Supplemental Material for "Testing the Unconfoundedness Assumption via Inverse Probability Weighted Estimators of (L)ATT"

Stephen G. Donald* Yu-Chin Hsu† Robert P. Lieli‡

October 20, 2013

*Department of Economics, University of Texas, Austin, donald@eco.utexas.edu.

†Institute of Economics, Academia Sinica, ychsu@econ.sinica.edu.tw

‡Department of Economics, Central European University, Budapest and the National Bank of Hungary, rlieli@gmail.com.

Appendix

A. Identification

The derivation of equations (1) and (2) We can write

$$\begin{aligned} E[W(1)-W(0)] &= E\{[D(1)-D(0)][Y(1)-Y(0)]\} \\ &= E[Y(1)-Y(0) \mid D(1)-D(0)=1]P[D(1)-D(0)=1] = \tau \cdot E[D(1)-D(0)], \end{aligned}$$

where the second equality holds because under monotonicity (Assumption 1(v)), $D(1)-D(0)$ is either zero or one. Similarly,

$$\begin{aligned} E[W(1)-W(0) \mid Z=1] &= E\{[D(1)-D(0)][Y(1)-Y(0) \mid Z=1]\} \\ &= E[Y(1)-Y(0) \mid D(1)-D(0)=1, Z=1]P[D(1)-D(0)=1 \mid Z=1] \\ &= \tau_t \cdot E[D(1)-D(0) \mid Z=1], \end{aligned}$$

where the third equality holds because under monotonicity, $D(1)-D(0)=1$ implies $D=Z$. ∎

B. The proof of Theorem 1

In order to simplify notation, we set $X_1 = X$. Let

$$\widehat{\Delta} = \frac{1}{n}\sum_{i=1}^{n}\left\{\frac{Z_iY_i}{\hat{q}(X_i)} - \frac{(1-Z_i)Y_i}{1-\hat{q}(X_i)}\right\}, \qquad \widehat{\Gamma} = \frac{1}{n}\sum_{i=1}^{n}\left\{\frac{Z_iD_i}{\hat{q}(X_i)} - \frac{(1-Z_i)D_i}{1-\hat{q}(X_i)}\right\}.$$

so that $\hat{\tau} = \widehat{\Delta}/\widehat{\Gamma}$. The asymptotic properties of $\hat{\Delta}$ and $\hat{\Gamma}$ are established in the following lemma.

Lemma 3 *Under the conditions of Theorem 1,* $\sqrt{n}(\widehat{\Delta}-\Delta) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\delta(Y_i, D_i, Z_i, X_i) + o_p(1)$ *and* $\sqrt{n}(\widehat{\Gamma}-\Gamma) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\gamma(Y_i, D_i, Z_i, X_i) + o_p(1)$*, where*

$$\begin{aligned} \delta(Y_i, D_i, Z_i, X_i) &= \frac{Z_iY_i}{q(X_i)} - \frac{(1-Z_i)Y_i}{1-q(X_i)} - \Delta - \left(\frac{m_1(X_i)}{q(X_i)} + \frac{m_0(X_i)}{1-q(X_i)}\right)(Z_i - q(X_i)) \\ \gamma(Y_i, D_i, Z_i, X_i) &= \frac{Z_iD_i}{q(X_i)} - \frac{(1-Z_i)D_i}{1-q(X_i)} - \Gamma - \left(\frac{\mu_1(X_i)}{q(X_i)} + \frac{\mu_0(X_i)}{1-q(X_i)}\right)(Z_i - q(X_i)) \end{aligned}$$

The proof of Lemma 3 Recall the definition of $W(z)$. By Assumption 1(ii), it is true that $E[W(z)|Z,X] = E[W(z)|X]$, $z = 0, 1$. That is, if we treat Z as the treatment assignment and $W(z)$ as the potential outcomes, $W(z)$ and Z are unconfounded given X. Also, it is straightforward to check that Assumptions 1-5 of Thm. 1 of HIR are satisfied. The result for $\widehat{\Delta}$ follows directly from it. A similar argument applies to $\widehat{\Gamma}$. ∎

Taking a first order Taylor expansion of $\hat{\Delta}/\hat{\Gamma}$ around the point (Δ, Γ) yields

$$\sqrt{n}(\hat{\tau}-\tau) = \sqrt{n}\left(\frac{\widehat{\Delta}}{\widehat{\Gamma}} - \frac{\Delta}{\Gamma}\right) = \frac{1}{\Gamma}\sqrt{n}(\widehat{\Delta}-\Delta) - \frac{\tau}{\Gamma}\sqrt{n}(\widehat{\Gamma}-\Gamma) + o_p(1). \tag{14}$$

Applying Lemma 3 to (14) gives (8). Under Assumption 1(i), we have $E[\psi(Y,D,Z,X)]=0$ and $E[\psi^2(Y,D,Z,X)]<\infty$. Applying the Lindeberg-Levy CLT to (8) shows $\sqrt{n}(\hat{\tau}-\tau)\xrightarrow{d}\mathcal{N}(0,\mathcal{V})$. Let

$$\widehat{\Delta}_t=\sum_{i=1}^{n}\Big\{\hat{q}(X_i)\Big(\frac{Z_iY_i}{\hat{q}(X_i)}-\frac{(1-Z_i)Y_i}{1-\hat{q}(X_i)}\Big)\Big\}\Big/\sum_{i=1}^{n}\hat{q}(X_i),$$
$$\widehat{\Gamma}_t=\sum_{i=1}^{N}\Big\{\hat{q}(X_i)\Big(\frac{Z_iD_i}{\hat{q}(X_i)}-\frac{(1-Z_i)D_i}{1-\hat{q}(X_i)}\Big)\Big\}\Big/\sum_{i=1}^{n}\hat{q}(X_i),$$

so that $\hat{\tau}_t=\widehat{\Delta}_t/\widehat{\Gamma}_t$. Then the second part Theorem 1(a) can be shown after we show the following lemma:

