Problem Set 12 Sarah Bass

27.1 (27.2) $m(x) = E[Y|X] = X'\beta\,\Phi\left(\frac{X'\beta}{\sigma}\right) + \sigma\,\phi\left(\frac{X'\beta}{\sigma}\right)$

$$
\begin{aligned}
m(x) &= E[Y|X] \\
&= E[Y^* | \{X > 0\}] \\
&= E[X'\beta \mid \{X > 0\}] \\
&= \mu\left(1 - \Phi\left(\frac{0-\mu}{\sigma}\right)\right) + \sigma\,\phi\left(\frac{0-\mu}{\sigma}\right) \\
&= X'\beta\left(1 - \Phi\left(\frac{-X'\beta}{\sigma}\right)\right) + \sigma\,\phi\left(\frac{-X'\beta}{\sigma}\right) \\
&= X'\beta\,\Phi\left(\frac{X'\beta}{\sigma}\right) + \sigma\,\phi\left(\frac{X'\beta}{\sigma}\right)
\end{aligned}
$$

(27.3) $m^{\#}(x) = E[Y^{\#}|X] = X'\beta + \sigma\lambda\left(\frac{X'\beta}{\sigma}\right)$

$$
\begin{aligned}
E[Y^{\#}|X] &= E[Y^* | X > 0] \\
&= \mu + \sigma\lambda\left(\frac{\mu - 0}{\sigma}\right) \\
&= X'\beta + \sigma\lambda\left(\frac{X'\beta}{\sigma}\right)
\end{aligned}
$$

27.2 $Y^* = X'\beta + e$

$e \sim N(0, \sigma^2)$

$$
Y = \begin{cases} Y^* & \text{if } Y \le \tau \\ \text{missing} & \text{if } Y > \tau \end{cases}
$$

$$
\begin{aligned}
\hat{\beta} &= E[XX']^{-1} E[X'Y] \\
&= E[XX']^{-1} E[X'Y^* \mid Y^* \leq T] \cdot P(Y^* \leq T) \\
&= E[XX']^{-1} E[X'Y^*] \cdot P(Y^* \leq T) \\
&= \beta \cdot P(Y^* \leq T)
\end{aligned}
$$

This is not a consistent estimator for β.
This estimator $\hat{\beta}$ will be below the true value of β

27.4. $\min\limits_{\beta, \sigma} \left(Y - X'\beta\, \Phi\left(\frac{X'\beta}{\sigma}\right) - \sigma\, \phi\left(\frac{X'\beta}{\sigma}\right)\right)^2$

27.8 (27.7) $E[Y|X, Z, S=1] = X'\beta + \sigma_{21}\lambda(z'\gamma)$

$$
\begin{aligned}
E[Y|X,Z,S=1] &= E[Y^*|X,Z] \cdot Pr(S=1) \\
&= E[Y^*|X,Z] \cdot Pr(S^*>0) \\
&= E[Y^*|X] \cdot Pr(z'\gamma + u > 0) \\
&= E[X'\beta + e|X] \cdot Pr(z'\gamma + u > 0) \\
&= X'\beta \cdot Pr(u > -z'\gamma) \\
&= X'\beta + \sigma_{21}\lambda(z'\gamma)
\end{aligned}
$$

27.9

a.

```
. *a
. reg transfers income Dincome
```

Source	SS	df	MS
Model	101549.791	2	50774.8957
Residual	3316402.29	8,681	382.029984
Total	3417952.08	8,683	393.637231

Number of obs	=	8,684
F(2, 8681)	=	132.91
Prob > F	=	0.0000
R-squared	=	0.0297
Adj R-squared	=	0.0295
Root MSE	=	19.546

transfers	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
income	-42.76098	2.623992	-16.30	0.000	-47.90462	-37.61733
Dincome	42.76697	2.624527	16.30	0.000	37.62228	47.91167
_cons	49.72515	2.600365	19.12	0.000	44.62781	54.82248

b. 22.46% of observations have transfers=0, so censoring could be an issue for these.

c. . reg transfers income Dincome

Source	SS	df	MS
Model	106525.362	2	53262.6812
Residual	3168438.69	6,731	470.723323
Total	3274964.05	6,733	486.404879

Number of obs	=	6,734
F(2, 6731)	=	113.15
Prob > F	=	0.0000
R-squared	=	0.0325
Adj R-squared	=	0.0322
Root MSE	=	21.696

transfers	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
income	-42.81576	2.946075	-14.53	0.000	-48.591	-37.04052
Dincome	42.85625	2.946915	14.54	0.000	37.07937	48.63314
_cons	50.53024	2.912459	17.35	0.000	44.82089	56.23958

d. Tobit regression

Limits: lower = 0.00
upper = +inf

Log likelihood = -30272.069

Number of obs	=	6,734
Uncensored	=	6,733
Left-censored	=	1
Right-censored	=	0
LR chi2(2)	=	222.66
Prob > chi2	=	0.0000
Pseudo R2	=	0.0037

transfers	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
income	-42.81828	2.9457	-14.54	0.000	-48.59278	-37.04377
Dincome	42.85877	2.94654	14.55	0.000	37.08262	48.63492
_cons	50.53051	2.912087	17.35	0.000	44.82189	56.23912
var(e.transfers)	470.6032	8.110936			454.9688	486.7749

e.

```
. clad transfers income Dincome, ll

Initial sample size = 6734
Final sample size = 6734
Pseudo R2 = .02248594

Bootstrap statistics
```

Variable	Reps	Observed	Bias	Std. Err.	[95% Conf.	Interval]	
income	100	-46.11266	-.4893832	7.185244	-60.36974	-31.85557	(N)
					-58.74292	-34.88314	(P)
					-58.70982	-30.40741	(BC)
Dincome	100	46.11775	.4897029	7.185284	31.86059	60.37492	(N)
					34.88919	58.74852	(P)
					30.41023	58.71731	(BC)
const	100	48	.4817194	7.18257	33.74822	62.25178	(N)
					36.80779	60.581	(P)
					36.80779	60.581	(BC)

N = normal, P = percentile, BC = bias-corrected

The regressions in A, C, and D have very similar coefficients. In E, the constant is slightly smaller and the coefficients on income and Dincome are larger in magnitude. The standard errors in A, C, and D are much smaller than in E.

28.12

	AIC	BIC
Model 1	4766.0346	4814.0936
Model 2	4379.6882	4427.7471
Model 3	4390.1824	4468.2782
Model 4	4756.7389	4816.8126
Model 5	4369.7779	4429.8516
Model 6	4379.712	4469.8225
Model 7	4758.6806	4824.7616
Model 8	4371.7655	4437.8466
Model 9	4381.664	4477.7819

For each criterion, the first best model is highlighted in pink, and the second best is highlighted in blue. Since Model 5 is the best for AIC and the second best for BIC, this is what I would choose.