APPENDIX A: CODEBOOK OF VARIABLES USED IN TRIAD ANALYSES

Variables Used in Outcome Analyses

NOTE: Unless otherwise indicated, all of the nominal variables are coded as effects vectors.

Age at Time of First Commitment
AGEFIRCO: A continuous variable.

Age at Time of Release from Current Incarceration
AGERLSE: A continuous variable.

Alcohol Treatment History
EPSTETOH: Coded 1 if there was no previous inpatient or outpatient alcohol treatment.
Excluded group had previous history of alcohol treatment.

Diagnoses of Depression and Antisocial Personality
EDIAGDEP: Coded 1 if diagnosis of depression only.
EDIAGASP: Coded 1 if diagnosis of antisocial personality only.
EDIAGBTH: Coded 1 if both diagnoses.
Excluded group has neither diagnosis.

Disciplinary Infraction Before Release
A) ERELIRY: Coded 1 if individual had one or more 100 or 200 level (lowest level is 400) disciplinary infractions within 6 months before release from BOP custody.
Excluded group had no 100 or 200 disciplinary infraction 6 months before release.

B) EDRUGIRY: Coded 1 if individual had one or more drug related disciplinary infractions within 6 months before release from BOP custody.
Excluded group had no drug related disciplinary infraction 6 months before release.

Drug Treatment History
EPSTDGTX: Coded 1 if no previous inpatient or outpatient drug treatment.
Excluded group had previous history of drug treatment.

Drug Use in the Year Before Arrest. (These are dummy-coded variables with no daily alcohol or illicit drug use during year before arrest serving as excluded group and coded as 0).
ALCONLY: Coded 1 if used alcohol only on a daily basis.
MJALC: Coded 1 if used alcohol and marijuana only on a daily basis.
MJNOALC: Coded 1 if used marijuana only on a daily basis.

ONEALCY: Coded 1 if used alcohol and only one illicit drug other than marijuana on a daily basis.
ONEALCN: Coded 1 if did not use alcohol but used only one illicit drug other than marijuana on a daily basis.
TWOALCY: Coded 1 if used alcohol and two or more illicit drugs other than marijuana on a daily basis.
TWOALCN: Coded 1 if did not use alcohol but used only two or more illicit drugs other than marijuana on a daily basis.

Drug Use - Lifetime (These is a dummy-coded variable with the excluded group comprised of those who never used an illicit drug on a daily basis and coded as 0).
DAILYNO: Coded 1 if ever used an illicit drug on a daily basis.

Education
GRADEA: Number of years of education: A continuous variable.

Employment Status in Month Before Incarceration
EWORKJOB: Coded 1 if working full- or part-time.
ELEGITUN: Coded 1 if unemployed because in school, a homemaker, retired, or disabled.
EUNEMP: Coded 1 if unemployed but looking for work.
Excluded group is composed of those unemployed because of involvement in illegal drug use or illegal activities, because the individual has never been employed, or due to other reasons.

Employment After Release (This is a dummy coded variable with those who were in the work force after release coded as 0)
WORKNA: Not in the labor force after release

Ethnicity
EHISP: Coded 1 for Hispanics.
Excluded group consists of non-Hispanics.

Illegal Source of Income
ESUPILL: Coded 1 if ever supported self mainly through illegal activity for at least one year.
Excluded group did not ever support self mainly through illegal activity for at least one year.

Involved in Post-Release Self-Help Group[1]
EAAYES: Coded 1 if involved in self-help group within first month after release to supervision.
Excluded group did not participate in self-help after release from prison.

Level of selection bias
COVARIAN—Ordinal variable.

Living with Spouse
ESPOUSE: Coded 1 if living with spouse upon release to supervision.
ECOM_LAW: Coded 1 if living with common-law partner.
Excluded group not living with spouse or common-law partner.

Mental Health Treatment History
EPSTMHTX: Coded 1 if subject received no previous inpatient or outpatient mental health treatment.
Excluded group had previous history of mental health treatment

Monthly Rate of Urine Testing[1]
UA_RATE: Average number of urinalysis tests per month during first 6 months of release.

