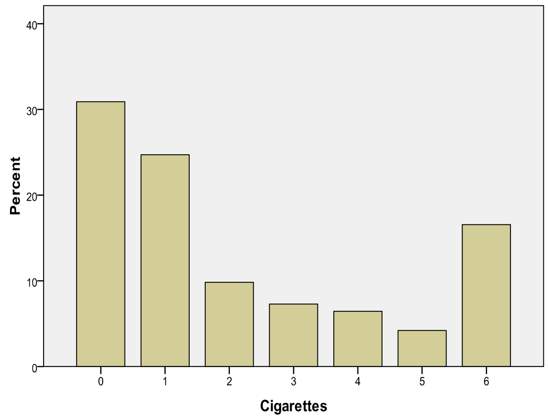
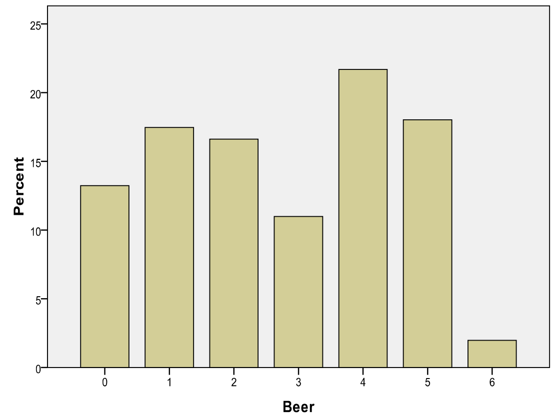
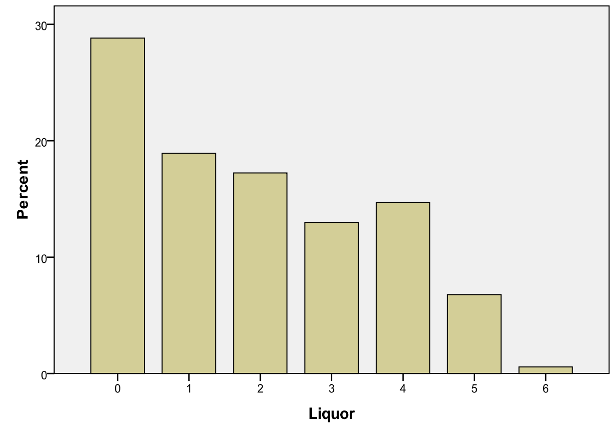
Measurement Models for Other Kinds of Continuous but Non-Normal Outcomes in Mplus version 7.11

This example examines alternative factor models for 6 outcomes that measure use of controlled substances on a scale of 0 to 6, where: 0 = Never used, 1 = Have used once or twice, 2 = Once or twice a year, 3 = Less than once a month, 4 = Once or twice a month, 5 = Once or twice a week, and 6 = Daily. Below are the distributions of the outcomes in a sample of 356 rural adolescents. It is admittedly not the best example because of the constrained 7-point ordinal scale rather than a true count, but it is what I have to illustrate these models…

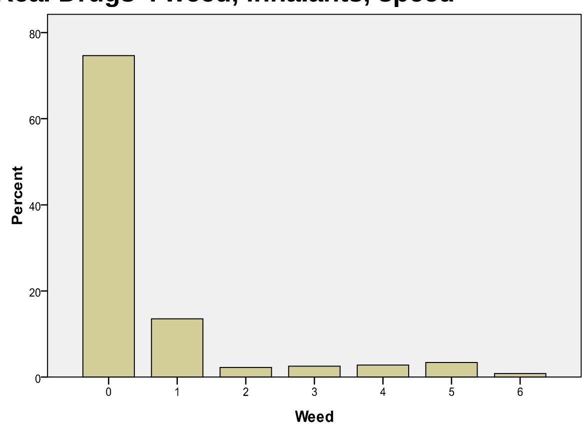
"Recreational Drugs": cigarettes, beer, and liquor

