Web Appendix for "A Simple, Flexible Estimator for Count and Other Ordered Discrete Data" by Thomas A. Mroz, Journal of Applied Econometrics

Table of Contents:

Web Appendix Section 1

Unconditional distributions for the two data generating processes used in the Monte Carlo experiments.

For the Monte Carlo experiments with the Poisson DGP in Section 3.1, unconditionally and approximately, E(C)=1.18, V(C)=2.06, P(C=0) = .42, P(C=1)=.27, P(C=2) =.15, P(C=3)=.08, P(C=4)=.04, P(C=5)=.02, P(C=6)=.01, and P(C>6)<.01.

For the set of Monte Carlo experiments calibrated using information from Mullahy's(1998) doctor visit data in Section 3.2, unconditionally and approximately, E(C)=1.75; V(C)=3.90; P(C=0)=.149; P(C=1) = .454; P(C=2)=.208; P(C=3)=.087; P(C=4) = .042; P(C=5) = .022; P(C=6)=.013; P(C=7)=.008 ;P(C=8)=.005; P(C=9)=.004; P(C=10)=.003; P(C=11)=.002; P(C=12)=.001; P(C>12)<.004; P(C>49)=.00002.

Web Appendix Section 2

Extending the Estimation Approach To Right Censored and Left Truncated Data

Suppose that the right censoring for an observation occurs immediately after count Q. This situation is depicted in the 5th column of Table 1 in the paper. In this instance one knows that the count does not take on any value less than Q+1, but there is no information about which count value greater than Q the observation might have achieved. In this instance, one would use the censored observations' replications corresponding to all counts less than Q+1 when estimating the hazard functions; the binary outcome variable for each of these Q+1 potential counts (0,1,...,Q), $d_{A,bi}$, would be 0. Because of the independence of the conditional hazard events, there is no systematic censoring for all count outcomes greater than Q that would violate the ignorability condition[1]. Therefore, failing to incorporate information about counts above the censoring point in the point data does not introduce any systematic bias in the estimates. One would use the sequence of potential outcomes as depicted by the $d_{A,bi}$ in Table 1 column 5 in the published paper. Note that observations with observed counts less than Q+1 would be treated exactly the same as observations in the model with an absence of censoring.

Suppose one observes no information about any counts greater than Q in the sample, other than the fact that they take value Q+1 or larger. In this instance it would be impossible to infer exactly the contribution to expected values for potential outcomes greater than Q. Using a framework like that suggested by equation (7) in the paper, one could assume that the estimated

[1] The reason that the outcomes for 0 through Q must be included for the censored observations is because their draws from the standard uniform distribution, as introduced in the paper, are known to come from the upper tail of the uniform distributions associated with the hazard functions.

hazard functions for counts greater than Q follow the same functional form. This would be identification through functional form assumption; it is precisely the type of assumption that researchers make with standard parametric count models like the Poisson. Alternatively, one could make other arbitrary assumptions to address this shortcoming or use some other source of information about the larger outcomes. For example, one might assume that for all potential outcomes b greater than Q, that $\lambda(b,x) = \lambda(Q,x)$; this would imply a geometric right tail.

The outcome data for left truncation (and without right censoring) is depicted in the 6th column of Table 1. There is a crucial difference between a truncated data set and a censored data set. For a data set with censored outcomes, one observes the explanatory variables for all observations, but only knows that some outcomes fall within particular ranges. For truncated data sets, one does not observe any information about truncated observations, and often one does not even know the fraction of population that would be truncated. This requires that the empirical model explicitly condition on observations only being observed if they are not truncated.

