Appendix for "Much Ado About Nothing: A Comparison of Missing Data Methods and Software to Fit Incomplete Data Regression Models," Nicholas J. Horton and Ken P. Kleinman (2007), *The American Statistician*, 61, 79–90. This appendix is available at `http://www.math.smith.edu/muchado-appendix.pdf`

Figure A.1: Code to install and run Amelia and Zelig

```
install.packages("Amelia",repos="http://gking.harvard.edu")
source("http://gking.harvard.edu/zelig/install.R")
library(Amelia)
library(Zelig)
AmeliaView()
ds1 <- read.csv("outdata1.csv")
ds2 <- read.csv("outdata2.csv")
ds3 <- read.csv("outdata3.csv")
ds4 <- read.csv("outdata4.csv")
ds5 <- read.csv("outdata5.csv")
ds6 <- read.csv("outdata6.csv")
ds7 <- read.csv("outdata7.csv")
ds8 <- read.csv("outdata8.csv")
ds9 <- read.csv("outdata9.csv")
ds10 <- read.csv("outdata10.csv")
z.out <- zelig(as.factor(ROUTINE) ~ AGE + as.factor(ATYPE) +
    AWEEKEND + FEMALE + LOS + NDX + as.factor(RACE) +
    TOTCHG + as.factor(NSEASON), model="logit",
    data = mi(ds1,ds2,ds3,ds4,ds5,ds6,ds7,ds8,ds9,ds10))
```

Figure A.2: Code to read and analyze Amelia datasets within SAS

```
%macro readin(n);
data imp&n;
infile "c:\projects\kid\ame_out_&n..csv" delimiter = ',';
input AGE AWEEKEND FEMALE LOS NDX RACE TOTCHG ROUTINE NEWATYPE NSEASON;
_imputation_ = &n;
run;
%mend readin;

%macro rall;
%do j = 1 %to 10;
%readin(&j);
%end;
%mend;

%rall;

data Amelia;
set imp1 imp2 imp3 imp4 imp5 imp6 imp7 imp8 imp9 imp10;
run;

proc sort data = amelia; by _imputation_; run;

proc logistic data = amelia descending ;
    by _imputation_;
    class routine race aweekend nseason female newatype/param = glm;
    model routine = age race aweekend nseason female newatype los totchg ndx/covb;
    ods output ParameterEstimates=kidsparms CovB=kidsscovb;
run;

data kp2;
set kidsparms;
    if df ne 0;
    if classval0 ne '' then variable = compress(variable||classval0);
run;

proc mianalyze parms=kp2 covb(effectvar=stacking)=kidsscovb;
    modeleffects Intercept age race1 race2 race3 aweekend0
    nseason1 nseason2 nseason3 female0
    newatype1 newatype2 los totchg ndx;
run;
```