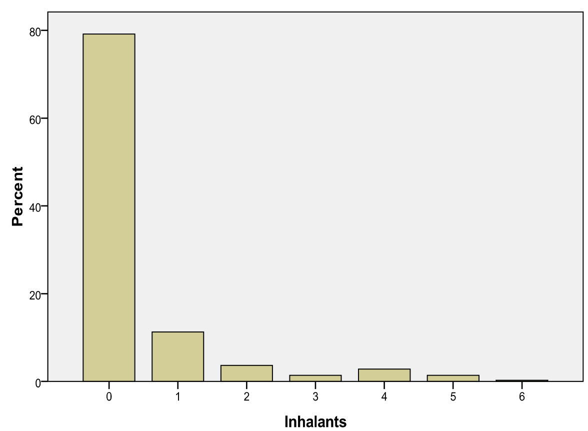


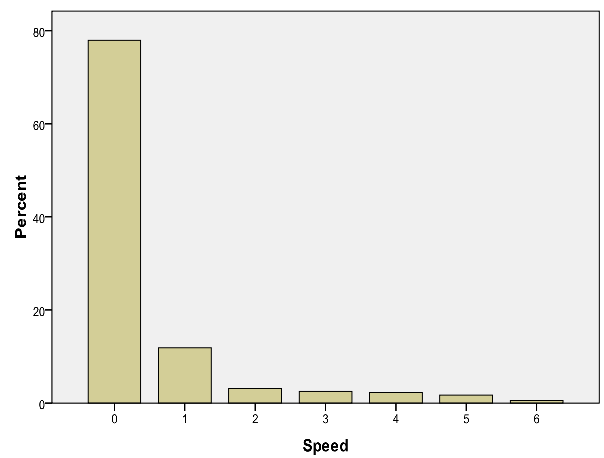




"Real Drugs": weed, inhalants, speed







Assuming we wish to model the distribution as some kind of continuum (i.e., not as graded response), there are several reasonable options described below for factor models that assume different conditional item response distributions. We will see examples of how to specify each of these in Mplus next, using MLR (robust ML) for all models. Model fit statistics will only be available for normal models that are easily summarized by a covariance matrix, though.

Normal (regular CFA) Model: We fit a linear model of the factor predicting the ORIGINAL item response and assume each item follows a conditionally normal distribution (i.e., the item residuals are normally distributed after controlling for the factor). The measurement model would thus include per item an intercept (the expected item response when factor = 0), a factor loading (change in item response per unit change in the factor), and a residual variance (amount of item variance not predicted by the factor). A normal model that assumes a linear relationship between the item response and the factor (i.e., an interval scaling of the response options) is not likely to be tenable for these kinds of data, but it's the most common approach. Although we can use MLR instead of ML for CFA models when item response distributions look non-normal to correct the fit statistics and standard errors accordingly, that doesn't solve the basic problem of whether it is reasonable to expect a linear relationship between the item response and the factor. The alternative models below address this latter problem.

Poisson and Negative Binomial Models: We fit a linear model of the factor predicting the LOG of the item response. We assume the items follow a **Poisson** distribution in which the mean is the same as the variance (a single parameter called "k"). The Poisson measurement model would thus include per item an intercept (the expected LOG of the item response when factor = 0) and a factor loading (change in the LOG of the item response per unit change in the factor), but no estimated residual variance (because it is determined by the conditional mean). In the closely related **Negative Binominal** model, we add to the Poisson model a scaling factor "α" that allows the residual variance to exceed the mean (called "over-dispersion"), such that the new variance = k(1+kα). In Mplus we can test if the scaling factor is different than 0 (because 0=Poisson), and thus we could do a nested model comparison as to whether a Negative Binominal fits better than a Poisson for each item response. These models works well for integer count data that can't be negative or data that are skewed, but they run into problems if the over-dispersion is caused by an excess of zeros. The alternative models below address this extra-zero problem.

