Supplement to: Binary Response Panel Data Models with Sample Selection and Self Selection
by Anastasia Semykina and Jeffrey M. Wooldridge

1. Asymptotic Variance of the Parametric Two-Step Estimator of Censored Selection Model

This Section discusses the estimation of the asymptotic variance of the estimator summarized in Procedure 3.2. Denote $\mathbf{w}_t = (1, \mathbf{x}_t, \mathbf{z}_t, \mathbf{v}_t)$, $\mathbf{q}_t = (1, \mathbf{z}_t, \mathbf{e}_t)$, and $\mathbf{v}_t = (\eta_1, \beta', \xi_1, \gamma)'$. Using the argument similar to that presented in Section 2 below, it can be shown that

$$\sqrt{N}(\hat{\theta} - \theta) \sim \text{Normal}(0, V),$$

where $V = A^{-1}BA^{-1}$ is the asymptotic variance of $\sqrt{N}(\hat{\theta} - \theta)$,

$$A = E\left[ -\sum_{t=1}^{T} 1[s_{it} > 0] \cdot H_{it}(\theta) \right],$$
$$B = E[p_i'p_i'],$$
$$p_i = \sum_{t=1}^{T} 1[s_{it} > 0] \cdot \left\{ \frac{\phi(w_{it}\theta)}{\Phi(w_{it}\theta)[1 - \Phi(w_{it}\theta)]} w_{it}'y_{it} - \Phi(w_{it}\theta) - Fr_i(\pi) \right\},$$
$$F = -E\left[ \sum_{t=1}^{T} 1[s_{it} > 0] \cdot \phi(w_{it}\theta)^2 \frac{w_{it}'q_{it}\gamma}{[1 - \Phi(w_{it}\theta)]} \right],$$
$$r_i(\pi) = \left[ E\left( -\sum_{t=1}^{T} H_{it}(\pi) \right) \right]^{-1} \sum_{t=1}^{T} S_{it}(\pi).$$

Here $H_{it}(\theta)$ is the Hessian matrix from the second-step probit estimation, while $H_{it}(\pi)$ and $S_{it}(\pi)$ are the Hessian matrix and score vector from the first-step Tobit estimation, respectively.
Then, $A\varphi(\hat{\theta})$ can be estimated as
\[
\hat{A} = -\frac{1}{N} \sum_{i=1}^{N} \sum_{t=1}^{T} 1[s_{it} > 0] \cdot H_{it}(\hat{\theta}),
\]
(3)

\[
\hat{B} = \frac{1}{N} \sum_{i=1}^{N} \hat{p}_i \hat{p}_i',
\]

\[
\hat{p}_i = \sum_{t=1}^{T} 1[s_{it} > 0] \cdot \left\{ \frac{\phi(\hat{w}_{it}\hat{\theta})}{\Phi(\hat{w}_{it}\hat{\theta})[1 - \Phi(\hat{w}_{it}\hat{\theta})]} \hat{w}_{it}' y_{it} - \Phi(\hat{w}_{it}\hat{\theta}) \right\},
\]

\[
\hat{F} = -\frac{1}{N} \sum_{i=1}^{N} \sum_{t=1}^{T} \left[ \frac{1}{1 - \Phi(\hat{w}_{it}\hat{\theta})} \hat{w}_{it}' q_{it} \hat{\gamma} \right],
\]

\[
r_i(\hat{\pi}) = \left[ -\frac{1}{N} \sum_{i=1}^{N} \sum_{t=1}^{T} H_{it}(\hat{\pi}) \right]^{-1} \sum_{t=1}^{T} S_{it}(\hat{\pi}),
\]
(4)

where $H_{it}(\hat{\theta})$ is the Hessian matrix from the second-step probit estimation evaluated at $\hat{\theta}$, while $H_{it}(\hat{\pi})$ and $S_{it}(\hat{\pi})$ are the Hessian matrix and score vector from the first-step Tobit estimation, respectively, evaluated at $\hat{\pi}$.

2. Asymptotic Properties of the Semiparametric Estimator

In this Section we discuss asymptotic properties of the semiparametric estimator proposed in Section 4 of the paper. The argument below is very similar to the one in Blundell and Powell (2004).