Lemma 4 *Under the conditions of Theorem 1,*

$$\sqrt{n}(\widehat{\Delta}_t-\Delta_t)=\frac{1}{\sqrt{n}}\frac{1}{E(Z)}\sum_{i=1}^{n}q(X_i)\Big\{\frac{Z_i(Y_i-m_1(X_i))}{q(X_i)}-\frac{(1-Z_i)(Y_i-m_0(X_i))}{1-q(X_i)}+\frac{(m_1(X_i)-m_0(X_i)-\Delta_q)Z_i}{q(X_i)}\Big\}+o_p(1),$$
$$\sqrt{n}(\widehat{\Gamma}_t-\Gamma_t)=\frac{1}{\sqrt{n}}\frac{1}{E(Z)}\sum_{i=1}^{n}q(X_i)\Big\{\frac{Z_i(D_i-\mu_1(X_i))}{q(X_i)}-\frac{(1-Z_i)(D_i-\mu_0(X_i))}{1-q(X_i)}+\frac{(\mu_1(X_i)-\mu_0(X_i)-\Gamma_q)Z_i}{q(X_i)}\Big\}+o_p(1).$$

The proof of Lemma 4 The proof is identical to Lemma 3 with Corollary 1 of HIR in place of Theorem 1 of HIR. ∎

C. The proof of Theorem 2 and Lemma 1

The proof of Theorem 2: The proof of Theorem 2 follows from Theorem 1 and Corollary 1 of HIR. ∎

The proof of Lemma 1: Note that $Var(Y(0))=0$ implies that $P[Y(0)=a]=1$ for some $a\in R$. Define $Y^*(1)=Y(1)-a$, $Y^*(0)=Y(0)-a$ and $Y^*=D(Y^*(1))+(1-D)(Y^*(0))$. It is obvious that $Y^*(0)=0$. Hence, $\hat{\tau}_t$ and $\hat{\beta}_t$ can be written as

$$\hat{\tau}_t=\frac{\sum_{i=1}^{n}Z_iY_i^*}{\sum_{i=1}^{n}Z_iD_i}+a\Bigg(\frac{\frac{1}{n}\sum_{i=1}^{n}Z_i-\frac{\hat{q}(X_{1i})(1-Z_i)}{1-\hat{q}(X_{1i})}\Big/\frac{1}{n}\sum_{i=1}^{n}\hat{q}(X_{1i})}{\frac{1}{n}\sum_{i}^{n}Z_iD_i-\frac{\hat{q}(X_{1i})(1-Z_i)D_i}{1-\hat{q}(X_{1i})}\Big/\frac{1}{n}\sum_{i=1}^{n}\hat{q}(X_{1i})}\Bigg),$$
$$\hat{\beta}_t=\frac{\sum_{i=1}^{n}D_iY_i^*}{\sum_{i=1}^{n}\hat{p}(X_{2i})}+a\Bigg(\frac{\frac{1}{n}\sum_{i=1}^{n}D_i-\frac{\hat{p}(X_{2i})(1-D_i)}{1-\hat{p}(X_{2i})}}{\frac{1}{n}\sum_{i=1}^{n}\hat{p}(X_{2i})}\Bigg)$$

The proof follows easily from the following four results:

$$\sqrt{n}\Big(\frac{\sum_{i=1}^n D_i Y_i^*}{\sum_{i=1}^n \hat{p}(X_{2i})} - \frac{\sum_{i=1}^n Z_i Y_i^*}{\sum_{i=1}^n Z_i D_i}\Big) = o_p(1), \tag{15}$$

$$\sqrt{n}\Big(\frac{1}{n}\sum_i^n D_i - \frac{\hat{p}(X_{2i})(1-D_i)}{1-\hat{p}(X_{2i})}\Big/\frac{1}{n}\sum_i^n \hat{p}(X_{2i})\Big) = o_p(1), \tag{16}$$

$$\sqrt{n}\Big(\frac{1}{n}\sum_i^n Z_i - \frac{\hat{q}(X_{1i})(1-Z_i)}{1-\hat{q}(X_{1i})}\Big/\frac{1}{n}\sum_i^n \hat{q}(X_{1i})\Big) = o_p(1), \tag{17}$$

$$\Big(\frac{1}{n}\sum_i^n Z_i D_i - \frac{\hat{q}(X_{1i})(1-Z_i)D_i}{1-\hat{q}(X_{1i})}\Big)\Big/\Big(\frac{1}{n}\sum_i^n \hat{q}(X_{1i})\Big) \xrightarrow{p} \Gamma_t > 0. \tag{18}$$

We verify each of these equations in turn. If $P[Y^*(0)=0]=1$, then $Y^* = DY^*(1)$, and so $(1-Z)Y^*=0$ and $(1-D)Y^*=0$. Hence,

$$\frac{\sum_{i=1}^n Z_i Y_i^*}{\sum_{i=1}^n Z_i D_i} = \frac{\sum_{i=1}^n Z_i Y_i^*}{\sum_{i=1}^n Z_i D_i} = \frac{\sum_{i=1}^n D_i Y_i^*}{\sum_{i=1}^n D_i},$$

where the second equality holds since $ZY = ZDY(1) = DY(1) = DY$ and $ZD = D$. We first claim that

$$\sqrt{n}\Big(\frac{\sum_{i=1}^n D_i}{\sum_{i=1}^n \hat{p}(X_{2i})} - 1\Big) = o_p(1), \tag{19}$$

which implies

$$\sqrt{n}\Big(\frac{1}{\frac{1}{n}\sum_{i=1}^n D_i} - \frac{1}{\frac{1}{n}\sum_{i=1}^n \hat{p}(X_{2i})}\Big) = o_p(1). \tag{20}$$