Zero-Inflated Poisson (ZIP) or Zero-Inflated Negative Binomial (ZINB): These models specify two underlying distributions in the observed item responses: "structural zeros" and "non-structural zeros" (includes expected zeros based on regular Poisson or negative binomial distributions). A structural zero would never do any of the behaviors in question, whereas an expected zero (who belongs in the regular distribution) might do the general behavior, just not that particular item (e.g., zero for use of speed but non-zero for use of weed). We can potentially fit a factor model to each part of the distribution. The structural zero measurement model would have a linear model of the factor predicting the LOGIT of being a structural zero (so the "higher category" being predicted is the structural zero as 1). The structural zero factor model would thus estimate thresholds (expected LOGIT of being a non-structural zero if the factor = 0) and factor loadings (change in the LOGIT of being a structural zero per unit change in the factor). The non-structural-zero measurement model would have a linear model of its factor predicting the LOG of the item response, and the ZINB would again have an added scaling parameter for over-dispersion. Thus the non-structural-zero measurement model would estimate intercepts (expected LOG of item response if factor = 0) and factor loadings (change in LOG of item response per unit change in factor). Just as the Poisson is nested within the Negative Binomial (tests if the scaling parameter for extra residual variance is needed), the ZIP is nested within the ZINB. In addition, the AIC and BIC can be compared between the Poisson and ZIP, or between the Negative Binomial and ZINB, to see if the zero-inflation parameters are helpful. It is not required to have a factor for the inflation, but one can do so in Mplus (very hard to estimate). However, the interpretation of two kinds of zeros can be confusing, and so the alternative models below address the issue of excess zeros more directly.

Negative Binomial Hurdle and Two-Part Models: Rather than trying to distinguish "structural zeros" from "non-structural zeros", these models simply split each observed item response into two new variables: "0 vs. something", and "how much if not 0". The models differ in how they accomplish this same idea. The **negative binomial hurdle model** for "0 vs. something" uses "0" as what is predicted. Thus, the measurement model for the "not 0 vs. 0" part would have a linear model of its factor predicting the LOGIT of being a 0. It thus estimates a threshold (expected LOGIT of not being 0 if factor = 0) and a factor loading (expected change in the LOGIT of being a 0 per unit change in the factor). The negative binomial hurdle measurement model for the "not 0" part would have a linear model of its factor predicting the LOG of the item response past 0 (a zero-truncated distribution). It thus estimates an intercept (expected LOG of the non-zero item response if factor = 0), a factor loading (expected change in the LOG of the non-zero item response per unit change in the factor), and a scaling parameter for the over-dispersion of the residual variance. **The two-part model** for the "0 vs. something" uses "something" as what is predicted. Thus, the measurement model for the "0 vs. not 0" part would have a linear model of its factor predicting the LOGIT of being not 0. It thus estimates a threshold (expected LOGIT of being 0 if factor = 0) and a factor loading (expected change in the LOGIT of being not 0 per unit change in the factor). The two-part measurement model for the "not 0" part would have a linear model of its factor predicting the LOG of the item response past 0. It thus estimates an intercept (expected LOG of the non-zero item response if factor = 0), a factor loading (expected change in the LOG of the non-zero item response per unit change in the factor), and a residual variance. The "not 0" model uses a LOG transformation by default, but other transformations (including none) are available as well.