Monthly Rate of Personal Contacts with Probation Officer[1]
PC_RATE: Average number of personal contacts – face-to-face visits at probation office, home or work or telephone contacts – per month during first 6 months of release.

Monthly Rate of Collateral Contacts by Probation Officer[1]
CC_RATE: Average number of collateral contacts – at home or office or by telephone – per month during first 6 months of release.

Prior Commitments
EPRIORCM: Coded 1 if had a major or minor prior commitment.
Excluded group had no prior commitment.

Prison Industries - Participation during incarceration[1]
PERCEN_U: Percent of time incarcerated employed in prison industries (UNICOR).

Race
EBLACK: Coded as 1 if black
ERACEOTH: Coded as 1 if of other race.
Excluded group are those of white race.

Received CCC Placement[1]
ENOCCC: Coded 1 if did not receive CCC placement.
Excluded group received a CCC placement.
Sex
EFEM: Coded 1 if female, with males serving as the excluded group.

Sentence Reduction
EYEARNO: Coded 1 if did not have treatment available when year off sentence reduction available.
EYEARYES: Coded 1 if had treatment available when year off sentence reduction available.

Excluded group comprised of those for whom the year off sentence reduction was not applicable. Non-DAP comparison subjects did not have treatment available.

Split Population
Q:

Spouse With Drug Problem
ESPOUDRG : Coded 1 if spouse ever had a drug problem (before incarceration of subject).
Excluded group did not have a spouse with a drug problem.

Supervised After Release[1]
ESUPRLNO: Coded 1 if not released to supervision.
Excluded group are those who were released to supervision.

Type of Post-Release Treatment[1]
ETXIND: Coded 1 if began receiving individual counseling services only during first month under supervision by a Probation officer.
ETXGRP: Coded 1 if began receiving group counseling services only during first month under supervision by a Probation officer.
ETXBTH: Coded 1 if began receiving both individual and group counseling services during first month under supervision by a Probation officer.
Excluded group had no post-release treatment upon release to supervision.

Type of Subject (used for Heckman model)
WASTREAT: Conditional probability of completing treatment if entered treatment.

Unemployed Before Incarceration
NVERUNEM: Coded 1 if never unemployed 30 or more days.
NVRWORK: Coded 1 if never in the work force.
Excluded group was unemployed at least once for 30 or more days.

Vocational education during incarceration[1]
EVOCEDUC: Coded 1 if participated in vocational education during incarceration.
Excluded group did not participate in vocational education.

[1] These variables serve as control for level of supervision and post-release services received and are not substantively interpreted.

APPENDIX B. MODELS FOR DEALING WITH SELECTION BIAS

This appendix develops two mathematical models of recidivism. One is based on a lognormal survival model and the other is based on an exponential survival model. We discuss both models in section B.1, because they raise similar analytic problems, which have similar solutions. In section B.2, we introduce a form of selection bias into both models and develop an estimation procedure (maximum likelihood) that yields consistent parameter estimates of the treatment effect provided the model is true. Deriving those estimates requires a model of the process by which subjects get into treatment, which we develop in section B.3. A probit model is developed in B.4; a two-limit tobit model appears in B.5.

B.1 Basic Recidivism Model

Upon release from prison (including confinement in a half-way house), every offender has a *propensity* to recidivate. Recidivism means either that the offender was rearrested or that he tested positive for an illegal drug. These two events are analyzed separately. The propensity to recidivate can be expressed as a non-negative, increasing function of an underlying latent propensity score, Z. This score is in turn assumed to be a linear function of a dummy variable (coded 1 when the offender was treated and coded zero otherwise) and a vector of control variables. Thus, the propensity score is written:

$$Z_i = \alpha_0 + \alpha_1 TR_i + \alpha_2 X_i + \sigma\epsilon_{1i} \qquad \textbf{(1)}$$

where:

- Zi — a latent variable, measured on a continuous scale, so that within a specified time the probability of recidivism for the ith individual decreases as Zi increases.
- TR_I — a dummy variable coded 1 when the i^{th} offender was treated and coded 0 otherwise.
- X_i — a column vector of control variables such as age, gender, and race.
- a_0 — a scalar parameter the constant term.
- a_1 — a scalar parameter the treatment effect.
- a_2 — a row vector of parameters associated with the control variables.

e_{1i} a random error term, identically and independently distributed as standard normal across the sample of offenders. We use e as an error term in other equations, so the superscript 1 is introduced to distinguish error terms across equations.

s A scalar parameter. Alternatively, we might drop s from (1) and assume that e is distributed as normal with a mean of zero and variance of s^2, but the derivations are simplified by using this first specification.