To demonstrate the consistency of the semiparametric estimator, first show that $\hat{S}^t$ is consistent for $\Sigma_0^t$, $t = 1, ..., T$, where $\Sigma_0^t$ is a particular form of matrix $\Sigma_0^t$ that uses the weighting matrix specified in equation (55) in the paper. Using the first-order mean-value
expansion, for each $t$ we can write:

$$
\hat{S}_t = S_0^t + S_1^t,
$$

where

$$
S_1^t = \left( \frac{n}{2} \right) \sum_{i<j} \omega_{ij}^l (w_{it} - w_{jt})'(w_{it} - w_{jt}), \quad l = 0, 1,
$$

where

$$
\omega_{ij}^0 = \frac{1}{h_\omega} \kappa_g \left( \frac{g_i - g_j}{h_\omega} \right) \kappa_v \left( \frac{v_{i2} - v_{j2}}{h_\omega} \right) d_{it} \cdot d_{jt} \cdot \tau_{it} \cdot \tau_{jt},
$$

and

$$
\omega_{ij}^1 = \frac{1}{h_\omega} \left\{ \kappa_g \left( \frac{g_i}{h_\omega} \right) \kappa_v \left( \frac{v_{i2}}{h_\omega} \right) (\hat{g}_{it} - g_{it} - \hat{g}_{jt} + g_{jt}) 
- \kappa_g \left( \frac{g_{ij}}{h_\omega} \right) \kappa_v \left( \frac{v_{ij2}}{h_\omega} \right) (q_{it} - q_{jt})(\hat{\pi}_t - \pi_t) \right\} d_{it} \cdot d_{jt} \cdot \tau_{it} \cdot \tau_{jt},
$$

where $\kappa_g(\cdot)$ and $\kappa_v(\cdot)$ are vectors of first derivatives of functions $\kappa_g(\cdot)$ and $\kappa_v(\cdot)$, respectively, $q_{it} = (1, z_{i1}, \ldots, z_{it})$, $\pi_t = (\eta_{2t}, \xi_{21}, \ldots, \delta_t + \xi_{2t}, \ldots, \xi_{2T})'$, and $\hat{\pi}_t$ is the first-step estimator of $\pi_t$.

Similar to Blundell and Powell (2004), if the first four moments of $r_{it}$ and $s_{it}$ are finite, and $\kappa_g(\cdot)$, $\kappa_v(\cdot)$, $\tau_{it}$, $\tau_{jt}$ are bounded, then $S_0^t = \Sigma_0^t + o_p(1), \quad t = 1, \ldots, T$, when $h_\omega \to 0$, $h_\omega N \to \infty$.

To show that $S_1^t$ converges in probability to zero, $t = 1, \ldots, T$, assume that functions $\kappa_g(\cdot)$, $\kappa_v(\cdot)$, $\kappa_g^{(1)}(\cdot)$, $\kappa_v^{(1)}(\cdot)$ are uniformly bounded, and the first two moments of $q_{it}$ exist.

For the first-step Powell’s censored least absolute deviations estimator (Powell, 1984) or symmetrically trimmed censored least squares estimator (Powell, 1986), assume that the appropriate regularity conditions hold, so that $\hat{\pi}_t$ is $\sqrt{N}$-consistent for all $t$. Moreover, assume that regularity conditions provided in Ahn and Powell (1993) are satisfied. These include smoothness assumptions for conditional expectation and density functions conditional on $g_{it} = g_{jt}$ and $v_{i2} = v_{j2}$, as well as the restrictions on the speed with which $h_g$ and $h_\omega$ converge to zero as $N \to \infty$, where both depend on the dimensionality of the continuous component of $w_{it}$. There is also a requirement that higher-order (bias-reducing) kernels are used at both steps. The second-step kernels, $\kappa_g$ and $\kappa_v$, are assumed to be fourth-order kernels with the first three moments being equal to zero. For the first-
step kernel, \( K \), the number of vanishing moments depends on the number of continuous variables in \( w_{it} \). If these assumptions hold, the biases resulting from the nonparametric estimation of \( g_{it} \) and \( \omega_{ijt} \) are of the order smaller than \( \sqrt{N} \).