Here is one alternative: (1) Normal CFA model with Robust ML

```
TITLE:  Model 1: Normal Response Distribution
DATA:   FILE IS deviance.dat;

VARIABLE:   NAMES ARE        cig beer liquor weed inhale speed;
            USEVARIABLES ARE cig beer liquor weed inhale speed;
            MISSING ARE .;
! No extra code here means we assume each item response is normal


ANALYSIS:   ESTIMATOR IS MLR;
OUTPUT:     RESIDUAL STDYX;

MODEL:
! Factor loadings all estimated
   Rec  BY cig* beer* liquor*;
   Drug BY weed* inhale* speed*;
! Intercepts all estimated
   [cig* beer* liquor* weed* inhale* speed*];
! Residual variances all estimated
    cig* beer* liquor* weed* inhale* speed*;
! Factor mean=0 and variance=1 for identification, factors correlate
   [Rec@0 Drug@0]; Rec@1 Drug@1; Rec WITH Drug*;


Number of Free Parameters                       19
Loglikelihood
          H0 Value                       -3408.733
          H0 Scaling Correction Factor      2.0743
            for MLR
          H1 Value                       -3388.791
          H1 Scaling Correction Factor      1.8171
            for MLR

Information Criteria
          Akaike (AIC)                    6855.466
          Bayesian (BIC)                  6929.090
          Sample-Size Adjusted BIC        6868.813
            (n* = (n + 2) / 24)

Chi-Square Test of Model Fit
          Value                             33.067*
          Degrees of Freedom                     8
          P-Value                           0.0001
          Scaling Correction Factor          1.206
            for MLR

RMSEA (Root Mean Square Error Of Approximation)
          Estimate                           0.094
          90 Percent C.I.                    0.062  0.128
          Probability RMSEA <= .05           0.014

CFI/TLI
          CFI                                0.947
          TLI                                0.900
```

```
MODEL RESULTS
                                                    Two-Tailed
                    Estimate       S.E.  Est./S.E.    P-Value
FACTOR LOADINGS: CHANGE IN ACTUAL Y PER SD CHANGE IN FACTOR
 REC      BY
    CIG                1.179      0.115     10.267      0.000
    BEER               1.561      0.062     25.195      0.000
    LIQUOR             1.359      0.066     20.667      0.000

 DRUG     BY
    WEED               1.037      0.121      8.582      0.000
    INHALE             0.805      0.096      8.412      0.000
    SPEED              0.857      0.096      8.930      0.000

CORRELATION BETWEEN KINDS OF DRUG USE
 REC      WITH
    DRUG               0.613      0.051     12.110      0.000

EXPECTED ACTUAL Y WHEN FACTOR IS 0
 Intercepts
    CIG                2.126      0.116     18.257      0.000
    BEER               2.726      0.093     29.221      0.000
    LIQUOR             1.886      0.088     21.486      0.000
    WEED               0.593      0.070      8.487      0.000
    INHALE             0.431      0.056      7.670      0.000
    SPEED              0.468      0.060      7.871      0.000

Residual Variances – AMOUNT OF ITEM VARIANCE THAT IS NOT THE FACTOR
    CIG                3.439      0.292     11.757      0.000
    BEER               0.658      0.141      4.669      0.000
    LIQUOR             0.894      0.108      8.282      0.000
    WEED               0.636      0.175      3.628      0.000
    INHALE             0.462      0.112      4.120      0.000
    SPEED              0.510      0.118      4.332      0.000

STDYX Standardization
                                                     Two-Tailed
                    Estimate       S.E.  Est./S.E.    P-Value

 RECAMT   BY: CORRELATION BETWEEN ACTUAL Y AND FACTOR
    CIG                0.537      0.048     11.096      0.000
    BEER               0.887      0.026     34.184      0.000
    LIQUOR             0.821      0.026     31.516      0.000

 DRUGAMT  BY
    WEED               0.793      0.062     12.834      0.000
    INHALE             0.764      0.053     14.322      0.000
    SPEED              0.768      0.048     15.922      0.000

 RECAMT   WITH
    DRUGAMT            0.613      0.051     12.110      0.000
```