Figure A.3: Screenshot of Amelia

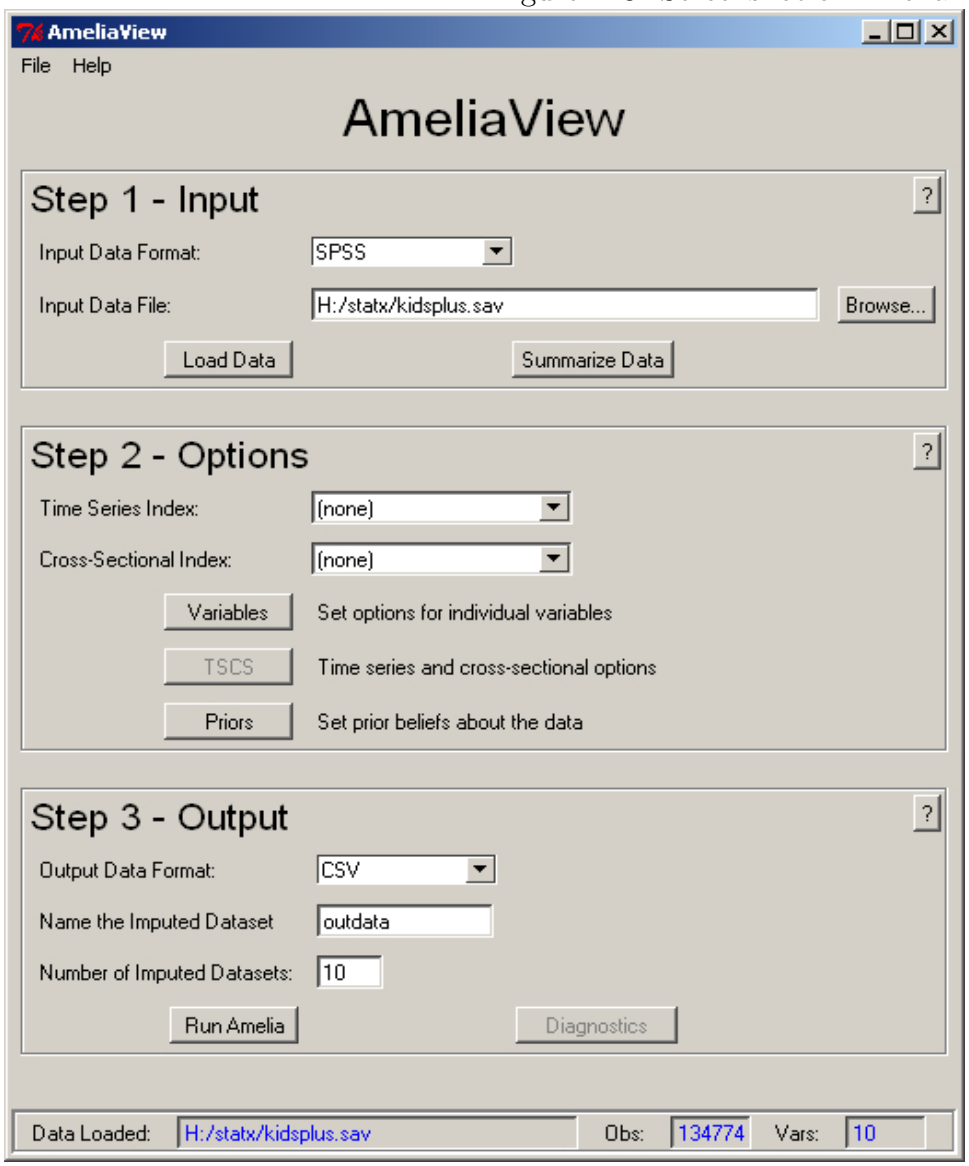


Figure A.4: Output from Zelig

```
> summary(z.out)
  Model: logit
  Number of multiply imputed data sets: 10

Combined results:
Call:
zelig(formula = as.factor(ROUTINE) ~ AGE + as.factor(ATYPE) +
    AWEEKEND + FEMALE + LOS + NDX + as.factor(RACE) + TOTCHG +
    as.factor(NSEASON), model = "logit", data = mi(ds1, ds2,
    ds3, ds4, ds5, ds6, ds7, ds8, ds9, ds10))

Coefficients:
                        Value Std. Error   t-stat    p-value
(Intercept)          2.759e+00  5.554e-02  49.6723  0.000e+00
AGE                 -3.810e-02  3.015e-03 -12.6341  1.374e-36
as.factor(ATYPE)2    2.871e-01  1.915e-02  14.9889  1.254e-44
as.factor(ATYPE)3    3.303e-01  2.489e-02  13.2728  2.865e-34
AWEEKEND            -2.677e-02  1.962e-02  -1.3644  1.724e-01
FEMALE               1.033e-01  1.584e-02   6.5216  6.956e-11
LOS                 -2.880e-03  7.144e-04  -4.0319  5.550e-05
NDX                 -1.058e-01  3.723e-03 -28.4276 9.734e-178
as.factor(RACE)2    -6.592e-02  2.381e-02  -2.7684  6.039e-03
as.factor(RACE)3    -1.163e-01  2.667e-02  -4.3621  1.641e-05
as.factor(RACE)4    -7.256e-02  3.880e-02  -1.8704  6.446e-02
TOTCHG              -4.743e-06  5.408e-07  -8.7699  2.293e-18
as.factor(NSEASON)1 -7.716e-02  2.331e-02  -3.3108  9.482e-04
as.factor(NSEASON)2 -6.684e-02  2.398e-02  -2.7878  5.344e-03
as.factor(NSEASON)3 -1.959e-02  2.465e-02  -0.7948  4.272e-01

For combined results from datasets i to j, use summary(x, subset = i:j).
For separate results, use print(summary(x), subset = i:j).
```