We eventually adopt two different assumption about how the latent variable Z affects the distribution of time until recidivism, but it is useful to first define the density and distribution functions for time until recidivism generically, and then substitute parametric distribution functions to get the lognormal and exponential models. Let:

t_i represent time until recidivism;

$f(t_i)$ represent the density function for time until recidivism; and

$F(t_i)$ represent the cumulative distribution function for time until recidivism.

The follow up period lasts M months. If recidivism occurs within M months, then we observe the time when it occurred. Otherwise we observe that recidivism did not occur within those M months. The generic likelihood function for recidivism during the first M months is written:

$$L_1 = \prod_i \phi(T_i)^{R_i}(1-\Phi(6))^{1-R_i} \quad \textbf{(2)}$$

where:

L_1 is the generic likelihood function for a survival model with censoring at M months;

T_i is the time (in months) until recidivism for the i^{th} subject when recidivism is observed;

R_i is coded 1 when recidivism happens within the six-month follow up period and is coded 0 otherwise.

This generic likelihood function is standard for survival models (Kalbfeisch and Prentice, 1980; Lancaster, 1990). It is readily changed into the likelihood for the lognormal survival model by substituting the lognormal density and distribution functions into the generic form, and likewise,

it is transformed into a variation of the exponential survival model by substituting density and distribution functions based on a modification of the exponential distribution. We take those steps below.

Following diagnostic tests, it might be reasonable to assume that the time until an arrest follows a lognormal distribution. In this case, $\ln(t_i) = Z_i$, and the density function for time until an arrest is written:

$$\phi_A(t_{A_i}) = \frac{e^{-0.5\frac{(\ln(t_{A_i})-\alpha_0-\alpha_1 TR_i-\alpha_2 X_i)^2}{\sigma^2}}}{t_{A_i}\sqrt{2\pi\sigma^2}} \qquad \textbf{(3)}$$

where:

$f_A(t_{Ai})$ represents the lognormal density function for the distribution of time until arrest;

t_{Ai} time of arrest.

Substituting the lognormal density (3) and its distribution function into the generic likelihood function (2) yields the likelihood function for the lognormal survival model.

Also using diagnostic tests, time until a positive urine screen might follow an exponential distribution. The propensity to recidivate (1) is now written in the form:

$$\lambda_i = e^{Z_i} \qquad \textbf{(4)}$$

Unlike the usual exponential model, this specification has an error term e_1 that must be taken into account in the analysis (see Heckman and Singer, 1985). This introduction of an error term into is a convenient and realistic[1] way to introduce selection bias into the model, although it does

[1]The models developed here are sometimes called mixture models (Lancaster, 1990), and the h(e) is sometimes called the mixture distribution. Estimates of the parameters in the distribution of greatest interest to us (e.g. the exponential) are sensitive to the assumptions made about the mixture distribution (Yamaguchi, 1986). A literature on criminal careers (Spelman, 1994) reports that offense rates have a skewed distribution across offenders, and this finding

complicate the mathematics behind the development of the survival model. Thus, the density function for the time until recidivism is now written as the integral of a mixture distribution:

$$\phi_U(t_{U_i}) = \int_{\epsilon=-\infty}^{\epsilon=+\infty} \lambda_i e^{-\lambda_i t_{U_i}} \eta(\epsilon_1) d\epsilon_1 \quad \textbf{(5)}$$

where:

$f_U(t_{Ui})$ represents the density function for the distribution of time until a positive urine test;

t_{Ui} time until a positive urine test;

$h(e_1)$ the standard normal density function.