Using standard conditional probability, the density function for a count c, given that it is at least equal to Q+1 and x is:

$$g(c \mid c > Q, x) = \frac{\lambda(c,x) \cdot \left\{ \prod_{b=0}^{c-1} [1 - \lambda(b,x)] \right\}}{S(Q)} = \frac{\lambda(c,x) \cdot \left\{ \prod_{b=0}^{c-1} [1 - \lambda(b,x)] \right\}}{\prod_{b=0}^{Q} [1 - \lambda(b,x)]}$$

$$= \lambda(c,x) \cdot \left\{ \prod_{b=Q+1}^{c-1} [1 - \lambda(b,x)] \right\}$$

Note that this conditional density function only depends on hazard rates for counts at least equal to Q+1. By the independence property for the events associated with the hazard rates, all observed hazard "events" for count greater than Q are from random samples from Bernoulli distributions with the probability given by the corresponding hazard rate.

Estimation with left truncated data is straightforward. One would model hazard events only for potential counts greater than Q and otherwise proceed exactly as above. The 6th column in Table 1 indicates which observations would be used in the binary outcome models. It is straightforward for the truncation point to vary across observations. Column 7 presents the sample sequence of outcomes one would model if there were left truncation and right censoring below an observation's observed count.

If all observations in a data set are truncated below some threshold, then it will be impossible to make data-driven inferences about unconditional expectations from this sample without imposing untestable assumptions or out-of-sample information. As with right censored data, one could assume that the process estimated over the non-truncated observations applies for the entire possible range of the count outcomes. This would again be identification through a functional form assumption.

Web Appendix Section 3

Model Selection Frequencies for the "Hazard" Models in the Monte Carlo Experiments

Web Appendix Table 1

Percent of times a polynomial model was selected by sequential upwards testing with likelihood ratio tests with Probability(Type I error), and by cross validation, BIC, HQIC, and AIC

A) True Model is Poisson: For Table 2 in the paper, and Web Appendix Tables 2A, 2B, 2C, and 2D

	Degree 1	Degree 2	Degree 3	Degree 4
Prob "Type I Error"				
Upwards Testing at "size"				
.05	0.0	70.9	25.3	3.8
.10	0.0	58.0	32.9	9.1
.25	0.0	36.9	38.1	25.0
.50	0.0	16.7	28.1	55.2
BIC	95.6	4.4	0.0	0.0
HQIC	6.1	93.9	1.6	0.0
AIC	0.0	97.0	3.0	0.0
Cross Validation	0.1	97.3	2.6	0.0

B) True Model is Neural Net based on Mullahy (1998) Data: for Table 3 in the paper and Web Appendix Tables 3A, 3B, 3C, and 3D

	Degree 1	Degree 2	Degree 3	Degree 4
Prob "Type I Error"				
Upwards Testing at "size"				
.05	0.5	9.7	13.6	76.2
.10	0.5	9.1	12.8	77.6
.25	0.4	8.7	11.8	79.1
.50	0.4	7.8	10.3	81.5
BIC	63.9	36.0	0.1	0.0
HQIC	1.5	88.2	10.3	0.0
AIC	0.0	10.7	28.5	60.8
Cross Validation	4.1	91.9	4.0	0.0

Web Appendix Section 4

Additional estimated effects in the Monte Carlo Studies and their empirical mean square errors and mean absolute deviations

Web Appendix Table 2A
Monte Carlo Estimates when the True Model is Poisson: The Impacts of Three Explanatory Variables on the Expected Count from Different Estimation and Model Selection Procedures