Here are two more alternatives: (2a) Poisson Factor Model and (2b) Negative Binomial/Poisson Factor Model

```
TITLE:  Model 2a: Poisson for all;
DATA:   FILE IS deviance.dat;
VARIABLE:   NAMES ARE         cig beer liquor weed inhale speed;
            USEVARIABLES ARE cig beer liquor weed inhale speed;
            MISSING ARE .;
! Tells Mplus which distribution each item response should get
COUNT ARE cig (p) beer (p) liquor (p) weed (p) inhale (p) speed (p);
ANALYSIS:   ESTIMATOR IS MLR;
OUTPUT:     RESIDUAL; ! STDYX ! standardized doesn't make any sense
MODEL:
! Factor loadings all estimated
   Rec  BY cig* beer* liquor*;
   Drug BY weed* inhale* speed*;
! Intercepts all estimated
   [cig* beer* liquor* weed* inhale* speed*];
! Factor mean=0 and variance=1 for identification, factors correlate
    [Rec@0 Drug@0]; Rec@1 Drug@1; Rec WITH Drug*;

Number of Free Parameters                       13
Loglikelihood
          H0 Value                       -2664.557
          H0 Scaling Correction Factor      0.8884
            for MLR
Information Criteria
          Akaike (AIC)                    5355.113
          Bayesian (BIC)                  5405.488
          Sample-Size Adjusted BIC        5364.246
            (n* = (n + 2) / 24)
                                                    Two-Tailed
                    Estimate       S.E.  Est./S.E.    P-Value
FACTOR LOADINGS: CHANGE IN LOG(Y) PER SD CHANGE IN FACTOR
 REC      BY
    CIG                0.879      0.065     13.562      0.000
    BEER               0.538      0.042     12.892      0.000
    LIQUOR             0.713      0.059     12.108      0.000
 DRUG     BY
    WEED               2.527      0.210     12.038      0.000
    INHALE             2.476      0.259      9.565      0.000
    SPEED              2.613      0.241     10.845      0.000

CORRELATION BETWEEN KINDS OF DRUG USE – MUCH LARGER NOW…
 REC      WITH
    DRUG               0.952      0.032     29.753      0.000
EXPECTED LOG(Y) WHEN FACTOR IS 0
 Intercepts
    CIG                0.425      0.084      5.066      0.000
    BEER               0.872      0.047     18.357      0.000
    LIQUOR             0.413      0.069      5.994      0.000
    WEED              -2.555      0.281     -9.107      0.000
    INHALE            -2.809      0.328     -8.571      0.000
    SPEED             -2.899      0.333     -8.711      0.000

NO RESIDUAL VARIANCES WERE ESTIMATED (ARE DETERMINED INSTEAD)
```

```
TITLE:  Model 2b: Poisson for all; Negative Binomial for CIG only
DATA:   FILE IS deviance.dat;
VARIABLE:   NAMES ARE         cig beer liquor weed inhale speed;
            USEVARIABLES ARE cig beer liquor weed inhale speed;
            MISSING ARE .;
! Tells Mplus which distribution each item response should get
COUNT ARE cig (nb) beer (p) liquor (p) weed (p) inhale (p) speed (p);
ANALYSIS:   ESTIMATOR IS MLR;
OUTPUT:     RESIDUAL;
MODEL:
! Factor loadings all estimated
   Rec  BY cig* beer* liquor*;
   Drug BY weed* inhale* speed*;
! Intercepts all estimated
   [cig* beer* liquor* weed* inhale* speed*];
! Factor mean=0 and variance=1 for identification, factors correlate
    [Rec@0 Drug@0]; Rec@1 Drug@1; Rec WITH Drug*;

Number of Free Parameters                       14
Loglikelihood
          H0 Value                       -2657.992
          H0 Scaling Correction Factor      0.9035
            for MLR
Information Criteria
          Akaike (AIC)                    5343.984
          Bayesian (BIC)                  5398.233
          Sample-Size Adjusted BIC        5353.819
            (n* = (n + 2) / 24)
                                                    Two-Tailed
                    Estimate       S.E.  Est./S.E.    P-Value
FACTOR LOADINGS: CHANGE IN LOG(Y) PER SD CHANGE IN FACTOR
 REC      BY
    CIG                0.805      0.077     10.497      0.000
    BEER               0.539      0.041     13.090      0.000
    LIQUOR             0.719      0.058     12.326      0.000
 DRUG     BY
    WEED               2.578      0.207     12.432      0.000
    INHALE             2.528      0.260      9.720      0.000
    SPEED              2.663      0.239     11.123      0.000

CORRELATION BETWEEN KINDS OF DRUG USE – STILL MUCH LARGER NOW…
 REC      WITH
    DRUG               0.989      0.032     31.109      0.000
EXPECTED LOG(Y) WHEN FACTOR IS 0
 Intercepts
    BEER               0.870      0.048     18.204      0.000
    LIQUOR             0.410      0.070      5.866      0.000
    WEED              -2.591      0.283     -9.141      0.000
    INHALE            -2.848      0.336     -8.484      0.000
    SPEED             -2.934      0.340     -8.639      0.000
    CIG                0.482      0.082      5.888      0.000

Dispersion – ALPHA MULTIPLIER TO INCREASE VARIANCE RELATIVE TO MEAN
    CIG                0.229      0.088      2.589      0.010
```