The integration removes the unobserved e_1 from the distribution. However, the presence of e_1 will not be innocuous in discussions to follow. Equation (5) has no closed-form equivalent expression and requires numerical integration. Of course, this is also true of its cumulative distribution function, which requires a second integration over t_{Ui} from 0 to T_i.

B.2 Introducing Selection Bias

A problem occurs when subjects who receive treatment are selected on a non-random basis. This may happen because subjects self-select for treatment or because treatment personnel are selective, or both. To build selection bias into the lognormal and exponential models, we introduce a second latent variable, the propensity to enter treatment:

$$Y_i = \beta_0 + \beta_1 X_i + \epsilon_{2i} \quad \textbf{(6)}$$

Here:

---

might be extended to assume that time until recidivism will be similarly skewed, so that the error distributions chosen for this analysis have some justification. Others (Schmidt and Witte, 1988; Rhodes, 1989) have found the lognormal to be a useful distribution for explaining recidivism. Nevertheless, future analyses will test the sensitivity of results to alternative assumptions made about the mixture distribution. For example, by using a power transformations (such as the Box-Cox power transformation), the distribution h(e) can be extremely flexible. Such tests are planned for the future.

Y_i a latent variable. The higher the value of Y, the more likely a person will enter treatment;

X_i a column vector of control variables, the same as defined earlier;

b_0 a scalar parameter;

b_1 a row vector of parameters conformable with X;

e_{2i} a random error term that is distributed as standard normal;

and

when Y 0, then treatment occurs (TR=1), and

when Y < 0, then treatment does not occur (TR=0).

Unless 1 and e_2 are statistically independent, the variable representing treatment (TR) will not be independent of e_1. It seems unlikely that the two will be independent, because they both are affected by excluded variables, such as motivation to change behavior. This correlation will cause the parameter estimate of the treatment effect (a_1) to be biased and inconsistent unless it is taken into account in the analysis.

One approach to overcoming this problem is to assume a parametric form for the joint distribution between e_1 and e_2, and to take that joint distribution into account in the likelihood functions (equation 2). Assuming that the two are distributed as bivariate normal, two cases are pertinent, the first for time until an arrest and the second for time until a positive urine test. Considering the first case (the lognormal distribution), the density function expressed previously as equation (3) is correct only for those cases that come from the non-DAP facility. For people who receive treatment, we use the conditional density function as represented by equation (7) in place of (3).

$$\phi_A(t_{A_i}|TR_i=1) = \frac{e^{-0.5\frac{(\ln(t_{A_i})-\alpha_0-\alpha_1 TR_i-\alpha_2 X_i)^2}{\sigma_i^2}}}{t_{A_i}\sqrt{2\pi\sigma_i^2}} \frac{H\left(\frac{\beta_0+\beta_1 X_i+\rho\frac{\ln(t_{A_i})-\alpha_0-\alpha_1 TR_i-\alpha_2 X_i}{\sigma_i}}{\sqrt{1-\rho^2}}\right)}{H(\beta_0+\beta_1 X_i)} \quad (7)$$

and for people who do not enter treatment and were members of the DAP comparison group, we

use the conditional density function represented by (8) in place of (3).

$$\phi_A(t_{A_i}|TR_i=0) = \frac{e^{-0.5\frac{(\ln(t_{A_i})-\alpha_0-\alpha_2 X_i)^2}{\sigma_i^2}}}{t_{A_i}\sqrt{2\pi\sigma_i^2}} \frac{H_c\left(\dfrac{\beta_0+\beta_1 X_i+\rho\dfrac{\ln(t_{A_i})-\alpha_0-\alpha_2 X_i}{\sigma_i}}{\sqrt{1-\rho^2}}\right)}{H_c(\beta_0+\beta_1 X_i)} \quad \textbf{(8)}$$

where:

H the standard normal cumulative distribution function;

H_c the complement of the standard normal cumulative distribution function;

r the correlation between e_1 and e_2.

The conditional density functions (7) and (8) have cumulative distribution counterparts, which must also be substituted into (2). We do not show those distribution functions because they are just the appropriate specification of the bivariate normal cdf divided by the probability that the subject was treated (7) or was not treated (8).