	Average X1 Derivative (sd)	Average X2 Derivative (sd)	Average X3 Derivative (sd)	Average Dummy Var Effect (sd)
True Model	1.211 (0.045)	1.212 (0.046)	1.211 (0.045)	0.801 (0.011)
Poisson	1.213 (0.081)	1.216 (0.079)	1.211 (0.078)	0.800 (0.037)
Negative Binomial	1.212 (0.081)	1.216 (0.079)	1.211 (0.078)	0.800 (0.037)
Logit Condit. Hazard, Polynomial Degree 1	1.040 (0.068)	1.042 (0.065)	1.038 (0.066)	0.705 (0.036)
Logit Condit. Hazard, Polyn. Degree 2	1.157 (0.083)	1.161 (0.082)	1.157 (0.817)	0.765 (0.050)
Logit Condit. Hazard, Polyn. Degree 3	1.221 (0.117)	1.216 (0.119)	1.209 (0.115)	0.789 (0.056)
Logit Condit. Hazard, Polyn. Degree 4	1.236 (0.139)	1.233 (0.136)	1.228 (0.136)	0.779 (0.093)
Logit Condit. Hazard, LRT Selection using p=.05*	1.179 (0.102)	1.183 (0.104)	1.177 (0.102)	0.773 (0.055)
Logit Condit. Hazard, LRT Selection using p=.10*	1.188 (0.111)	1.192 (0.112)	1.183 (0.108)	0.776 (0.058)
Logit Condit. Hazard, LRT Selection using p=.25*	1.205 (0.122)	1.204 (0.122)	1.200 (0.120)	0.782 (0.060)
Logit Condit. Hazard, LRT Selection using p=.50*	1.221 (0.133)	1.219 (0.130)	1.215 (0.127)	0.784 (0.071)
BIC -to select logit model*	1.047 (0.074)	1.048 (0.071)	1.044 (0.073)	0.709 (0.040)
HQIC-to select logit model*	1.152 (0.087)	1.155 (0.0854)	1.151 (0.086)	0.762 (0.051)
AIC -to select logit model*	1.160 (0.085)	1.163 (0.085)	1.160 (0.085)	0.766 (0.050)
Cross Validation*	1.158 (0.084)	1.163 (0.083)	1.159 (0.084)	0.765 (0.050)

Web Appendix Table 2B

Monte Carlo Estimates of Effects Conditional on the Sign of x_1 when the True Model is Poisson: The Impacts of Three Explanatory Variables on the Expected Count from Different Estimation and Model Selection Procedures

	Average Derivative For X1 (sd)		Average Derivative For X2 (sd)		Average Effect for Dummy Variable(sd)	
	x1<0	x1>0	x1<0	x1>0	x1<0	x1>0
True Model	0.629 (0.047)	1.794 (0.076)	0.624 (0.048)	1.802 (0.078)	0.447 (0.011)	1.156 (0.018)
Poisson	0.629 (0.033)	1.797 (0.129)	0.627 (0.044)	1.806 (0.116)	0.447 (0.024)	1.155 (0.053)
Negative Binomial	0.629 (0.033)	1.797 (0.129)	0.628 (0.045)	1.806 (0.116)	0.446 (0.024)	1.154 (0.053)
Logit Condit. Hazard, Polyn. Degree 1	0.773 (0.043)	1.308 (0.093)	0.682 (0.045)	1.403 (0.088)	0.438 (0.025)	0.974 (0.052)
Logit Condit. Hazard, Polyn. Degree 2	0.627 (0.072)	1.689 (0.176)	0.621 (0.063)	1.702 (0.136)	0.447 (0.028)	1.082 (0.087)
Logit Condit. Hazard, Polyn. Degree 3	0.631 (0.072)	1.811 (0.239)	0.638 (0.078)	1.796 (0.194)	0.447 (0.030)	1.131 (0.104)
Logit Condit. Hazard, Polyn. Degree 4	0.630 (0.075)	1.844 (0.280)	0.635 (0.097)	1.833 (0.234)	0.444 (0.032)	1.115 (0.181)
Logit Condit. Hazard, LRT Selection using p = 0.05	0.628 (0.072)	1.7432 (0.213)	0.627 (0.070)	1.742 (0.172)	0.447 (0.029)	1.099 (0.101)
Logit Condit. Hazard, LRT Selection using p = 0.10	0.629 (0.072)	1.748 (0.229)	0.630 (0.075)	1.756 (0.182)	0.447 (0.029)	1.106 (0.106)
Logit Condit. Hazard, LRT Selection using p = 0.25	0.630 (0.073)	1.782 (0.248)	0.632 (0.080)	1.777 (0.204)	0.446 (0.030)	1.118 (0.112)
Logit Condit. Hazard, LRT Selection using p = 0.50	0.630 (0.074)	1.814 (0.268)	0.634 (0.087)	1.806 (0.220)	0.445 (0.031)	1.124 (0.135)
BIC -to select logit model	0.765 (0.060)	1.329 (0.138)	0.678 (0.050)	1.419 (0.114)	0.438 (0.025)	0.980 (0.060)
HQIC -to select logit model	0.633 (0.079)	1.671 (0.195)	0.624 (0.064)	1.687 (0.149)	0.447 (0.028)	1.077 (0.089)
AIC -to select logit model	0.627 (0.072)	1.693 (0.179)	0.621 (0.064)	1.705 (0.140)	0.448 (0.028)	1.084 (0.088)
Cross Validation	0.627 (0.072)	1.690 (0.178)	0.622 (0.064)	1.705 (0.138)	0.447 (0.028)	1.084 (0.088)