To see (19), let $L(1)=1$ and $L(0)=0$, then in this case $E[L(1)-L(0)|D=1]=1$ and the unconfoundedness assumption holds automatically. Then $\sum_{i=1}^n D_i/\sum_{i=1}^n \hat{p}(X_{2i})$ is an estimator for $E[L(1)-L(0)|D=1]$ and by Corollary 1 of HIR, its asymptotic variance equal to 0 because $E[L(1)-L(0)|X_2]=1=ATT$, and $Var(L(1)|X_2) = Var(L(0)|X_2) = 0$ a.s. in X_2. Equation (20) further implies that

$$\begin{aligned}\sqrt{n}(\hat{\tau}_t - \hat{\beta}_t) &= \sqrt{n}\Big(\frac{\sum_{i=1}^n D_i Y_i}{\sum_{i=1}^n D_i} - \frac{\sum_{i=1}^n D_i Y_i}{\sum_{i=1}^n \hat{p}(X_{2i})}\Big) \\ &= \sqrt{n}\Big(\frac{1}{\frac{1}{n}\sum_{i=1}^n D_i} - \frac{1}{\frac{1}{n}\sum_{i=1}^n \hat{p}(X_{2i})}\Big)\big(\frac{1}{n}\sum_{i=1}^n D_i Y_i\big) = o_p(1)\cdot O_p(1) = o_p(1)\end{aligned}$$

and this shows (15).

Let $L^*(1)=1$, $L^*(0)=1$ and $L^* = DL^*(1) + (1-D)L^*(0)$. Then $\frac{1}{n}\sum_i^n (D_i - \frac{\hat{p}(X_{2i})(1-D_i)}{1-\hat{p}(X_{2i})})\big/\frac{1}{n}\sum_i^n \hat{p}(X_{2i})$ is an estimator for $E[L^*(1)-L^*(0)|D=1]=0$. The result in (16) follows from the same argument we used to establish (19). Equation (17) can be shown in a similar way and, finally, (18) follows from Lemma 4.

Taken together, equations (15)-(18) imply that whether or not the unconfoundedness assumption holds,

$$\begin{aligned}\sqrt{n}(\hat{\tau}_t - \hat{\beta}_t) =& a\sqrt{n}\Big(\frac{1}{n}\sum_i^n Z_i - \frac{\hat{q}(X_{1i})(1-Z_i)}{1-\hat{q}(X_{1i})}\Big)\cdot\frac{1}{n}\sum_i^n Z_i D_i - \frac{\hat{q}(X_{1i})(1-Z_i)D_i}{1-\hat{q}(X_{1i})}\Big) \\ &- a\sqrt{n}\Big(\frac{1}{n}\sum_i^n D_i - \frac{\hat{p}(X_{2i})(1-D_i)}{1-\hat{p}(X_{2i})}\Big) + \sqrt{n}\Big(\frac{\sum_{i=1}^n D_i Y_i^*}{\sum_{i=1}^n \hat{p}(X_{2i})} - \frac{\sum_{i=1}^n Z_i Y_i^*}{\sum_{i=1}^n Z_i D_i}\Big) \\ &= o_p(1)\cdot O_p(1) - o_p(1) + o_p(1) = o_p(1).\end{aligned}$$

This completes the proof. ∎

D. Additional simulation results

Here we present additional simulation results on the power properties of the unconfoundedness test. The results correspond to those presented in Section 5.2 for $b = 0.5$ except here $b = 0.25$ and fewer cases are reported. As the violation of the unconfoundedness assumption is less severe, power is generally lower. The table is a particularly good illustration of the phenomenon described in the last paragraph of Section 5.2. If this table was included in the paper, logically it would be Table C.5 in Appendix C.

Table C.5: Properties of the unconfoundedness test and the distribution of $\sqrt{n}[\widehat{(L)ATT} - (L)ATT]$ for various estimators: $b = 0.25$ (unconfoundedness does not hold)

