w, z). Such restrictions are that

$$\begin{aligned} F_1\left(w^{(1)} - g_w^{(1)}, w^{(1)} - g_r^{(1)}\right) &= F_1\left(w^{(2)} - g_w^{(2)}, w^{(2)} - g_r^{(2)}\right) \text{ if} \\ w^{(1)} - g_w^{(1)} &= w^{(2)} - g_w^{(2)} \text{ and} \\ w^{(1)} - g_r^{(1)} &= w^{(2)} - g_r^{(2)} \end{aligned}$$

and

$$\begin{aligned} F^*\left(g_r^{(1)} - g_w^{(1)}\right) &= F^*\left(g_r^{(2)} - g_w^{(2)}\right) \text{ if} \\ g_r^{(1)} - g_w^{(1)} &= g_r^{(2)} - g_w^{(2)}. \end{aligned}$$

Once the algorithm has converged, we can stack the equations in equation (5.3) over all values of (w, z) as $W = AQ$ and write restrictions (given estimates of g_w and g_r) as $BQ = 0$. Putting these together, we want to solve the first order conditions for the Lagrangian equation

$$\pounds = (W - AQ)'(W - AQ) + \lambda(BQ).$$

The solution is

$$Q - (A'A)^{-1} B' \left[B(A'A)^{-1} B'\right]^{-1} (BQ)$$

(I think). Note that this involves inverting $(A'A)$ (which is easy because it is block diagonal) and $\left[B(A'A)^{-1} B'\right]$ (which is not block diagonal but is pretty sparse, so it may not be that hard).