The general approach to deriving this likelihood is explained in Maddala (1983, p. 266). Briefly, we start with the bivariate normal density involving e_1 and e_2. This can be written as $h(e_1)h(e_2|e_1)$. We integrate this over the appropriate range for e_2 to get the joint probability of t_A and entering treatment (equation (7)) or not entering treatment (equation (8)). We divide the results by the unconditional probability of entering treatment (equation (7)) or not entering treatment (equation (8)).

In essence, then, the likelihood function is different depending on whether the subject came from a non-DAP facility, came from a DAP facility but did not enter treatment, or came from a DAP facility and entered treatment. Nevertheless, the generic likelihood (2) holds; we just substitute the correct density and distribution function depending on whether the subject is a member of the non-DAP control group, the DAP comparison group, or the DAP treatment group.

The generic likelihood function also has to be modified when the exponential model is used. When a subject comes from a non-DAP facility, equation (5) represents the density function. When the subject comes from a DAP facility and receives treatment, we use (9) in place of (5)

$$\phi_U(t_{U_i}|TR_i=1) = \int_{\epsilon=-\infty}^{\epsilon=+\infty} \lambda_i e^{-\lambda_i t_{U_i}} \eta(\epsilon_1|TR_i=1)d\epsilon_1 \quad \textbf{(9)}$$

and when the subject comes from a DAP facility but does not receive treatment then we use (10) in place of (5)

$$\phi_U(t_{U_i}|TR_i=0) = \int_{\epsilon=-\infty}^{\epsilon=+\infty} \lambda_i e^{-\lambda_i t_{U_i}} \eta(\epsilon_1|TR_i=0)d\epsilon_1 \quad \textbf{(10)}$$

where:

$h(e_1|TR_i=1)$ is the normal density function conditional on $TR_i = 1$, and

$h(e_1|TR_i=0)$ is the normal density function conditional on $TR_i = 0$.

and numerical integration was used to get these conditional distributions, because there is no closed-form expression. The density function for the error terms in (9) and (10) conditional on TR can be written:

$$\eta(\epsilon_1|TR_i=1) = \frac{\int_{\epsilon_2=-\beta_0-\beta_1 X_1}^{\infty} \eta_b(\epsilon_1,\epsilon_2,\rho)d\epsilon_2}{\int_{\epsilon_2=-\beta_0-\beta_1 X_i}^{\infty} \eta(\epsilon_2)d\epsilon_2} \quad \textbf{(11)}$$

$$\eta(\epsilon_1|TR_i=1) = \frac{\int_{\epsilon_2=-\beta_0-\beta_1 X_1}^{\infty} \eta_b(\epsilon_1,\epsilon_2,\rho)d\epsilon_2}{\int_{\epsilon_2=-\beta_0-\beta_1 X_i}^{\infty} \eta(\epsilon_2)d\epsilon_2} \quad \textbf{(12)}$$

where:

h_b represents the density function for the bivariate normal (standard normal in this case), and

r represents the correlation between e_1 and e_2;

and a similar expression exists for $h(e_1|TR_i=0)$. As before, the density functions have cumulative distribution (over t_U) function counterparts. These must be numerically computed with a double integral and substituted, as appropriate, into (2).

The likelihood function is different depending on whether the subject came from the non-DAP control group, the DAP comparison group, or the DAP treatment group. The generic likelihood (2) holds; we substitute the correct density and distribution function depending on whether the subject is a member of the non-DAP control group, the DAP comparison group, or the DAP treatment group.