				$n = 250$				$n = 500$				$n = 2500$			
q	Series ($\hat{q}$)	Trim.	Estimator	$\frac{\text{Mean}}{\sqrt{n}}$	s.e.	$\text{E}(\widehat{s.e.})$	MSE	$\frac{\text{Mean}}{\sqrt{n}}$	s.e.	$\text{E}(\widehat{s.e.})$	MSE	$\frac{\text{Mean}}{\sqrt{n}}$	s.e.	$\text{E}(\widehat{s.e.})$	MSE
Const.	quad.	.5%	$\hat{\tau}_t$ (LATT)	0.00	3.85	5.14	14.84	-0.00	2.81	2.80	7.89	-0.00	2.49	2.45	6.20
			$\hat{\beta}_t$ (ATT)	0.16	2.18	2.09	10.88	0.15	1.71	1.56	14.68	0.15	1.50	1.48	60.92
			Combined	0.15	2.10	1.76	10.38	0.15	1.74	1.54	14.07	0.15	1.51	1.48	60.38
			Pre-tested	0.13	2.64	1.88	11.06	0.07	3.17	1.92	12.43	0.00	2.67	2.43	7.16
			Power/$E(\hat{a})$	0.075/0.08				0.307/0.04				0.974/0.00			
Lin. 1	quad.	.5%	$\hat{\tau}_t$ (LATT)	0.02	7.32	11.41	53.69	0.01	4.42	4.40	19.53	-0.00	3.06	2.75	9.37
			$\hat{\beta}_t$ (ATT)	0.16	2.72	2.40	13.64	0.15	2.04	1.67	15.53	0.15	1.60	1.51	56.50
			Combined	0.17	2.56	1.94	14.00	0.15	1.99	1.61	15.42	0.15	1.62	1.51	57.21
			Pre-tested	0.17	3.24	2.06	17.98	0.11	3.28	1.89	16.41	0.01	3.67	2.60	13.57
			Power/$E(\hat{a})$	0.062/0.08				0.158/0.04				0.874/0.00			
Lin. 2	quad.	.5%	$\hat{\tau}_t$ (LATT)	0.00	6.71	5.19	44.97	0.01	4.50	3.88	20.28	0.00	3.80	3.65	14.48
			$\hat{\beta}_t$ (ATT)	0.13	9.33	4.95	91.47	0.13	2.93	2.53	16.63	0.12	2.58	2.50	40.75
			Combined	0.16	4.14	2.70	23.21	0.13	2.87	2.51	16.44	0.12	2.59	2.50	41.09
			Pre-tested	0.09	5.03	3.01	27.50	0.07	4.12	2.79	19.57	0.03	4.76	3.18	24.92
			Power/$E(\hat{a})$	0.154/0.11				0.176/0.02				0.578/0.00			
Rat. 1	quad.	.5%	$\hat{\tau}_t$ (LATT)	0.00	4.66	5.48	21.69	-0.01	3.24	2.96	10.58	-0.01	2.62	2.52	7.32
			$\hat{\beta}_t$ (ATT)	0.15	2.52	2.08	11.92	0.14	1.86	1.58	13.13	0.14	1.55	1.48	49.87
			Combined	0.15	2.34	1.78	11.33	0.14	1.88	1.55	12.79	0.14	1.58	1.48	49.29
			Pre-tested	0.13	3.00	1.93	12.99	0.06	3.41	1.95	13.20	-0.01	2.91	2.48	8.72
			Power/$E(\hat{a})$	0.085/0.08				0.279/0.04				0.955			
Rat. 2	quad.	.5%	$\hat{\tau}_t$ (LATT)	-0.00	3.32	3.83	11.00	-0.00	2.69	2.67	7.23	-0.00	2.50	2.49	6.29
			$\hat{\beta}_t$ (ATT)	0.15	2.07	1.89	9.68	0.14	1.74	1.62	13.46	0.14	1.60	1.57	53.72
			Combined	0.14	2.01	1.74	9.11	0.14	1.75	1.62	12.72	0.14	1.61	1.57	52.42
			Pre-tested	0.11	2.65	1.88	9.83	0.06	3.12	1.96	11.53	-0.00	2.74	2.46	7.50
			Power/$E(\hat{a})$	0.110/0.77				0.328/0.04				0.964/0.01			
Rat. 2	cube	.5%	$\hat{\tau}_t$ (LATT)	0.10	10.19	54.80	106.31	0.02	4.33	9.80	18.95	0.00	2.68	2.57	7.18
			$\hat{\beta}_t$ (ATT)	0.22	8.04	30.86	76.68	0.17	2.71	4.11	21.02	0.14	1.65	1.62	54.22
			Combined	0.32	4.50	5.46	46.16	0.16	2.56	2.33	19.99	0.14	1.70	1.60	51.63
			Pre-tested	0.33	4.64	5.48	48.20	0.15	2.92	2.43	19.35	0.01	3.05	2.50	9.39
			Power/$E(\hat{a})$	0.014/0.23				0.061/0.15				0.931/0.01			

E. Propensity score estimates

We present the propensity score estimates used to construct the second panel of Table 6.1 (X=three dummies, estimation method=sample splitting). Pages 7 and 8 show the model estimates for $p(X) = P(D = 1|X)$ for males and females and the distribution of the random variable $p(X)$. Pages 9 and 10 show the model estimates for $q(X) = P(Z = 1|X)$ for males and females and the distribution of the random variable $q(X)$. Of course, as Z is completely random, $\hat{q}(X) = E(Z)$+estimation error, and $\hat{q}(X)$ is centered tightly around $E(Z) = 0.67$.

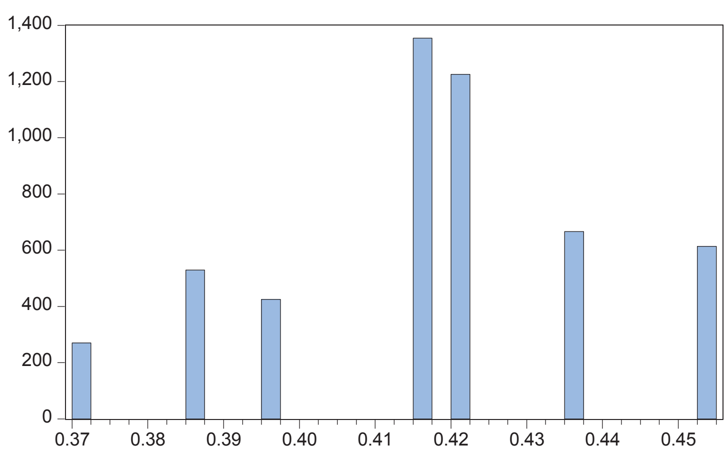
On pages 11 to 14 we present the propensity score estimates for the fourth panel of Table 6.1 in the paper. The order is $p(X)$ for males, $p(X)$ for females, $q(X)$ for males, $q(X)$ for females.

MALES

Dependent Variable: D (PARTICIPATION DUMMY)
Method: ML - Binary Logit (Quadratic hill climbing)
Date: 12/15/12 Time: 15:47
Sample: 1 15000 IF SEX=1 AND GOODOBS=1
Included observations: 5083
Convergence achieved after 4 iterations
Covariance matrix computed using second derivatives

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	-0.465705	0.089325	-5.213628	0.0000
HS	0.122646	0.104980	1.168273	0.2427
MINORITY	-0.064923	0.154471	-0.420293	0.6743
BELOW30	0.040593	0.133500	0.304067	0.7611
HS*MINORITY	0.148274	0.181689	0.816087	0.4145
HS*BELOW30	-0.019478	0.158146	-0.123162	0.9020
MINORITY*BELOW30	0.172500	0.226183	0.762658	0.4457
HS*MINORITY*BELOW30	-0.123397	0.266510	-0.463012	0.6434