From above, it follows that under the specified conditions,

\[
\hat{S} = \sum_{t=1}^{T} \hat{S}_t = \sum_{t=1}^{T} \Sigma_t + o_p(1) \equiv \Sigma_0 + o_p(1). \tag{9}
\]

Moreover, using the law of iterated expectations:

\[
\Sigma_t = E[f_{it} \cdot d_{it} \cdot d_{jt} \cdot \tau_{it} \cdot \tau_{jt} \cdot (w_{it} - w_{jt})'(w_{it} - w_{jt})|g_{it} = g, v_{it2} = v] = E \{2f_{it} \cdot (g_{it}\mu_{w, it} - \mu'_{w, it}\mu_{w, it})\}, \quad t = 1, \ldots, T, \tag{10}
\]

\[
\varrho_{it} = E[d_{it} \cdot \tau_{it}|g_{it} = g, v_{it2} = v],
\]

\[
\mu_{w, it} = E[d_{it} \cdot \tau_{it} \cdot w_{it}|g_{it} = g, v_{it2} = v],
\]

\[
\mu_{w, w, it} = E[d_{it} \cdot \tau_{it} \cdot w'_{it}w_{it}|g_{it} = g, v_{it2} = v]. \tag{11}
\]

Furthermore, \( \Sigma_0 \theta = 0 \) because

\[
\sum_{t=1}^{T} [\varrho_{it}\mu_{w, w, it} - \mu'_{w, w, it}\mu_{w, it}] \theta = \sum_{t=1}^{T} [\varrho_{it}E(w'_{it}w_{it}\theta|g_{it}, v_{it2}) - \mu'_{w, it}E(w_{it}\theta|g_{it}, v_{it2})] = \sum_{t=1}^{T} (\varrho_{it}\mu'_{w, w, it}g_{it} - \varrho_{it}\mu'_{w, it}g_{it}) = 0, \tag{12}
\]

where we use the fact that \( w_{it}\theta = g_{it}, t = 1, \ldots, T. \)

Finally, the identification condition has to hold. Regarding the first-step estimation, necessary identification conditions for the censored least absolute deviations estimator and symmetrically trimmed least squares estimator are provided in Powell (1984) and Powell (1986), respectively. The second part of the identification condition is that in the population, \( \theta \) is a unique nontrivial solution to \( \Sigma_0 \theta = 0 \) after the normalization \( \theta = (1, \alpha')' \) is imposed. Specifically, assume that matrix \( \Sigma_0^{22} \), which is the lower-right \((M + L - 1) \times (M + L - 1)\) sub-matrix of matrix \( \Sigma_0 \), has full rank. This completes the consistency argument.

In order to establish \( \sqrt{N} \)-asymptotic normality, we first use the second order mean
value expansion to write

\[ \hat{S} = S_0 + S_1 + S_2 = \sum_{t=1}^{T} S'_{0t} + \sum_{t=1}^{T} S'_{1t} + \sum_{t=1}^{T} S'_{2t}, \]  

(13)

where

\[ S'_l = \left( \begin{array}{c} n \end{array} \right)^{-1} \sum_{i<j} \omega'_{ijl} (w_{it} - w_{jt})' (w_{it} - w_{jt}), \quad l = 0, 1, \]  

(14)

\[ \omega^0_{ijl} = \frac{1}{h^2_{\omega}} \kappa_g \left( \frac{g_{it} - g_{jt}}{h_{\omega}} \right) \kappa_v \left( \frac{v_{it2} - v_{jt2}}{h_{\omega}} \right) d_{it} \cdot d_{jt} \cdot \tau_{it} \cdot \tau_{jt}, \]  

(15)

\[ \omega^1_{ijl} = \frac{1}{h^3_{\omega}} \left\{ \kappa_g^{(2)} \left( \frac{g_{it} - g_{jt}}{h_{\omega}} \right) \kappa_v^{(1)} \left( \frac{v_{it2} - v_{jt2}}{h_{\omega}} \right) (\hat{g}_{it} - g_{it} - \hat{g}_{jt} + g_{jt}) \right. \]  

\[ - \left. \kappa_g \left( \frac{g_{it} - g_{jt}}{h_{\omega}} \right) \kappa_v^{(1)} \left( \frac{v_{it2} - v_{jt2}}{h_{\omega}} \right) (q_{it} - q_{jt}) (\hat{\pi}_t - \pi_t) \right\} d_{it} \cdot d_{jt} \cdot \tau_{it} \cdot \tau_{jt}, \]  