Figure A.5: Code to fit aregImpute within the HMisc package in R

```
library(Hmisc)
library(acepack)
source('http://biostat.mc.vanderbilt.edu/tmp/getLatestSource.s')
getLatestSource(avail=TRUE)
kidsplus <- read.csv("kidsplus.csv")
kidfact <-
data.frame(kidsplus$age,as.factor(kidsplus$atype),
    as.factor(kidsplus$nseason),kidsplus$female,
    kidsplus$los,kidsplus$routine,kidsplus$totchg,
    as.factor(kidsplus$race),kidsplus$aweekend,kidsplus$ndx)
names(kidfact) <- c("AGE","ATYPE","NSEASON",
    "FEMALE","LOS","ROUTINE","TOTCHG","RACE","AWEEKEND","NDX")
n <- naclus(kidfact)
plot(n)
naplot(n)

f  <- aregImpute(~ ROUTINE + AGE + FEMALE + ATYPE + NSEASON + LOS +
   TOTCHG + RACE + AWEEKEND + NDX, n.impute=10, defaultLinear=TRUE,
   data=kidfact)
par(mfrow=c(2,3))
plot(f, diagnostics=TRUE, maxn=2)
fmi <- fit.mult.impute(ROUTINE ~ AGE + FEMALE + ATYPE + NSEASON + LOS +
   TOTCHG + RACE + AWEEKEND+NDX, glm, f, family="binomial",data=kidfact)
impse <- sqrt(diag(Varcov(fmi)))
fcc <- glm(ROUTINE ~ AGE + FEMALE + ATYPE + NSEASON + LOS + TOTCHG +
   RACE + AWEEKEND + NDX, family=binomial,data=kidfact)
summary(fmi)
summary(fcc)
```

Figure A.6: Output from aregImpute routine within the HMisc package

```
> summary(fmi)
Call:
fitter(formula = formula, family = "binomial", data = completed.data)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.3491   0.4531   0.5162   0.5793   2.3334

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)  2.737e+00  5.496e-02  49.793  < 2e-16 ***
AGE         -3.750e-02  3.014e-03 -12.443  < 2e-16 ***
FEMALE       9.931e-02  1.584e-02   6.269 3.64e-10 ***
ATYPE2       3.120e-01  1.813e-02  17.206  < 2e-16 ***
ATYPE3       3.760e-01  2.388e-02  15.743  < 2e-16 ***
NSEASON1    -6.399e-02  2.226e-02  -2.875  0.00404 **
NSEASON2    -5.019e-02  2.313e-02  -2.170  0.03002 *
NSEASON3    -4.734e-03  2.252e-02  -0.210  0.83349
LOS         -3.044e-03  7.058e-04  -4.312 1.62e-05 ***
TOTCHG      -4.510e-06  5.280e-07  -8.541  < 2e-16 ***
RACE2       -9.938e-02  2.195e-02  -4.529 5.94e-06 ***
RACE3       -1.437e-01  2.542e-02  -5.655 1.56e-08 ***
RACE4       -1.086e-01  3.476e-02  -3.125  0.00178 **
AWEEKEND    -2.030e-02  1.962e-02  -1.035  0.30077
NDX         -1.063e-01  3.723e-03 -28.553  < 2e-16 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 109941  on 134773  degrees of freedom
Residual deviance: 107907  on 134759  degrees of freedom
AIC: 107937

Number of Fisher Scoring iterations: 4
```

Figure A.7: Plots of missing value patterns using aregImpute and the Hmisc package