McFadden R-squared	0.001396	Mean dependent var	0.416486
S.D. dependent var	0.493025	S.E. of regression	0.492897
Akaike info criterion	1.359516	Sum squared resid	1232.961
Schwarz criterion	1.369799	Log likelihood	-3447.210
Hannan-Quinn criter.	1.363117	Deviance	6894.421
Restr. deviance	6904.061	Restr. log likelihood	-3452.031
LR statistic	9.640533	Avg. log likelihood	-0.678184
Prob(LR statistic)	0.209876		
Obs with Dep=0	2966	Total obs	5083
Obs with Dep=1	2117		



Series: PHAT
Sample 1 15000 IF SEX=1 AND GOODOBS=1
Observations 5083

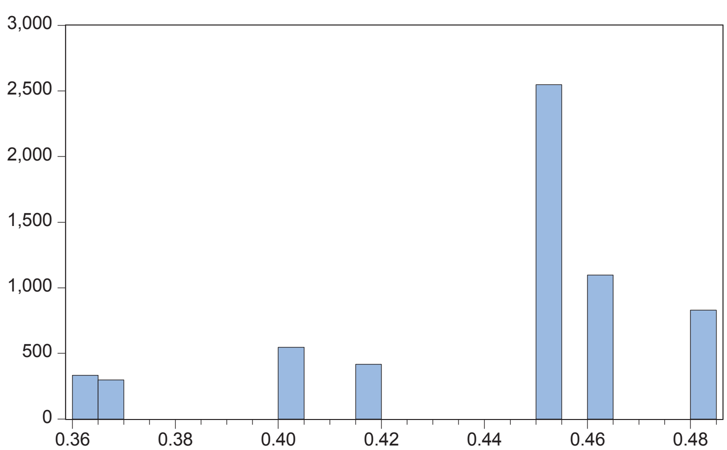
Mean	0.416486
Median	0.415066
Maximum	0.452769
Minimum	0.370370
Std. Dev.	0.021448
Skewness	-0.241952
Kurtosis	2.730917
Jarque-Bera	64.92868
Probability	0.000000

FEMALES

Dependent Variable: D (PARTICIPATION DUMMY)
Method: ML - Binary Logit (Quadratic hill climbing)
Date: 12/15/12 Time: 15:50
Sample: 1 15000 IF SEX=0 AND GOODOBS=1
Included observations: 6067
Convergence achieved after 4 iterations
Covariance matrix computed using second derivatives

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	-0.396333	0.087198	-4.545183	0.0000
HS	0.214662	0.099957	2.147545	0.0317
MINORITY	-0.177468	0.143637	-1.235526	0.2166
BELOW30	0.056563	0.132284	0.427585	0.6690
HS*MINORITY	0.176817	0.166495	1.061996	0.2882
HS*BELOW30	-0.033841	0.153481	-0.220488	0.8255
MINORITY*BELOW30	-0.027856	0.212181	-0.131286	0.8955
HS*MINORITY*BELOW30	0.124785	0.246177	0.506892	0.6122

McFadden R-squared	0.003584	Mean dependent var	0.442888
S.D. dependent var	0.496768	S.E. of regression	0.495842
Akaike info criterion	1.370934	Sum squared resid	1489.660
Schwarz criterion	1.379783	Log likelihood	-4150.728
Hannan-Quinn criter.	1.374005	Deviance	8301.456
Restr. deviance	8331.317	Restr. log likelihood	-4165.659
LR statistic	29.86169	Avg. log likelihood	-0.684148
Prob(LR statistic)	0.000101		
Obs with Dep=0	3380	Total obs	6067
Obs with Dep=1	2687		



Series: PHAT
Sample 1 15000 IF SEX=0 AND GOODOBS=1
Observations 6067

Mean	0.442888
Median	0.454707
Maximum	0.484337
Minimum	0.360360
Std. Dev.	0.034692
Skewness	-1.156645
Kurtosis	3.371242
Jarque-Bera	1387.606
Probability	0.000000

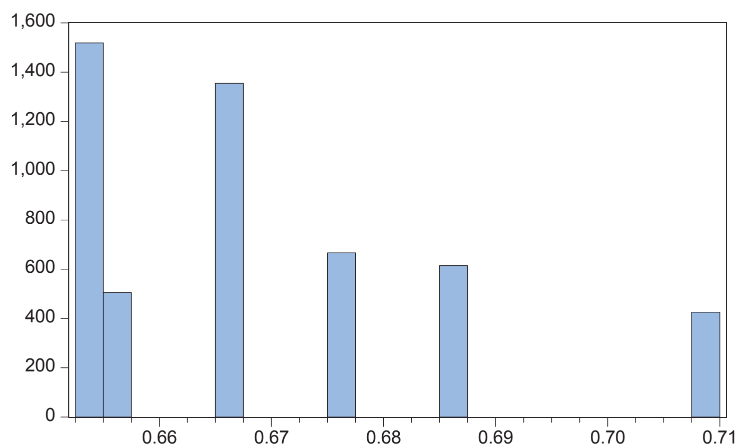
MALES

Dependent Variable: Z (RANDOM OFFER OF SERVICES)
Method: ML - Binary Logit (Quadratic hill climbing)
Date: 12/15/12 Time: 16:11
Sample: 1 15000 IF SEX=1 AND GOODOBS=1
Included observations: 5083
Convergence achieved after 4 iterations
Covariance matrix computed using second derivatives

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	0.636953	0.091404	6.968557	0.0000
HS	0.053980	0.108054	0.499562	0.6174
MINORITY	0.006598	0.157344	0.041931	0.9666
BELOW30	0.249876	0.140504	1.778424	0.0753
HS*MINORITY	0.036439	0.186896	0.194971	0.8454
HS*BELOW30	-0.301728	0.165921	-1.818511	0.0690
MINORITY*BELOW30	-0.250923	0.234485	-1.070104	0.2846
HS01*MINORITY*BELOW30	0.348749	0.277812	1.255339	0.2094