(16)

\[ \omega^2_{ijl} = \frac{1}{2h^4_{\omega}} \left\{ \kappa_g^{(2)} \left( \frac{g_{it} - g_{jt}}{h_{\omega}} \right) \kappa_v^{(1)} \left( \frac{v_{it2} - v_{jt2}}{h_{\omega}} \right) (\hat{g}_{it} - g_{it} - \hat{g}_{jt} + g_{jt})^2 \right. \]  

\[ - 2 \kappa_g^{(1)} \left( \frac{g_{it} - g_{jt}}{h_{\omega}} \right) \kappa_v^{(1)} \left( \frac{v_{it2} - v_{jt2}}{h_{\omega}} \right) (q_{it} - q_{jt}) (\hat{\pi}_t - \pi_t) (\hat{g}_{it} - g_{it} - \hat{g}_{jt} + g_{jt}) \]  

\[ + \left. \kappa_g \left( \frac{g_{it} - g_{jt}}{h_{\omega}} \right) \kappa_v^{(2)} \left( \frac{v_{it2} - v_{jt2}}{h_{\omega}} \right) (q_{it} - q_{jt}) (\hat{\pi}_t - \pi_t) (\hat{\pi}_t - \pi_t)' (q_{it} - q_{jt})' \right\} d_{it} \cdot d_{jt} \cdot \tau_{it} \cdot \tau_{jt}, \]  

(17)

Under assumptions stated in Ahn and Powell (1993), using \( \sqrt{N} \)-consistency of the first-step estimator \( \hat{\pi} \), and following the same argument as in Blundell and Powell (2004), it should be the case that

\[ \sqrt{N} S_0 \theta = o_p(1), \quad \sqrt{N} S_2 \theta = o_p(1). \]  

(18)

Furthermore, when the selection equation is estimated using either Powell’s censored least absolute deviations estimator or symmetrically trimmed censored least squares estimator, \( \hat{\pi} \) satisfies

\[ \sqrt{N} (\hat{\pi} - \pi) = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} m_i + o_p(1), \]  

where \( \text{E}(m_i) = 0 \), and \( \text{E}(m_i m_i') \) exists and is nonsingular.

Then, we can show that

\[ \sqrt{N} S \theta = \sqrt{N} S_1 \theta + o_p(1) = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} (e_{i1} + e_{i2}) + o_p(1), \]  

(19)
where
\[ e_{i1} = \sum_{t=1}^{T} 2f_{it} \psi(g_{it}w_{it} - \mu_{w,it})' \cdot \frac{\partial \psi(g_{it}, \mu_{w})}{\partial g_{it}} \cdot [y_{it} - g(r_{it})], \]
\[ e_{i2} = -Fm_{i}(\pi), \]
\[ F = E \left[ \sum_{t=1}^{T} 2f_{it} \psi(g_{it}w_{it} - \mu_{w,it})' \cdot \frac{\partial \psi(g_{it}, \mu_{w})}{\partial \mu_{w}} \cdot q_{it} \right]. \]  

(20)

If the censored least absolute deviations estimator (Powell, 1984) is used as the estimator of \( \pi \), and the first-step estimation is performed separately for each \( t \), then
\[ m_{it}(\pi) = \begin{pmatrix} m_{i1}(\pi_1) \\ \vdots \\ m_{iT}(\pi_T) \end{pmatrix}, \]
\[ m_{it}(\pi_t) = [f_t(0) \cdot J_t]^{-1} \cdot 1[q_{it} \pi_t > 0] \cdot q_{it}' \left( \frac{1}{2} - 1[v_{it2} > 0] \right), \]
\[ J_t = E \left( 1[q_{it} \pi_t > 0] \cdot q_{it}' q_{it} \right), \quad t = 1, \ldots, T, \]  

(21)

where \( f_t(\cdot) \) is the density function of error \( v_{it2} \) in period \( t \).