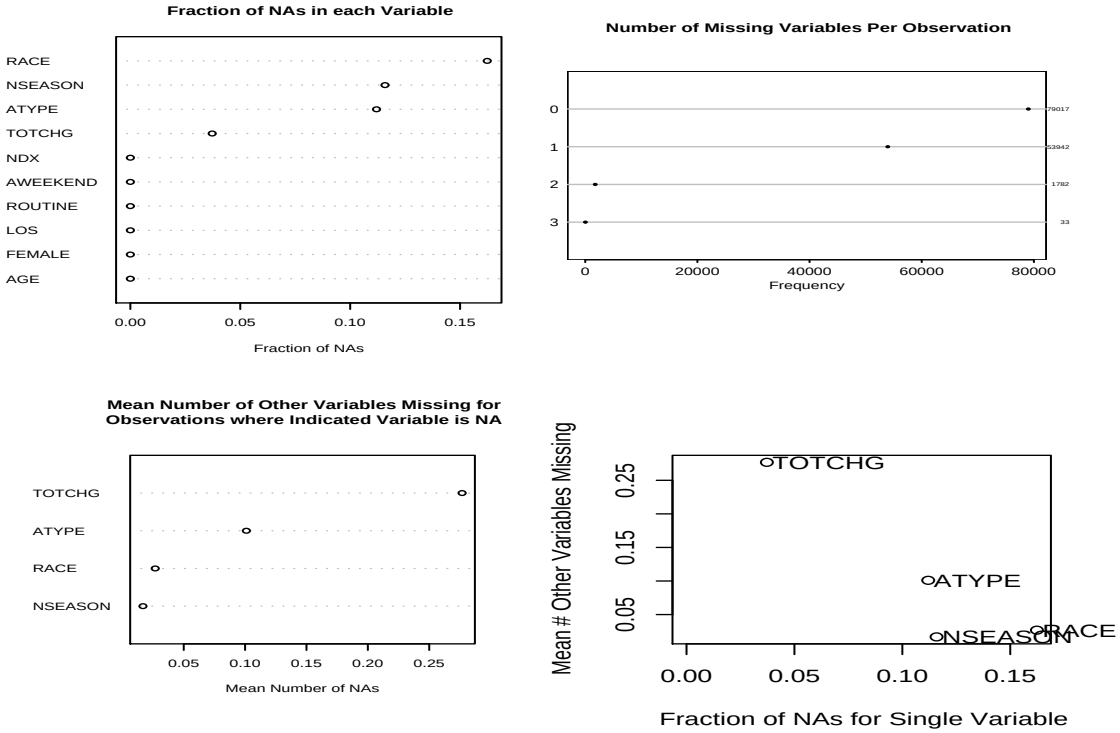


Figure A.8: Code to fit ICE within Stata

```
ssc install ice
set memory 64m
use kidsplus, clear
compress

generate byte nseas1=(nseason==1)
generate byte nseas2=(nseason==2)
generate byte nseas3=(nseason==3)
replace nseas1=. if nseason==.
replace nseas2=. if nseason==.
replace nseas3=. if nseason==.

generate byte a2=(atype==2)
generate byte a3=(atype==3)
replace a2=. if atype==.
replace a3=. if atype==.

generate byte r2=(race==2)
generate byte r3=(race==3)
generate byte r4=(race==4)
replace r2=. if race==.
replace r3=. if race==.
replace r4=. if race==.
replace totchg=. if totchg==.c

capture erase imputed.dta

#delimit ;
ice routine age nseason nseas1 nseas2 nseas3 aweekend race
  r2 r3 r4 female atype a2 a3 los totchg ndx using imputed,
  m(10) cmd(nseason:mlogit, race:mlogit, atype:mlogit, totchg:regress)
  passive(nseas1:nseason==1 \ nseas2:nseason==2 \ nseas3:nseason==3
  \ a2:atype==2 \ a3:atype==3 \ r2:race==2 \ r3:race==3
  \ r4:race==4) substitute(nseason:nseas1 nseas2 nseas3, race:r2 r3 r4,
  atype:a2 a3 ) ;
#delimit cr

use imputed, clear

micombine logistic routine age nseas1 nseas2 nseas3 aweekend
     r2 r3 r4 female a2 a3 los totchg ndx
```