McFadden R-squared	0.000906	Mean dependent var	0.668700
S.D. dependent var	0.470727	S.E. of regression	0.470783
Akaike info criterion	1.272188	Sum squared resid	1124.808
Schwarz criterion	1.282472	Log likelihood	-3225.267
Hannan-Quinn criter.	1.275789	Deviance	6450.533
Restr. deviance	6456.383	Restr. log likelihood	-3228.192
LR statistic	5.849964	Avg. log likelihood	-0.634520
Prob(LR statistic)	0.557374		

Obs with Dep=0	1684	Total obs	5083
Obs with Dep=1	3399		



Series: QHAT
Sample 1 15000 IF SEX=1 AND GOODOBS=1
Observations 5083

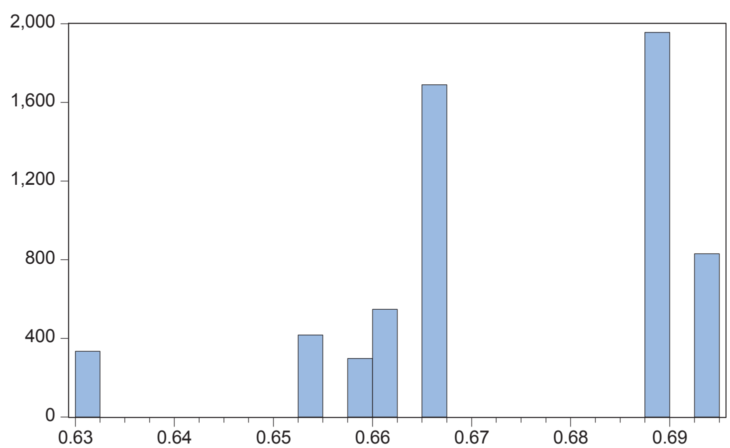
Mean	0.668700
Median	0.666174
Maximum	0.708235
Minimum	0.654064
Std. Dev.	0.015886
Skewness	1.169575
Kurtosis	3.617201
Jarque-Bera	1239.524
Probability	0.000000

FEMALES

Dependent Variable: Z (RANDOM OFFER OF SERVICES)
Method: ML - Binary Logit (Quadratic hill climbing)
Date: 12/15/12 Time: 16:12
Sample: 1 15000 IF SEX=0 AND GOODOBS=1
Included observations: 6067
Convergence achieved after 4 iterations
Covariance matrix computed using second derivatives

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	0.671288	0.090376	7.427722	0.0000
HS	0.019196	0.104066	0.184456	0.8537
MINORITY	-0.136365	0.145120	-0.939672	0.3474
BELOW30	-0.035300	0.137072	-0.257526	0.7968
HS*MINORITY	0.245892	0.170793	1.439702	0.1500
HS*BELOW30	0.140975	0.160331	0.879272	0.3793
MINORITY*BELOW30	0.163371	0.216064	0.756121	0.4496
HS01*MINORITY*BELOW30	-0.255949	0.254382	-1.006161	0.3143

McFadden R-squared	0.001067	Mean dependent var	0.673809
S.D. dependent var	0.468857	S.E. of regression	0.468810
Akaike info criterion	1.264187	Sum squared resid	1331.663
Schwarz criterion	1.273035	Log likelihood	-3826.910
Hannan-Quinn criter.	1.267258	Deviance	7653.820
Restr. deviance	7661.992	Restr. log likelihood	-3830.996
LR statistic	8.172498	Avg. log likelihood	-0.630775
Prob(LR statistic)	0.317629		
Obs with Dep=0	1979	Total obs	6067
Obs with Dep=1	4088		



Series: QHAT
Sample 1 15000 IF SEX=0 AND GOODOBS=1
Observations 6067

Mean	0.673809
Median	0.666075
Maximum	0.692771
Minimum	0.630631
Std. Dev.	0.017249
Skewness	-0.623027
Kurtosis	2.754264
Jarque-Bera	407.7626
Probability	0.000000

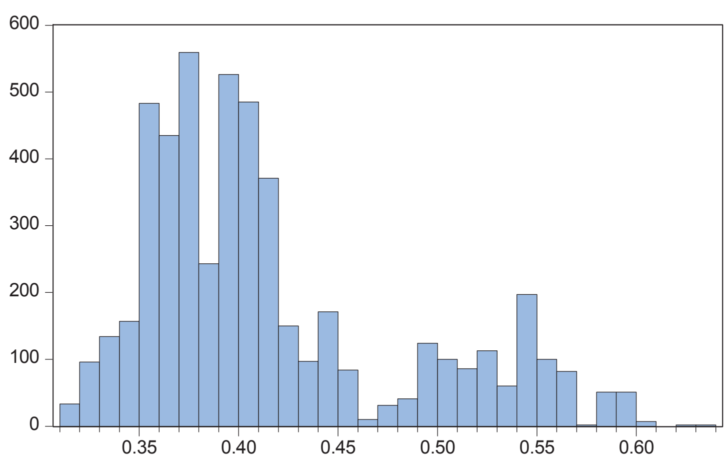
MALES

Dependent Variable: D (PARTICIPATION DUMMY)
Method: ML - Binary Logit (Quadratic hill climbing)
Date: 12/15/12 Time: 15:37
Sample: 1 15000 IF SEX=1 AND GOODOBS=1
Included observations: 5083
Convergence achieved after 4 iterations
Covariance matrix computed using second derivatives