Here is another alternative: (3) Zero-Inflated Negative Binomial or Poisson Factor Model

```
TITLE:  Model 3: Zero-Inflated Poisson and Negative Binomial (FOR CIG)
DATA:   FILE IS deviance.dat;

VARIABLE:   NAMES ARE        cig beer liquor weed inhale speed;
            USEVARIABLES ARE cig beer liquor weed inhale speed;
            MISSING ARE .;
! Tells Mplus which distribution each item response should get
COUNT ARE cig (nbi) beer (pi) liquor (pi) weed (pi) inhale (pi)
          speed (pi);

ANALYSIS:   ESTIMATOR IS MLR;
OUTPUT:     RESIDUAL;  ! STDYX ! standardized doesn't make any sense
MODEL:
! Factor loadings all estimated for AMOUNT if Structural Non-Zero
   RecAmt  BY cig* beer* liquor*;
   DrugAmt BY weed* inhale* speed*;
! Means all estimated for inflation variables (not predicted)
   [cig#1* beer#1* liquor#1* weed#1* inhale#1* speed#1*];
! Intercepts all estimated for AMOUNT factor
   [cig* beer* liquor* weed* inhale* speed*];
! Factor mean=0 and variance=1 for identification, factors correlate
    [RecAmt@0 DrugAmt@0]; RecAmt@1 DrugAmt@1; RecAmt WITH DrugAmt*;


MODEL FIT INFORMATION
Number of Free Parameters                      20

Loglikelihood
          H0 Value                      -2654.559
          H0 Scaling Correction Factor     0.9560
            for MLR

Information Criteria
          Akaike (AIC)                   5349.118
          Bayesian (BIC)                 5426.616
          Sample-Size Adjusted BIC       5363.167
            (n* = (n + 2) / 24)
```

ZIP AND ZINB Inflation factors: Although we could have fit factors for the zero-inflation part (the logit of being a structural zero is predicted by each factor), those models showed severe convergence problems, most likely because the probability of being a structural zero was so small in this particular sample. For instance, the largest probability is for the mean of CIG#1 (logit of -2.597 = prob of .07). So we proceed with a single factor for each item for now.

Further, the AIC and BIC are higher in this zero-inflated model, suggesting that most of the items do not need “structural zeros”, or that including inflation parameters for the extra zeros does not help model fit.

```
MODEL RESULTS
                                                    Two-Tailed
                    Estimate       S.E.  Est./S.E.    P-Value

FACTOR LOADINGS: CHANGE IN LOG(Y) PER SD CHANGE IN FACTOR
APPLIES TO NON-STRUCTURAL ZEROS ONLY
 RECAMT      BY
    CIG                0.787      0.075     10.538      0.000
    BEER               0.542      0.041     13.094      0.000
    LIQUOR             0.725      0.059     12.358      0.000

DRUGAMT     BY
    WEED               2.618      0.219     11.946      0.000
    INHALE             2.472      0.297      8.315      0.000
    SPEED              2.707      0.245     11.041      0.000

CORRELATION BETWEEN KINDS OF DRUG USE IN NON-STRUCTURAL ZEROS
 REC      WITH
    DRUG               0.983      0.034     28.730      0.000

EXPECTED LOGIT OF BEING A STRUCTURAL ZERO, -15 = “TOO SMALL TO FIND”
 Means
    RECAMT             0.000      0.000    999.000    999.000
    DRUGAMT            0.000      0.000    999.000    999.000

    BEER#1           -15.000      0.000    999.000    999.000
    LIQUOR#1         -15.000      0.000    999.000    999.000
    WEED#1            -2.835      0.848     -3.344      0.001
    INHALE#1          -2.621      2.184     -1.200      0.230
    SPEED#1           -4.123      5.086     -0.811      0.418
    CIG#1             -2.597      0.598     -4.341      0.000

EXPECTED LOG(Y) WHEN FACTOR IS 0 IN NON-STRUCTURAL ZEROS
 Intercepts
    BEER               0.869      0.048     18.183      0.000
    LIQUOR             0.407      0.070      5.820      0.000
    WEED              -2.565      0.305     -8.414      0.000
    INHALE            -2.715      0.457     -5.934      0.000
    SPEED             -2.967      0.375     -7.905      0.000
    CIG                0.555      0.094      5.892      0.000

Variances
    RECAMT             1.000      0.000    999.000    999.000
    DRUGAMT            1.000      0.000    999.000    999.000

 Dispersion – ALPHA MULTIPLIER TO INCREASE VARIANCE RELATIVE TO MEAN
    CIG                0.101      0.093      1.091      0.275
```