Figure A.9: Output from ICE (missing values and prediction equations)

```
   #missing |
     values |      Freq.     Percent        Cum.
------------+-----------------------------------
          0 |     79,017       58.63       58.63
          1 |      3,662        2.72       61.35
          3 |     13,601       10.09       71.44
          4 |     37,892       28.12       99.55
          5 |        110        0.08       99.63
          7 |        246        0.18       99.82
          8 |        246        0.18      100.00
------------+-----------------------------------
      Total |    134,774      100.00

   Variable | Command | Prediction equation
------------+---------+-------------------------------------------------------
    routine |         | [No missing data in estimation sample]
        age |         | [No missing data in estimation sample]
    nseason | mlogit  | routine age aweekend r2 r3 r4 female a2 a3 los totchg
            |         | ndx
     nseas1 |         | [Passively imputed from nseason==1]
     nseas2 |         | [Passively imputed from nseason==2]
     nseas3 |         | [Passively imputed from nseason==3]
   aweekend |         | [No missing data in estimation sample]
       race | mlogit  | routine age nseas1 nseas2 nseas3 aweekend female a2 a3
            |         | los totchg ndx
         r2 |         | [Passively imputed from race==2]
         r3 |         | [Passively imputed from race==3]
         r4 |         | [Passively imputed from race==4]
     female |         | [No missing data in estimation sample]
      atype | mlogit  | routine age nseas1 nseas2 nseas3 aweekend r2 r3 r4
            |         | female los totchg ndx
         a2 |         | [Passively imputed from atype==2]
         a3 |         | [Passively imputed from atype==3]
        los |         | [No missing data in estimation sample]
     totchg | regress | routine age nseas1 nseas2 nseas3 aweekend r2 r3 r4
            |         | female a2 a3 los ndx
        ndx |         | [No missing data in estimation sample]
-------------------------------------------------------------------------------
Imputing 1..2..3..4..5..6..7..8..9..10..file imputed.dta saved
. use imputed, clear
```

Figure A.10: Output from ICE (micombine)

```
Multiple imputation parameter estimates (10 imputations)
------------------------------------------------------------------------------
     routine |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |  -.0374547   .0030187   -12.41   0.000    -.0433713   -.0315381
      nseas1 |  -.0725107   .0233962    -3.10   0.002    -.1183664    -.026655
      nseas2 |  -.0569509   .0248921    -2.29   0.022    -.1057385   -.0081633
      nseas3 |  -.0046107   .0252203    -0.18   0.855    -.0540415    .0448202
    aweekend |  -.0201544   .0196609    -1.03   0.305    -.0586891    .0183802
          r2 |  -.0821431   .0239809    -3.43   0.001    -.1291449   -.0351413
          r3 |  -.1457935   .0272348    -5.35   0.000    -.1991726   -.0924143
          r4 |  -.1024608    .035986    -2.85   0.004     -.172992   -.0319295
      female |   .0993698    .015858     6.27   0.000     .0682886     .130451
          a2 |   .3109377   .0205179    15.15   0.000     .2707234     .351152
          a3 |   .3816381   .0267408    14.27   0.000      .329227    .4340491
         los |  -.0032354   .0007087    -4.57   0.000    -.0046245   -.0018463
      totchg |  -4.34e-06   5.36e-07    -8.10   0.000    -5.40e-06   -3.29e-06
         ndx |  -.1061707   .0037324   -28.45   0.000    -.1134862   -.0988553
       _cons |   2.735714   .0558411    48.99   0.000     2.626267     2.84516
------------------------------------------------------------------------------
134774 observations.
```

Figure A.11: Code to fit iveWare

```
kid.sas:
options set = SRCLIB "." sasautos = ('!SRCLIB' sasautos)
mautosource;

options ls=80 nocenter;

libname mylib '.';
libname kid '..';
data one; set kid.kidsplus;
      routine = 1-routine;
run;
%regress(name=mysetup,dir=.);
run;

mysetup.set:
    datain work.one;
    mdata impute;
    iterations 10;
    multiples 10;
    seed 42;
    estout mylib.est;
    repout mylib.rep;
    link logistic;
    categorical atype nseason race;
    dependent routine;
    predictor age female los totchg ndx aweekend;
    estimates
         race1: race (1) /
         race2: race (0 1) /
         race3: race (0 0 1) /
         atype1: atype (1) /
         atype2: atype (0 1) /
         nseason1: nseason (1) /
         nseason2: nseason (0 1) /
         nseason3: nseason (0 0 1);
    print details;
 run;
```