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	-0.461464	0.180062	-2.562806	0.0104
A2225	-0.001881	0.169248	-0.011112	0.9911
A2629	-0.170449	0.171725	-0.992565	0.3209
A3035	-0.102333	0.169061	-0.605305	0.5450
A3644	-0.167305	0.170562	-0.980904	0.3266
A4554	-0.224944	0.187597	-1.199083	0.2305
BLACK	0.004379	0.070537	0.062084	0.9505
HISP	0.169274	0.098865	1.712173	0.0869
MARITAL STATUS	0.176134	0.061464	2.865633	0.0042
HIGH SCHOOL	0.105632	0.064905	1.627497	0.1036
WORKED LAST 12WK	-0.034243	0.059403	-0.576463	0.5643
CLASS_TR	0.548423	0.082963	6.610460	0.0000
OTJ_JSA	-0.052369	0.068131	-0.768650	0.4421
F2SMS	-0.000466	0.063170	-0.007379	0.9941
AFDC	0.014355	0.140523	0.102156	0.9186

McFadden R-squared	0.013606	Mean dependent var	0.416486
S.D. dependent var	0.493025	S.E. of regression	0.489106
Akaike info criterion	1.345687	Sum squared resid	1212.390
Schwarz criterion	1.364968	Log likelihood	-3405.063
Hannan-Quinn criter.	1.352439	Deviance	6810.125
Restr. deviance	6904.061	Restr. log likelihood	-3452.031
LR statistic	93.93592	Avg. log likelihood	-0.669892
Prob(LR statistic)	0.000000		

Obs with Dep=0	2966	Total obs	5083
Obs with Dep=1	2117		



Series: PHAT	
Sample 1 15000 IF SEX=1 AND GOODOBS=1	
Observations 5083	
Mean	0.416486
Median	0.398893
Maximum	0.630909
Minimum	0.315725
Std. Dev.	0.067228
Skewness	1.041525
Kurtosis	3.071458
Jarque-Bera	920.0658
Probability	0.000000

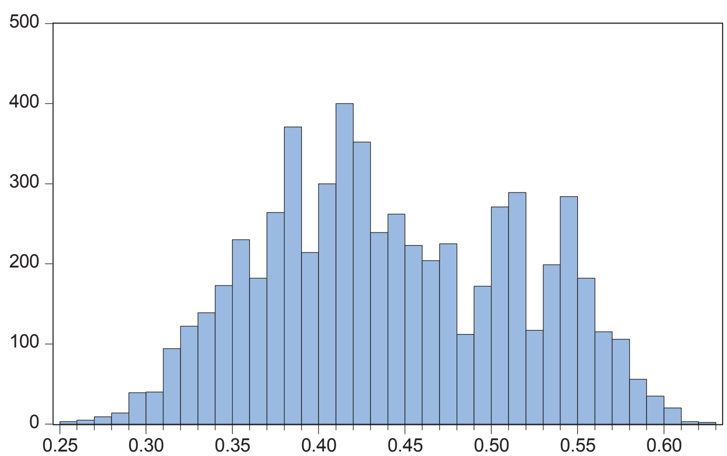
FEMALES

Dependent Variable: D (PARTICIPATION DUMMY)
Method: ML - Binary Logit (Quadratic hill climbing)
Date: 12/15/12 Time: 15:28
Sample: 1 15000 IF SEX=0 AND GOODOBS=1
Included observations: 6067
Convergence achieved after 4 iterations
Covariance matrix computed using second derivatives

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	-0.388253	0.146430	-2.651450	0.0080
A2225	-0.080431	0.138201	-0.581985	0.5606
A2629	-0.080184	0.138676	-0.578212	0.5631
A3035	-0.038998	0.136760	-0.285160	0.7755
A3644	-0.193794	0.138918	-1.395017	0.1630
A4554	-0.259554	0.154590	-1.678989	0.0932
BLACK	-0.112487	0.064172	-1.752905	0.0796
HISP	0.054035	0.083920	0.643894	0.5196
MARITAL STATUS	0.123547	0.059682	2.070098	0.0384
HIGH SCHOOL	0.264405	0.060822	4.347208	0.0000
WORKED LAST 12WK	-0.152918	0.055866	-2.737205	0.0062
CLASS_TR	0.381487	0.068025	5.608041	0.0000
OTJ_JSA	-0.149417	0.069171	-2.160118	0.0308
F2SMS	0.064533	0.059918	1.077023	0.2815
AFDC	0.101150	0.063362	1.596384	0.1104

McFadden R-squared	0.016155	Mean dependent var	0.442888
S.D. dependent var	0.496768	S.E. of regression	0.491802
Akaike info criterion	1.355979	Sum squared resid	1463.790
Schwarz criterion	1.372570	Log likelihood	-4098.363
Hannan-Quinn criter.	1.361738	Deviance	8196.726
Restr. deviance	8331.317	Restr. log likelihood	-4165.659
LR statistic	134.5914	Avg. log likelihood	-0.675517
Prob(LR statistic)	0.000000		

Obs with Dep=0	3380	Total obs	6067
Obs with Dep=1	2687		



Series: PHAT	
Sample 1 15000 IF SEX=0 AND GOODOBS=1	
Observations 6067	
Mean	0.442888
Median	0.435799
Maximum	0.621842
Minimum	0.256807
Std. Dev.	0.073734
Skewness	0.119635
Kurtosis	2.106597
Jarque-Bera	216.2428
Probability	0.000000

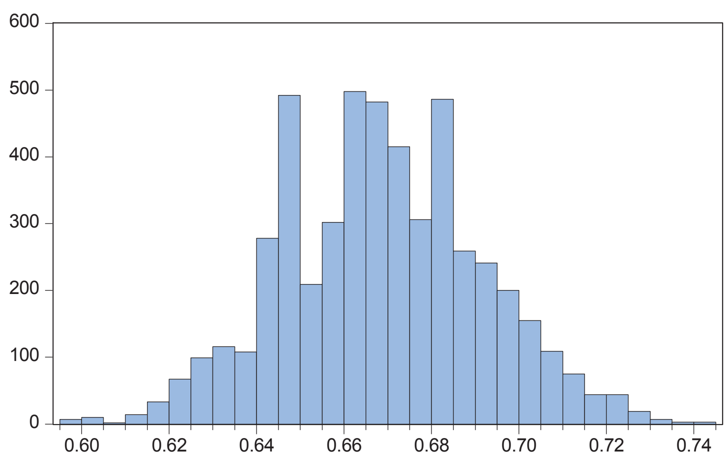
MALES

Dependent Variable: Z (RANDOM OFFER OF SERVICES)
Method: ML - Binary Logit (Quadratic hill climbing)
Date: 12/15/12 Time: 16:02
Sample: 1 15000 IF SEX=1 AND GOODOBS=1
Included observations: 5083
Convergence achieved after 4 iterations
Covariance matrix computed using second derivatives