Web Appendix Table 2C

Root Mean Square Errors and Mean Absolute Errors for Estimates in Web Appendix Table 2A
True Model is Poisson
(Root Mean Square Error) [Mean Absolute Deviation]

	Average X1 Derivative (sd)	Average X2 Derivative (sd)	Average X3 Derivative (sd)	Average Dummy Var Effect (sd)
True Model	(0.000),[0.000]	(0.000),[0.000]	(0.000),[0.000]	(0.000),[0.000]
Poisson	(0.087),[0.069]	(0.087),[0.069]	(0.085),[0.067]	(0.034),[0.028]
Logit Cond. Hazard, Polynomial Degree 1	(0.187),[0.171]	(0.187),[0.171]	(0.189),[0.174]	(0.102),[0.096]
Logit Cond. Hazard, Polynomial Degree 2	(0.105),[0.084]	(0.104),[0.083]	(0.104),[0.082]	(0.060),[0.048]
Logit Condit. Hazard, Polynomial 1 Degree 3	(0.121),[0.096]	(0.122),[0.099]	(0.119),[0.094]	(0.056),[0.043]
Logit Condit. Hazard, Polynomial Degree 4	(0.144),[0.114]	(0.139),[0.112]	(0.140),[0.112]	(0.095),[0.059]
Logit Condit. Hazard, LRT Selection using p=.05*	(0.112),[0.089]	(0.114),[0.091]	(0.111),[0.088]	(0.061),[0.047]
Logit Condit. Hazard, LRT Selection using p=.10*	(0.119),[0.094]	(0.118),[0.094]	[0.115),[0.090]	(0.062),[0.047]
Logit Condit. Hazard, LRT Selection using p=.25*	(0.126),[0.097]	(0.126),[0.101]	(0.123),[0.959]	(0.062),[0.047]
Logit Condit. Hazard, LRT Selection using p=.50*	(0.136),[0.106]	(0.132),[0.106]	(0.130),[0.102]	(0.071),[0.050]
BIC - for logit models*	(0.184),[0.166]	(0.183),[0.166]	(0.186),[0.169]	(0.100),[0.093]
HQIC - for logit models*	(0.110),[0.088]	(0.109),[0.087]	(0.110),[0.087]	(0.063),[0.071]
AIC - for logit models*	(0.105),[0.083]	(0.105),[0.084]	(0.104),[0.082]	(0.060),[0.048]
Cross Validation*	(0.105),[0.084]	(0.104),[0.083]	(0.104),[0.082]	(0.060),[0.048]

Web Appendix Table 2D

Root Mean Square Errors and Mean Absolute Errors for Web Appendix Table 2B Estimates:
True Model is Poisson, For X1>0 and X1<0
(Root Mean Square Error) [Mean Absolute Deviation]