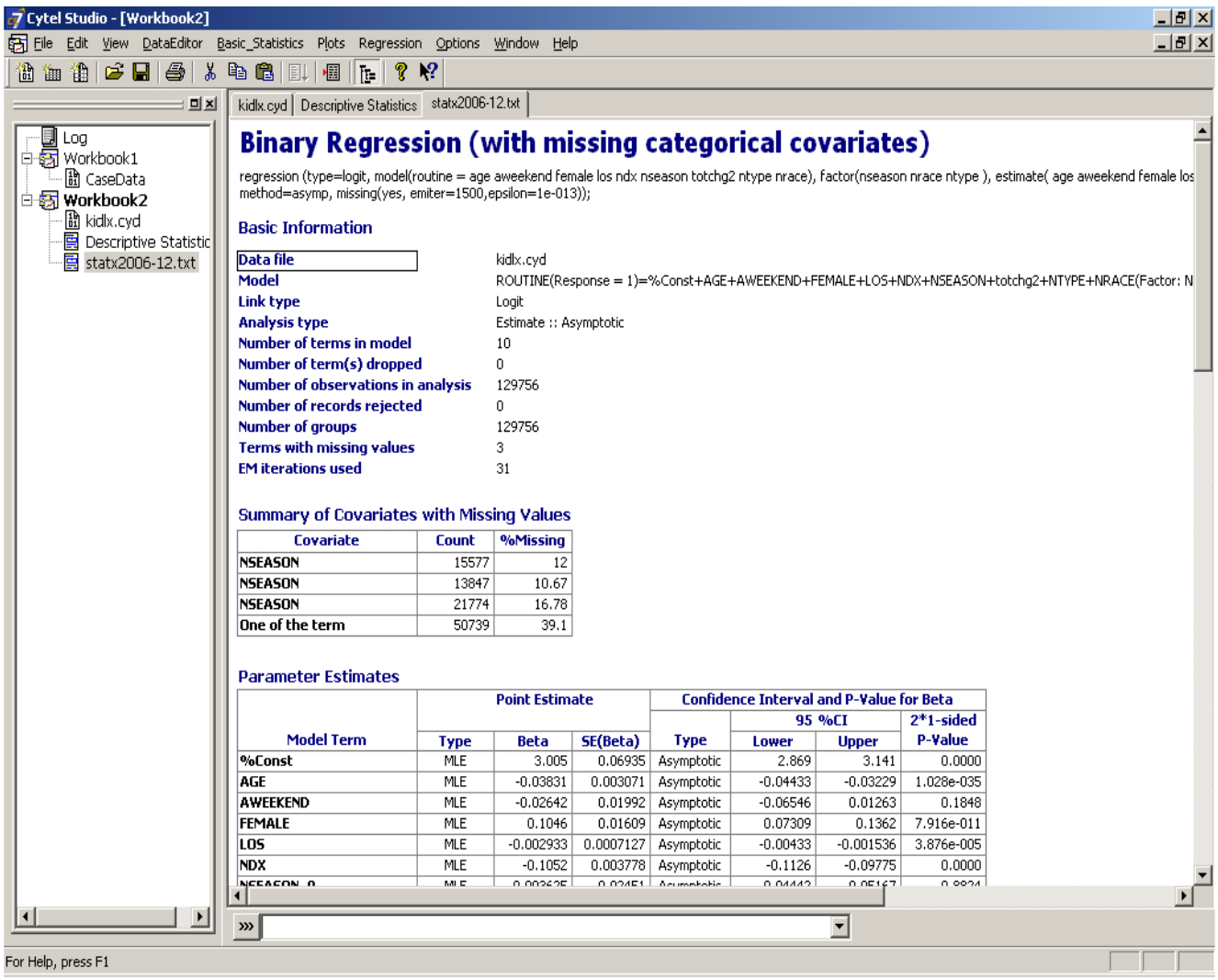
Figure A.12: Output of iveWare

```
(Results from separate imputations pruned)

All imputations
Valid cases              134774
Degr freedom        39.24283948
-2 LogLike          107917.8379

Variable   Estimate     Std Error  Wald test    Prob > Chi
Intercept  3.0025395   0.0681594  1940.55546   0.00000
AGE       -0.0375023   0.0030186   154.34713   0.00000
FEMALE     0.0997349   0.0158474    39.60739   0.00000
LOS       -0.0032319   0.0007054    20.98991   0.00000
TOTCHG    -0.0000043   0.0000005    66.43989   0.00000
NDX       -0.1063283   0.0037271   813.88402   0.00000
AWEEKEND  -0.0208605   0.0196453     1.12754   0.28830
RACE.1     0.1074252   0.0374141     8.24409   0.00409
RACE.2     0.0197358   0.0413594     0.22770   0.63324
RACE.3    -0.0443307   0.0438037     1.02421   0.31152
ATYPE.1   -0.3768551   0.0251527   224.48145   0.00000
ATYPE.2   -0.0697954   0.0268183     6.77318   0.00925
NSEASON.0  0.0076150   0.0236622     0.10357   0.74759
NSEASON.1 -0.0635770   0.0225465     7.95137   0.00481
NSEASON.2 -0.0515117   0.0229675     5.03018   0.02491
```