B.3 Estimating the Probability of Selection into Treatment

Applying the adjustment described above for selection bias requires an estimate of b. Although the a and b parameters could be estimated jointly, it is easier (although less efficient) to estimate the b parameters from the probit model (equation (6)) and then maximize the likelihood expression (equation 2, after the appropriate substitutions) conditional on those estimates of b. Estimation of the probit model was not straightforward. Because we sampled the DAP comparison cases, we had to take that sampling into account by including the probability of being sampled as part of the likelihood function for the probit model. Thus, the probit model needs to be based on the joint probability of two events: entering treatment or not entering treatment, and being selected into the study sample. DAP treatment cases were selected with certainty, so they have a conditional selection probability equal to one, and non-DAP cases do not enter into this estimation, because those cases have a zero probability of entering treatment.[2] The likelihood for this model is written:

$$L_2 = \prod_i \frac{\mathrm{H}(\beta_0+\beta_1 X_i)^{TR_i}\left(PS_i\left(1-\mathrm{H}(\beta_0+\beta_1 X_i)\right)\right)^{1-TR_i}}{\mathrm{H}(\beta_0+\beta_1 X_i)+(PS_i(1-\mathrm{H}(\beta_0+\beta_1 X_i))} \qquad \textbf{(13)}$$

[2]Actually, the DAP cases that received treatment were sampled with less than certainty. Assuming a sampling probability of one is convenient however, provided PS is adjusted accordingly.

where:

PS_I is the probability of selection into the study sample for the i^{th} case. When the subject received treatment, the probability is 1, because all treated subjects were included in the sample.

The logic of this approach is that the probit model represents the probability of occurrence of two events. In the first event, a subject either is selected for treatment or he is not selected for treatment. The second event being included in the sample is then conditional on the outcome of the first event. If the subject entered treatment, then he was included in the sample, but if he did not enter treatment, he was included in the sample with a probability of PS_i. The likelihood function reflect the joint probability of those two events.

B.4 A Probit Model of Halfway House Failures

We chose to analyze failures in halfway house assignments as a dichotomous dependent variable failure (coded 1) and success (coded 0). This decision suggested that a probit model would be an appropriate way to analyze outcomes. As before, we assume that every individual who is placed in a halfway house has a propensity to fail, expressed as a latent variable:
We have reused notation from above because there seems to be little risk of confusion. The Z again represents the latent variable, but now it applies to the propensity to fail in a halfway house confinement.

$$Z_i = \alpha_0 + \alpha_1 TR_i + \alpha_2 X_i + \epsilon_{3i}$$

An inmate fails when:

$Z_i \quad 0$

and he succeeds when:

$Z_i < 0$

Assuming that e_3 and e_2 are distributed as bivariate normal, the likelihood function for estimating the a can be written as

$$L_{41} = \prod_{i \in nonDAP\ CONTROL} \mathrm{H}(\alpha_0 + \alpha_2 X_i)^{h_i}\left[1 - \mathrm{H}(\alpha_0 + \alpha_2 X_i)\right]^{1-h_i}$$

for the non-DAP control group; as

$$L_{42} = \prod_{i \in DAP\ COMPARISON} \mathrm{N}(\alpha_0 + \alpha_2 X_i | TR_i = 0)^{h_i}\left[1 - \mathrm{N}(\alpha_0 + \alpha_2 X_i | TR_i = 0)\right]^{1-h_i}$$

for the DAP comparison group; and as

$$L_{43} = \prod_{i \in DAP\ TREATED} \mathrm{N}(\alpha_0 + \alpha_1 TR_i + \alpha_2 X_i | TR_i = 1)^{h_i}\left[1 - \mathrm{N}(\alpha_0 + \alpha_1 TR_i + \alpha_2 X_i | TR_i = 1)\right]^{1-h_i}$$

for the DAP treatment group, where:

h_i equals 1 when the subject failed and equals zero otherwise.

$N(a_0+a_2X_i|TR_i=0)$ represents the distribution function conditional on the i^{th} subject s not being treated and $N(a_0+a_1TR_i+a_2X_i|TR_i=1)$ represents the distribution conditional on the i^{th} subject s being treated:

$$\mathrm{N}(\alpha_0 + \alpha_1 TR_i + \alpha_2 X_i | TR_i = 1) = \frac{\mathrm{H}_b(\alpha_0 + \alpha_1 TR_i + \alpha_2 X_i, \beta_0 + \beta_1 X_i, \rho)}{\mathrm{H}(\beta_0 + \beta_1 X_i)}$$

where:

H_b is the bivariate normal distribution function (standard normal in this case);

and a similar expression exists for $N(a_0+a_2X_i|TR_i=0)$.