	Average Derivative for X1 (sd) x1<0	x1>0	Average Derivative for X2 (sd) x1<0	x1>0	Average Dummy Variable effect(sd) x1<0	x1>0
True Model	(0.000) [0.000]	(0.000) [0.000]	(0.000) [0.000]	(0.000) [0.000]	(0.000) [0.000]	(0.000) [0.000]
Poisson	(0.056) [0.045]	(0.141) [0.112]	(0.064) [0.051]	(0.132) [0.105]	(0.022) [0.018]	(0.048) [0.039]
Logit Cond. Hazard, Polynomial Degree 1	(0.157) [0.145]	(0.499) [0.489]	(0.087) [0.071]	(0.414) [0.399]	(0.025) [0.020]	(0.189) [0.182]
Logit Cond. Hazard, Polyn. Degree 2	(0.084) [0.066]	(0.212) [0.170]	(0.077) [0.062]	(0.181) [0.146]	(0.026) [0.021]	(0.113) [0.089]
Logit Condit. Hazard, Polyn. Degree 3	(0.084) [0.066]	(0.246) [0.193]	(0.089) [0.071]	(0.201) [0.162]	(0.028) [0.022]	(0.106) [0.082]
Logit Condit. Hazard, Polyn. Degree 4	(0.086) [0.068]	(0.288) [0.227]	(0.105) [0.084]	(0.238) [0.189]	(0.031) [0.024]	(0.185) [0.115]
Logit Condit. Hazard, LRT Selection using p = 0.05*	(0.083) [0.066]	(0.228) [0.181]	(0.083) [0.066]	(0.193) [0.156]	(0.027) [0.022]	(0.114) [0.089]
Logit Condit. Hazard, LRT Selection using p = 0.10*	(0.084) [0.066]	(0.240) [0.188]	(0.087) [0.069]	(0.197) [0.159]	(0.027) [0.022]	(0.116) [0.088]
Logit Condit. Hazard, LRT Selection using p = 0.25*	(0.084) [0.067]	(0.255) [0.195]	(0.090) [0.071]	(0.212) [0.169]	(0.028) [0.022]	(0.116) [0.089]
Logit Condit. Hazard, LRT Selection using p = 0.50*	(0.085) [0.067]	(0.274) [0.211]	(0.097) [0.077]	(0.223) [0.177]	(0.029) [0.023]	(0.138) [0.097]
BIC - for logit models	(0.155) [0.141]	(0.488) [0.469]	(0.087) [0.071]	(0.405) [0.386]	(0.025) [0.020]	(0.185) [0.177]
HQIC - for logit models	(0.084) [0.071]	(0.236) [0.186]	(0.078) [0.063)	(0.199) [0.159]	(0.026) [0.021]	(0.112) [0.094]
AIC - for logit models	(0.084) [0.066]	(0.211) [0.170]	(0.078) [0.062]	(0.182) [0.147]	(0.026) [0.021]	(0.118) [0.089]
Cross Validation*	(0.084) [0.066]	(0.212) [0.171]	(0.078) [0.062]	(0.181) [0.146]	(0.026) [0.021]	(0.112) [0.089]

Web Appendix Table 3A

Monte Carlo Estimates When the True Model is Derived from Mullahy's (1999) Data: Average Effects for Several Covariates from Different Estimation and Model Selection Procedures