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	0.585331	0.189618	3.086902	0.0020
A2225	-0.045532	0.179182	-0.254113	0.7994
A2629	0.024758	0.181664	0.136283	0.8916
A3035	-0.032771	0.178880	-0.183200	0.8546
A3644	-0.060331	0.180260	-0.334687	0.7379
A4554	-0.168260	0.196154	-0.857796	0.3910
BLACK	0.019934	0.073138	0.272548	0.7852
HISP	0.044805	0.104237	0.429835	0.6673
MARITAL STATUS	0.103766	0.063519	1.633623	0.1023
HIGH SCHOOL	-0.014380	0.067004	-0.214614	0.8301
WORKED LAST 12WK	0.056571	0.061430	0.920893	0.3571
CLASS_TR	0.177956	0.087620	2.030989	0.0423
OTJ_JSA	0.072630	0.069937	1.038509	0.2990
F2SMS	0.072452	0.065907	1.099311	0.2716

McFadden R-squared	0.001860	Mean dependent var	0.668700
S.D. dependent var	0.470727	S.E. of regression	0.470771
Akaike info criterion	1.273338	Sum squared resid	1123.417
Schwarz criterion	1.291334	Log likelihood	-3222.189
Hannan-Quinn criter.	1.279640	Deviance	6444.378
Restr. deviance	6456.383	Restr. log likelihood	-3228.192
LR statistic	12.00565	Avg. log likelihood	-0.633915
Prob(LR statistic)	0.527180		

Obs with Dep=0	1684	Total obs	5083
Obs with Dep=1	3399		



Series: ZHAT
Sample 1 15000 IF SEX=1 AND GOODOBS=1
Observations 5083

Mean	0.668700
Median	0.668497
Maximum	0.743766
Minimum	0.599334
Std. Dev.	0.022849
Skewness	0.065305
Kurtosis	2.820900
Jarque-Bera	10.40648
Probability	0.005499

FEMALES

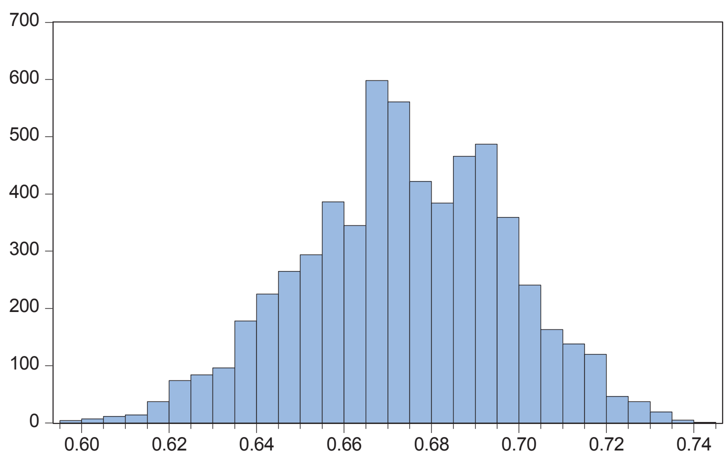
Dependent Variable: Z (RANDOM OFFER OF

SERVICES)
Method: ML - Binary Logit (Quadratic hill climbing)
Date: 12/15/12 Time: 16:00
Sample: 1 15000 IF SEX=0 AND GOODOBS=1
Included observations: 6067
Convergence achieved after 4 iterations
Covariance matrix computed using second derivatives

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	0.656508	0.153737	4.270332	0.0000
A2225	0.022293	0.145666	0.153043	0.8784
A2629	0.063588	0.146409	0.434320	0.6641
A3035	-0.066029	0.143969	-0.458630	0.6465
A3644	-0.011535	0.146148	-0.078924	0.9371
A4554	0.015763	0.162011	0.097295	0.9225
BLACK	0.096094	0.067612	1.421255	0.1552
HISP	-0.059614	0.087560	-0.680841	0.4960
MARITAL STATUS	0.088647	0.063016	1.406742	0.1595
HIGH SCHOOL	0.133949	0.062589	2.140132	0.0323
WORKED LAST 12WK	0.029925	0.058581	0.510836	0.6095
CLASS_TR	-0.093183	0.072291	-1.289010	0.1974
OTJ_JSA	-0.115807	0.072556	-1.596099	0.1105
F2SMS	0.022982	0.063140	0.363988	0.7159
AFDC	-0.054486	0.066384	-0.820772	0.4118

McFadden R-squared	0.001979	Mean dependent var	0.673809
S.D. dependent var	0.468857	S.E. of regression	0.468816
Akaike info criterion	1.265342	Sum squared resid	1330.157
Schwarz criterion	1.281933	Log likelihood	-3823.415
Hannan-Quinn criter.	1.271101	Deviance	7646.831
Restr. deviance	7661.992	Restr. log likelihood	-3830.996
LR statistic	15.16143	Avg. log likelihood	-0.630199
Prob(LR statistic)	0.367209		

Obs with Dep=0	1979	Total obs	6067
Obs with Dep=1	4088		



Series: ZHAT
Sample 1 15000 IF SEX=0 AND GOODOBS=1
Observations 6067

Mean	0.673809
Median	0.673874
Maximum	0.744352
Minimum	0.596402
Std. Dev.	0.023442
Skewness	-0.169383
Kurtosis	2.800856
Jarque-Bera	39.03613
Probability	0.000000