If \( \pi_t, t = 1, \ldots, T, \) is estimated using the symmetrically trimmed censored least squares estimator (Powell, 1986), then
\[ m_{it}(\pi_t) = C_t^{-1} \cdot 1[q_{it} \pi_t > 0] \cdot q_{it}' \left( \min\{s_{it}, 2q_{it} \pi_t\} - q_{it} \pi_t \right), \]
\[ C_t = E \left\{ 1[-q_{it} \pi_t < v_{it2} < q_{it} \pi_t] \cdot q_{it}' q_{it} \right\}, \quad t = 1, \ldots, T. \]  

(22)

From (12) and (19) it follows that
\[ \sqrt{N} \hat{\theta} \tilde{\theta} = o_p(1), \]  

(23)

so that for the subvector \( \hat{\alpha} \) of \( \hat{\theta} = (1, \hat{\alpha}')' \), we obtain
\[ \sqrt{N}(\hat{\alpha} - \alpha) \xrightarrow{d} Normal(0, \Sigma_{22}^{-1}V_{22}^{-1}), \]  

(24)

where \( \Sigma_{22} \) is the lower \((M + L - 1) \times (M + L - 1)\) diagonal submatrix of \( \Sigma_0 \), and \( V_{22} \) is the lower \((M + L - 1) \times (M + L - 1)\) diagonal submatrix of \( V \),
\[ V = \text{Var}(e_{i1} + e_{i2}) = E[(e_{i1} + e_{i2})(e_{i1} + e_{i2})']. \]  

(25)

Note that this is a robust form of the variance that accounts for serial dependence in the errors.
To obtain a consistent estimator of $\text{Avar}[\sqrt{N}(\hat{\alpha} - \alpha)]$, first note that $\hat{S}$ is consistent for $\Sigma_0$. Furthermore, using the argument similar to the one in Ahn and Powell (1993), a consistent estimator of $V$ would be

$$
\hat{V} = \frac{1}{N} \sum_{i=1}^{N} [ (\hat{e}_{i1} + \hat{e}_{i2}) (\hat{e}_{i1} + \hat{e}_{i2}) ]' \tag{26}
$$

where

$$
\hat{e}_{i1} = \sum_{t=1}^{T} \frac{1}{N-1} \sum_{j=1}^{N} \left[ \frac{2}{h_w^3} K_g' \left( \frac{\hat{g}_{it} - \hat{g}_{jt}}{h_w} \right) K_v \left( \frac{\hat{v}_{it2} - \hat{v}_{jt2}}{h_w} \right) d_{it} d_{jt} \tau_{it} \tau_{jt} \hat{\delta}_{jt} \right] (y_{it} - \hat{g}_{it}),
$$

$$
\hat{\delta}_{jt} = \frac{\sum_{l=1}^{N} K \left( \frac{\hat{r}_{jl} - \hat{r}_{lt}}{h_w} \right) (w_{jt} - w_{lt})'}{\sum_{l=1}^{N} K \left( \frac{\hat{r}_{jl} - \hat{r}_{lt}}{h_w} \right)}, \tag{27}
$$

and

$$
\hat{e}_{i2} = -\hat{F} m_{i}(\hat{\pi}), \tag{28}
$$

$$
\hat{F} = \frac{1}{N} \sum_{i=1}^{N} \sum_{t=1}^{T} \frac{1}{N-1} \sum_{j=1}^{N} \left[ \frac{2}{h_w^3} K_g \left( \frac{\hat{g}_{it} - \hat{g}_{jt}}{h_w} \right) K_v \left( \frac{\hat{v}_{it2} - \hat{v}_{jt2}}{h_w} \right) d_{it} d_{jt} \tau_{it} \tau_{jt} \hat{\delta}_{jt} \right] q_{it},
$$

for $m_{i}(\hat{\pi}) = [m_{i1}(\hat{\pi}_1), \ldots, m_{iT}(\hat{\pi}_T)]'$, and $m_{i}(\hat{\pi}_t)$ defined as in either (21) or (22), but evaluated at $\hat{\pi}_t$.

References


Table 1: Simulation results for $\hat{\beta}_2/\hat{\beta}_1 (\beta_2/\beta_1 = 0.6), u_{it1} \sim \chi^2_3$