	Average Age Derivative (sd)	Average Male Effect (sd)	Average Educ. Derivative (sd)	Average Health Status Effect (sd)
True Model	-0.0056 (0.0008)	-0.554 (0.010)	0.090 (0.004)	-0.324 (0.011)
Poisson	-0.0021 (0.0025)	-0.527 (0.052)	0.067 (0.010)	-0.522 (0.028)
Negative Binomial	-0.0017 (0.0023)	-0.533 (0.049)	0.063 (0.009)	-0.488 (0.021)
Logit Condit. Hazard, Polyn. Degree 1	-0.0028 (0.0022)	-0.501 (0.055)	0.036 (0.007)	-0.397 (0.029)
Logit Condit. Hazard, Polyn. Degree 2	-0.0030 (0.0029)	-0.512 (0.054)	0.070 (0.010)	-0.356 (0.030)
Logit Condit. Hazard, Polyn. Degree 3	-0.0045 (0.0029)	-0.519 (0.052)	0.076 (0.01)	-0.319 (0.030)
Logit Condit. Hazard, Polyn. Degree 4	-0.0051 (0.0039)	-0.519 (0.054)	0.079 (0.012)	-0.306 (0.034)
Logit Condit. Hazard, LRT Selection using p=.05*	-0.0048 (0.0037)	-0.519 (0.055)	0.077 (0.013)	-0.314 (0.037)
Logit Condit. Hazard, LRT Selection using p=.10*	-0.0048 (0.0037)	-0.519 (0.055)	0.078 (0.013)	-0.313 (0.037)
Logit Condit. Hazard, LRT Selection using p=.25*	-0.0048 (0.0038)	-0.519 (0.055)	0.078 (0.013)	-0.313 (0.037)
Logit Condit. Hazard, LRT Selection using p=.50*	-0.0048 (0.0038)	-0.519 (0.055)	0.078 (0.013)	-0.313 (0.036)
BIC -to select logit model*	-0.0028 (0.0025)	-0.504 (0.055)	0.049 (0.019)	-0.380 (0.037)
HQIC-to select logit model*	-0.0031 (0.0029)	-0.514 (0.054)	0.070 (0.011)	-0.353 (0.033)
AIC -to select logit model*	-0.0047 (0.0036)	-0.520 (0.054)	0.0771 (0.0122)	-0.316 (0.037)
Cross Validation*	-0.0030 (0.029)	-0.512 (0.054)	0.069 (0.012)	-0.357 (0.032)

Web Appendix Table 3 B

Monte Carlo Estimates by Age When the True Model is Derived from Mullahy's (1999) Data: Average Effects from Different Estimation an Model Selection Procedures

	Average Derivative For Age (sd)		Average Education Effect (sd)		Average Health Status Effect (sd)	
	Age<50	Age≥50	Age<50	Age≥50	Age<50	Age≥50
True Model	-0.0084 (0.0010)	0.0026 (0.0007)	0.0923 (0.0051)	0.0830 (0.0076)	-0.2834 (0.0117)	-0.4399 (0.0219)
Poisson	-0.0021 (0.0025)	-0.0022 (0.0027)	0.0648 (0.0100)	0.0716 (0.0110)	-0.5136 (0.0280)	-0.5465 (0.0320)
Negative Binomial	-0.0017 (0.0023)	-0.0018 (0.0024)	0.0617 (0.0091)	0.0678 (0.0101)	-0.4805 (0.0208)	-0.5094 (0.0238)
Logit Condit. Hazard, Polyn. Degree 1	-0.0030 (0.0022)	-0.0022 (0.0017)	0.0340 (0.0068)	0.0422 (0.0091)	-0.3754 (0.0251)	-0.4571 (0.0444)
Logit Condit. Hazard, Polyn. Degree 2	-0.0083 (0.0034)	0.0123 (0.0059)	0.0710 (0.0106)	0.0671 (0.0153)	-0.3194 (0.0303)	-0.4618 (0.0522)
Logit Condit. Hazard, Polyn. Degree 3	-0.0090 (0.0031)	0.0084 (0.0092)	0.0792 (0.0118)	0.0683 (0.0183)	-0.2785 (0.0329)	-0.4369 (0.0483)
Logit Condit. Hazard, Polyn. Degree 4	-0.0087 (0.0033)	0.0051 (0.0133)	0.0821 (0.0135)	0.0702 (0.0247)	-0.2660 (0.0337)	-0.4221 (0.0562)
Logit Condit. Hazard, LRT Selection using p=.05	-0.0087 (0.0033)	0.0065 (0.0125)	0.0804 (0.0138)	0.0690 (0.0233)	-0.2737 (0.0405)	-0.4290 (0.0566)
Logit Condit. Hazard, LRT Selection using p=.10	-0.0087 (0.0033)	0.0064 (0.0125)	0.0805 (0.0138)	0.0690 (0.0234)	-0.2732 (0.0404)	-0.4286 (0.0563)
Logit Condit. Hazard, LRT Selection using p=.25	-0.0087 (0.0033)	0.0064 (0.0126)	0.0806 (0.0138)	0.0691 (0.0234)	-0.2728 (0.0401)	-0.4284 (0.0562)
Logit Condit. Hazard, LRT Selection using p=.50	-0.0088 (0.0033)	0.0064 (0.0127)	0.0807 (0.0138)	0.0692 (0.0235)	-0.2721 (0.0400)	-0.4275 (0.0561)
BIC -to select logit model	-0.0050 (0.0038)	0.0034 (0.0083)	0.0481 (0.0206)	0.0519 (0.0177)	-0.3530 (0.0407)	-0.4572 (0.0465)
HQIC -to select logit model	-0.0083 (0.0033)	0.0112 (0.0065)	0.0712 (0.0118)	0.0671 (0.0158)	-0.3156 (0.0347)	-0.4603 (0.0520)
AIC -to select logit model	-0.0088 (0.0033)	0.0069 (0.0118)	0.0800 (0.0133)	0.0689 (0.0225)	-0.2760 (0.0397)	-0.4315 (0.0564)
Cross Validation	-0.0081 (0.0035)	0.0116 (0.0067)	0.0698 (0.0129)	0.0662 (0.0162)	-0.3203 (0.0326)	-0.4615 (0.0514)