Here is another alternative: (4) Two-Part Factor Model (here, with a log transformation of the continuous part)

```
TITLE:  Model 4: Two-Part Distributions  (0 vs. log something)
DATA:   FILE IS deviance.dat;

DATA TWOPART: ! Instructs Mplus to cut up each into 0/log of amount
    NAMES ARE       cig  beer  liquor  weed  inhale  speed;
    BINARY ARE     Bcig Bbeer Bliquor Bweed Binhale Bspeed;
    CONTINUOUS ARE Ccig Cbeer Cliquor Cweed Cinhale Cspeed;
    CUTPOINT IS 0;
    TRANSFORM IS LOG;  ! Could also use "NONE" for no transformation

VARIABLE:   NAMES ARE         cig  beer l iquor  weed  inhale  speed;
            USEVARIABLES ARE Bcig Bbeer Bliquor Bweed Binhale Bspeed
                             Ccig Cbeer Cliquor Cweed Cinhale Cspeed;
            CATEGORICAL ARE  Bcig Bbeer Bliquor Bweed Binhale Bspeed;
            MISSING ARE .;

ANALYSIS:   ESTIMATOR IS MLR;
OUTPUT:     RESIDUAL STDYX TECH4; ! TECH4 gives factor correlation matrix
MODEL:
! Factor loadings all estimated for 2 separate factors (0/amount)
   RecNot0  BY Bcig* Bbeer* Bliquor*;
   DrugNot0 BY Bweed* Binhale* Bspeed*;

   RecAmt   BY Ccig* Cbeer* Cliquor*;
   DrugAmt  BY Cweed* Cinhale* Cspeed*;

! Thresholds all estimated for binary part
    [Bcig$1* Bbeer$1* Bliquor$1* Bweed$1* Binhale$1* Bspeed$1*];
! Intercepts all estimated for continuous part
    [Ccig* Cbeer* Cliquor* Cweed* Cinhale* Cspeed*];

! Residual variances all estimated for continuous part
    Ccig* Cbeer* Cliquor* Cweed* Cinhale* Cspeed*;

! Factor mean=0 and factor variance=1 for identification
   [RecNot0@0 RecAmt@0 DrugNot0@0 DrugAmt@0];
    RecNot0@1 RecAmt@1 DrugNot0@1 DrugAmt@1;

! All factors correlated by default
    RecNot0  WITH RecAmt* DrugNot0* DrugAmt*;
    RecAmt   WITH DrugNot0* DrugAmt*;
    DrugNot0 WITH DrugAmt*;
```

```
Number of Free Parameters                       36
Loglikelihood
          H0 Value                       -1727.508
          H0 Scaling Correction Factor      0.9497
            for MLR

Information Criteria
          Akaike (AIC)                    3527.016
          Bayesian (BIC)                  3666.513
          Sample-Size Adjusted BIC        3552.305
            (n* = (n + 2) / 24)
```

```
                                                     Two-Tailed
                    Estimate       S.E.  Est./S.E.    P-Value
FACTOR LOADINGS FOR "NOT 0": CHANGE IN LOGIT(Y=SOMETHING INSTEAD OF
0) PER SD CHANGE IN  FACTOR (APPLIES TO ALL 0 VALUES)
RECNOT0     BY
    BCIG               1.350      0.210      6.415      0.000
    BBEER              3.614      0.882      4.099      0.000
    BLIQUOR            3.079      0.694      4.439      0.000
DRUGNOT0    BY
    BWEED              4.415      1.075      4.106      0.000
    BINHALE            2.712      0.474      5.716      0.000
    BSPEED             4.313      0.976      4.419      0.000

FACTOR LOADINGS FOR "AMT": CHANGE IN LOG(AMOUNT Y) PER SD CHANGE IN
FACTOR (APPLIES TO ALL NON-ZEROS)
RECAMT   BY
    CCIG               0.385      0.052      7.365      0.000
    CBEER              0.565      0.028     20.241      0.000
    CLIQUOR            0.500      0.031     16.175      0.000
DRUGAMT  BY
    CWEED              0.916      0.111      8.293      0.000
    CINHALE            0.434      0.113      3.846      0.000
    CSPEED             0.554      0.107      5.162      0.000

Thresholds – EXPECTED LOGIT(Y=0) FOR 0 VS SOMETHING WHEN FACTOR IS 0
    BCIG$1            -1.078      0.165     -6.523      0.000
    BBEER$1           -4.545      0.921     -4.934      0.000
    BLIQUOR$1         -2.012      0.420     -4.787      0.000
    BWEED$1            3.215      0.733      4.386      0.000
    BINHALE$1          2.636      0.403      6.542      0.000
    BSPEED$1           3.643      0.768      4.745      0.000
Intercepts - EXPECTED LOG(AMOUNT Y) IF NON-ZERO WHEN FACTOR IS 0
    CCIG               0.789      0.049     16.256      0.000
    CBEER              0.911      0.038     23.994      0.000
    CLIQUOR            0.676      0.038     17.737      0.000
    CWEED             -0.272      0.164     -1.657      0.097
    CINHALE            0.116      0.139      0.840      0.401
    CSPEED            -0.018      0.157     -0.116      0.908
Residual Variances – AMOUNT OF ITEM VARIANCE "NOT THE FACTOR"
    CCIG               0.413      0.036     11.384      0.000
    CBEER              0.066      0.020      3.317      0.001
    CLIQUOR            0.131      0.020      6.653      0.000
    CWEED              0.074      0.084      0.882      0.378
    CINHALE            0.260      0.060      4.345      0.000
    CSPEED             0.252      0.054      4.656      0.000

TECH4 OUTPUT:
          ESTIMATED CORRELATION MATRIX FOR THE LATENT VARIABLES
              RECNOT0       DRUGNOT0      RECAMT        DRUGAMT
              ________      ________      ________      ________
 RECNOT0         1.000
 DRUGNOT0        0.922         1.000
 RECAMT          0.659         0.707         1.000
 DRUGAMT         0.766         0.834         0.757         1.000
```