<table>
<thead>
<tr>
<th></th>
<th>Probit time means</th>
<th>Probit, $s_{it}$ censored, $\sigma_a^2 = 0$, $\xi_1 = 0$, $\rho = 0$</th>
<th>2-step MLE</th>
<th>full MLE</th>
<th>Semiparametric</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>N=500</strong> Bias</td>
<td>0.0006</td>
<td>0.0006</td>
<td>0.0052</td>
<td>0.0051</td>
<td></td>
</tr>
<tr>
<td>RMSE</td>
<td>0.0482</td>
<td>0.0584</td>
<td>0.0612</td>
<td>0.0594</td>
<td>0.0698</td>
</tr>
<tr>
<td>Average se</td>
<td>0.0477</td>
<td>0.0583</td>
<td>0.0613</td>
<td>0.0594</td>
<td>0.0693</td>
</tr>
<tr>
<td>Bootstrap se</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.0815</td>
</tr>
</tbody>
</table>

| **N=500** Bias| -0.0630           | -0.0008                                                            | 0.0016     | 0.0098   |                |
| RMSE           | 0.0912            | 0.0654                                                             | 0.0698     | 0.0672   | 0.0812         |
| Average se     | 0.0641            | 0.0673                                                             | 0.0711     | 0.0684   | 0.0795         |
| Bootstrap se   |                   |                                                                    |            |          | 0.0945         |

| **N=500** Bias| -0.1237           | -0.0459                                                            | 0.0011     | 0.0042   | -0.0048        |
| RMSE           | 0.1402            | 0.0816                                                             | 0.0698     | 0.0677   | 0.0779         |
| Average se     | 0.0680            | 0.0690                                                             | 0.0712     | 0.0687   | 0.0764         |
| Bootstrap se   |                   |                                                                    |            |          | 0.0988         |

| **N=1000** Bias| 0.0009            | 0.0021                                                             | 0.0007     | 0.0052   | 0.0057         |
| RMSE           | 0.0350            | 0.0426                                                             | 0.0446     | 0.0438   | 0.0471         |
| Average se     | 0.0338            | 0.0413                                                             | 0.0434     | 0.0420   | 0.0509         |
| Bootstrap se   |                   |                                                                    |            |          | 0.0539         |

| **N=1000** Bias| -0.0617           | -0.0006                                                            | 0.0020     | 0.0070   |                |
| RMSE           | 0.0775            | 0.0495                                                             | 0.0517     | 0.0506   | 0.0557         |
| Average se     | 0.0455            | 0.0476                                                             | 0.0502     | 0.0484   | 0.0476         |
| Bootstrap se   |                   |                                                                    |            |          | 0.0623         |

| **N=1000** Bias| -0.1213           | -0.0455                                                            | 0.0004     | 0.0037   | -0.0026        |
| RMSE           | 0.1310            | 0.0672                                                             | 0.0510     | 0.0501   | 0.0529         |
| Average se     | 0.0482            | 0.0489                                                             | 0.0503     | 0.0485   | 0.0573         |
| Bootstrap se   |                   |                                                                    |            |          | 0.0643         |

| **N=2500** Bias| 0.0017            | 0.0018                                                             | -0.00002   | 0.0049   | 0.0033         |
| RMSE           | 0.0218            | 0.0262                                                             | 0.0274     | 0.0270   | 0.0294         |
| Average se     | 0.0214            | 0.0261                                                             | 0.0274     | 0.0265   | 0.0469         |
| Bootstrap se   |                   |                                                                    |            |          | 0.0463         |

| **N=2500** Bias| -0.0605           | 0.0023                                                             | 0.0008     | 0.0033   | 0.0044         |
| RMSE           | 0.0672            | 0.0303                                                             | 0.0326     | 0.0310   | 0.0357         |
| Average se     | 0.0288            | 0.0301                                                             | 0.0318     | 0.0306   | 0.0275         |

| **N=2500** Bias| -0.1215           | -0.0458                                                            | 0.0002     | 0.0032   | -0.0047        |
| RMSE           | 0.1255            | 0.0550                                                             | 0.0309     | 0.0301   | 0.0327         |
| Average se     | 0.0305            | 0.0309                                                             | 0.0317     | 0.0306   | 0.0476         |
Table 2: Descriptive Statistics for NLSY79 data

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>31.07</td>
</tr>
<tr>
<td></td>
<td>(2.63)</td>
</tr>
<tr>
<td>Education (years)</td>
<td>13.51</td>
</tr>
<tr>
<td></td>
<td>(2.35)</td>
</tr>
<tr>
<td>AFQT score</td>
<td>54.05</td>
</tr>
<tr>
<td></td>
<td>(26.16)</td>
</tr>
<tr>
<td>Married (%)</td>
<td>69.39</td>
</tr>
<tr>
<td>Number of observations</td>
<td>8,340</td>
</tr>
</tbody>
</table>

Sample standard deviations are in parentheses below the sample means