Web Appendix Table 3C
Root Mean Square Errors and Mean Absolute Errors for Estimates in Web Appendix Table 3A
True Model is Neural Net Estimated with Mullahy (1998) Data
(Root Mean Square Error) [Mean Absolute Deviation]

	Average Age Derivative (sd)	Average Male Effect (sd)	Average Educ. Derivative (sd)	Average Health Status Effect(sd)
True Model	(0.000) [0.000]	(0.000) [0.000]	(0.000) [0.000]	(0.000) [0.000]
Poisson	(0.004) [0.003]	(0.058) [0.047]	(0.026) [0.024]	(0.199) [0.198]
Negative Binomial	(0.005) [0.004]	(0.053) [0.043]	(0.029) [0.027]	(0.165) [0.164]
Logit Cond. Hazard, Poly. Degree 1	(0.004) [0.003]	(0.076) [0.064]	(0.054) [0.054]	(0.078) [0.073]
Logit Cond. Hazard, Poly. Degree 2	(0.004) [0.003]	(0.067) [0.055]	(0.023) [0.020]	(0.042) [0.036]
Logit Condit. Hazard, Poly. 1 Degree 3	(0.003) [0.002]	(0.062) [0.050]	(0.018) [0.015]	(0.027) [0.022]
Logit Condit. Hazard, Poly. Degree 4	(0.004) [0.004]	(0.064) [0.051]	(0.017) [0.015]	(0.036) [0.028]
Logit Condit. Hazard, LRT Selection using p=.05*	(0.004) [0.003]	(0.064) [0.051]	(0.018) [0.015]	(0.037) [0.028]
Logit Condit. Hazard, LRT Selection using p=.10*	(0.004) [0.003]	(0.064) [0.051]	(0.018) [0.015]	(0.037) [0.028]
Logit Condit. Hazard, LRT Selection using p=.25*	(0.004) [0.003]	(0.064) [0.051]	(0.018) [0.015]	(0.036) [0.028]
Logit Condit. Hazard, LRT Selection using p=.50*	(0.004) [0.003]	(0.064) [0.051]	(0.018) [0.015]	(0.037) [0.028]
BIC - for logit models*	(0.004) [0.003]	(0.074) [0.061]	(0.045) [0.041]	(0.037) [0.058]
HQIC - for logit models*	(0.004) [0.003]	(0.067) [0.054]	(0.023) [0.020]	(0.042) [0.035]
AIC - for logit models*	(0.004) [0.003]	(0.064) [0.050]	(0.018) [0.015]	(0.035) [0.027]
Cross Validation*	(0.004) [0.003]	(0.068) [0.056]	(0.025) [0.022]	(0.044) [0.037]