The likelihood function is then written:

$$L_3 = L_{31}L_{32}L_{33}$$

B.5 A Two-Limit Tobit Model of Employment

We measured post-release employment as percentage of time employed during the six-month follow up period. This could range from 0 for those who were never employed to 100 percent for those who were always employed. Both extremes were observed in the data.

Although an ordinary least squares regression might be used to analyze this outcome, OLS regression suffers from three problems when applied in this context. The first problem is that parameter estimates will be biased and inconsistent, because the outcomes have upper and lower limits, which are not taken into account by the estimation procedure. The second problem is that the standard errors will be inconsistent, because the error terms will necessarily be heteroscedastic. The third problem is that selection bias still needs to be taken into account. Although feasible generalized least squares can be used to deal with all these problems, an alternative approach is to use a two-limit tobit model (Maddala, p. 160).

As used here, this model assumes that the every offender has a propensity to be employed. Reusing the earlier notation, we write this propensity as:

$$Z_i = \alpha_0 + \alpha_1 TR_i + \alpha_2 X_i + \epsilon_{4i}$$

The subject is unemployed at all times when

$Z_i < 0$,

and he is employed full time when

$Z_i > 100$,

and otherwise, time employed (TE_i) equals the latent variable, so:

$TE_i = Z_i$ when Z_i 0 and Z_i 100

The unknown parameters can be estimated by maximum likelihood. As before, we have to account for three conditions. When the study subject comes from the non-DAP control group,

the likelihood is:

$$L_{41} = \prod_{i \in non\text{-}DAP\ CONTROL} \mathrm{H}\left(\frac{-\alpha_0 - \alpha_2 X_i}{\sigma} \right)^{E_{1i}} \left[\frac{\eta\left(\frac{TE_i - \alpha_2 X_i}{\sigma} \right)}{\sigma} \right]^{E_{2i}}$$

$$\left[1 - \mathrm{H}\left(\frac{100 - \alpha_0 - \alpha_1 TR_i - \alpha_2 X_i}{\sigma} \right) \right]^{1 - E_{1i} - E_{2i}}$$

When the subject come from the DAP treatment group, the likelihood is:

$$L_{42} = \prod_{i \in DAP\ TREATED} \mathrm{N}\left(\frac{-\alpha_0 - \alpha_1 TR_i - \alpha_2 X_i}{\sigma} | TR_i = 1 \right)^{E_{1i}} \left[\frac{\nu\left(\frac{TE_i - \alpha_0 - \alpha_1 TR_i - \alpha_2 X_i}{\sigma} | TR_i = 1 \right)}{\sigma} \right]^{E_{2i}}$$

$$\left[1 - \mathrm{N}\left(\frac{100 - \alpha_0 - \alpha_1 TR_i - \alpha_2 X_i}{\sigma} | TR_i = 1 \right) \right]^{1 - E_{1i} - E_{2i}}$$

and when the subject comes from the DAP comparison group, the likelihood is:

$$L_{43} = \prod_{i \in DAP\ COMPARISON} \mathrm{N}\left(\frac{-\alpha_0 - \alpha_2 X_i}{\sigma} \middle| TR_i = 0 \right)^{E_{1i}} \left[\frac{\nu\left(\frac{TE_i - \alpha_0 - \alpha_2 X_i}{\sigma} | TR_i = 0 \right)}{\sigma} \right]^{E_{2i}}$$

where:

E_{1i} equals 1 when the subject was unemployed for the entire follow up period;

$$\left[1-\mathrm{N}\left(\frac{100-\alpha_0-\alpha_1 TR_i-\alpha_2 X_i}{\sigma}\middle| TR_i=0\right)\right]^{1-E_{1i}-E_{2i}}$$

E_{2i} equals 1 when the subject was employed part (but not all) the follow up period.

As before, N and n represent the conditional distribution and density function, respectively. The conditional distribution function has already been presented as part of the probit model, and the density is similar to that for the uncensored part of the lognormal model, except that the dependent variable is in natural rather than logarithmic units.

Thus:

The likelihood function for the two-limit tobit model is written:

$$L_4 = L_{41}L_{42}L_{43}$$