Unfortunately, absolute model fit statistics are not given for the non-normal models, and relative fit statistics (AIC and BIC) are not comparable across the normal, Poisson/NB/ZIP/ZINB, and two-part families. What we can do is examine the predicted item response across factor levels for each alternative model and see what seems reasonable. Here are the plots (made in excel) for cigarettes and for weed, with scale ends noted with the horizontal lines.

As we can see, the Negative Binomial (for cigarettes) and Poisson (for weed) dramatically overshoot the possible item response at higher levels of the factor. The same is true for the zero-inflated versions of these models. But the normal model extends below the possible scale for both items.

The two-part models seems to have the best fit – results are shown for models with either a log transformation (model 4) or no transformation of the "how much" part (those model results were not shown). They both "shut off" towards the 0 end of the scale as needed (because "0 vs. something" is covered by the other part not plotted), but the predicted "how much" doesn't have the dramatic upswing at higher factor levels like the other models. Plus they have a more straightforward interpretation than the inflated models: Here, this is the relationship between answering "how much if not 0" and the factor.

Not shown is the model for the other factor that predicts the probability of "0 vs. something" instead. Finally, we could have had the binary "0 vs. something" items and the "how much if not 0" items load onto the same factor (but fit got worse for that in these data).

```
STDYX Standardization – STANDARDIZED LOADINGS are available

                                                    Two-Tailed
                    Estimate       S.E.  Est./S.E.    P-Value

 RECNOT0  BY – CORRELATION BETWEEN LOGIT(SOMETHING) AND FACTOR
    BCIG               0.597      0.060      9.970      0.000
    BBEER              0.894      0.044     20.369      0.000
    BLIQUOR            0.862      0.050     17.227      0.000

 DRUGNOT0 BY – CORRELATION BETWEEN LOGIT(SOMETHING) AND FACTOR
    BWEED              0.925      0.033     28.436      0.000
    BINHALE            0.831      0.045     18.500      0.000
    BSPEED             0.922      0.031     29.402      0.000

 RECAMT   BY – CORRELATION BETWEEN LOG(AMOUNT) AND FACTOR
    CCIG               0.514      0.066      7.820      0.000
    CBEER              0.910      0.029     31.146      0.000
    CLIQUOR            0.810      0.035     23.391      0.000

 DRUGAMT  BY – CORRELATION BETWEEN LOG(AMOUNT) AND FACTOR
    CWEED              0.959      0.050     19.117      0.000
    CINHALE            0.649      0.133      4.888      0.000
    CSPEED             0.741      0.093      7.937      0.000
```