Web Appendix Table 3D
Root Mean Square Errors and Mean Absolute Errors for Estimatesby Age in Web Appendix Table 3B:True Model is Neural Net Estimated with Mullahy (1998) Data
(Root Mean Square Error) [Mean Absolute Deviation]

	Average Age Effect Age<50	Age≥50	Average Education Effect Age<50	Age≥50	Average Health Status Effect Age<50	Age≥50
True Model	(0.0000) [0.0000]	(0.0000) [0.0000]	(0.0000) [0.0000]	(0.0000) [0.0000]	(0.0000) [0.0000]	(0.0000) [0.0000]
Poisson	(0.0068) [0.0063]	(0.0056) [0.0049]	(0.0297) [0.0277]	(0.0172) [0.0141]	(0.2313) [0.2302]	(0.1100) [0.1066]
Negative Binomial	(0.0071) [0.0067]	(0.0051) [0.0045]	(0.0324) [0.0307]	(0.0195) [0.0165]	(0.1977) [0.1970]	(0.0727) [0.0696]
Logit Cond. Hazard, Poly. Degree 1	(0.0059) [0.0055]	(0.0051) [0.0048]	(0.0589) [0.0584]	(0.0425) [0.0408]	(0.0948) [0.0920]	(0.0468) [0.0382]
Logit Cond. Hazard, Poly. Degree 2	(0.0034) [0.0027]	(0.0113) [0.0098]	(0.0244) [0.0218]	(0.0230) [0.0193]	(0.0447) [0.0384]	(0.0520) [0.0407]
Logit Condit. Hazard, Poly. Degree 3	(0.0031) [0.0024]	(0.0109) [0.0087]	(0.0183) [0.0153]	(0.0242) [0.0200]	(0.0303) [0.0241]	(0.0428) [0.0336]
Logit Condit. Hazard, Poly. Degree 4	(0.0033) [0.0026]	(0.0136) [0.0107]	(0.0177) [0.0143]	(0.0282) [0.0218]	(0.0394) [0.0308]	(0.0547) [0.0421]
Logit Condit. Hazard, LRT Selection using p=.05	(0.0033) [0.0026]	(0.0131) [0.0104]	(0.0189) [0.0154]	(0.0278) [0.0215]	(0.0399) [0.0311]	(0.0529) [0.0407]
Logit Condit. Hazard, LRT Selection using p=.10	(0.0033) [0.0025]	(0.0131) [0.0104]	(0.0188) [0.0153)	(0.0279) [0.0216]	(0.0400) [0.0312]	(0.0528) [0.0405]
Logit Condit. Hazard, LRT Selection using p=.25	(0.0033) [0.0025]	(0.0132) [0.0105]	(0.0187) [0.0152]	(0.0278) [0.0215]	(0.0397) [0.0309]	(0.0527) [0.0405]
Logit Condit. Hazard, LRT Selection using p=.50	(0.0033) [0.0025]	(0.0133) [0.0106]	(0.0186) [0.0151]	(0.0278) [0.0214]	(0.0398) [0.0310]	(0.0529) [0.0406]
BIC - for logit models	(0.0051) [0.0045]	(0.0084) [0.0070]	(0.0490) [0.0445]	(0.0366) [0.0328]	(0.0793) [0.0709]	(0.0478) [0.0384]
HQIC - for logit models	(0.0033) [0.0026]	(0.0113) [0.0097]	(0.0246) [0.0217]	(0.0233) (0.0208)	(0.0387) [0.0384]	(0.0512) [0.0401]
AIC - for logit models	(0.0033) [0.0026]	(0.0126) [0.0101]	(0.0187) [0.0153]	(0.0273) [0.0195]	(0.0453) [0.0299]	(0.0521) [0.0397]
Cross Validation	(0.0035) [0.0027]	(0.0112) [0.0096]	(0.0265) [0.0230]	(0.0242) [0.0202]	(0.0476) [0.0403]	(0.0516) [0.0406]