Figure A.13: Screenshot from LogXact



Cytel Studio - [Workbook2]

File Edit View DataEditor Basic_Statistics Plots Regression Options Window Help

Log
Workbook1
CaseData
Workbook2
kidlx.cyd
Descriptive Statistic
statx2006-12.txt

kidlx.cyd | Descriptive Statistics | statx2006-12.txt

Binary Regression (with missing categorical covariates)

regression (type=logit, model(routine = age aweekend female los ndx nseason totchg2 ntype nrace), factor(nseason nrace ntype), estimate(age aweekend female los method=asymp, missing(yes, emiter=1500,epsilon=1e-013));

Basic Information

Data file	kidlx.cyd
Model	ROUTINE(Response = 1)=%Const+AGE+AWEEKEND+FEMALE+LOS+NDX+NSEASON+totchg2+NTYPE+NRACE(Factor: N
Link type	Logit
Analysis type	Estimate :: Asymptotic
Number of terms in model	10
Number of term(s) dropped	0
Number of observations in analysis	129756
Number of records rejected	0
Number of groups	129756
Terms with missing values	3
EM iterations used	31

Summary of Covariates with Missing Values

Covariate	Count	%Missing
NSEASON	15577	12
NSEASON	13847	10.67
NSEASON	21774	16.78
One of the term	50739	39.1

Parameter Estimates

Model Term	Point Estimate			Confidence Interval and P-Value for Beta			
					95 %CI		2*1-sided
	Type	Beta	SE(Beta)	Type	Lower	Upper	P-Value
%Const	MLE	3.005	0.06935	Asymptotic	2.869	3.141	0.0000
AGE	MLE	-0.03831	0.003071	Asymptotic	-0.04433	-0.03229	1.028e-035
AWEEKEND	MLE	-0.02642	0.01992	Asymptotic	-0.06546	0.01263	0.1848
FEMALE	MLE	0.1046	0.01609	Asymptotic	0.07309	0.1362	7.916e-011
LOS	MLE	-0.002933	0.0007127	Asymptotic	-0.00433	-0.001536	3.876e-005
NDX	MLE	-0.1052	0.003778	Asymptotic	-0.1126	-0.09775	0.0000

For Help, press F1

Figure A.14: Code to fit MICE within R

```
library(mice)
source("patch1.14.R")
kidsplus <- read.csv("kidsplus.csv")
kidfact <- data.frame(kidsplus$age,as.factor(kidsplus$atype),
    as.factor(kidsplus$nseason),kidsplus$female,kidsplus$los,
    kidsplus$routine,kidsplus$totchg,as.factor(kidsplus$race),
    kidsplus$aweekend,kidsplus$ndx)
names(kidfact) <- c("AGE","ATYPE","NSEASON","FEMALE",
    "LOS","ROUTINE","TOTCHG","RACE","AWEEKEND","NDX")
imp <- mice(kidfact,im=c("","polyreg","polyreg","","","",
    "norm","polyreg","",""),m=10,seed=456)
fit <- glm.mids(ROUTINE ~ AGE + FEMALE + ATYPE + NSEASON +
    LOS + TOTCHG + RACE + AWEEKEND + NDX, family=binomial, data=imp)
result <- pool(fit)
fitcc <- glm(ROUTINE ~ AGE + FEMALE + ATYPE + NSEASON + LOS +
    TOTCHG + RACE + AWEEKEND + NDX, family=binomial, data=kidfact)
```

Figure A.15: Output from MICE (lower and upper 95% CI pruned)

```
> summary(result)
                      est           se           t          df
(Intercept)  2.737864e+00 5.573048e-02  49.1268757  12724.5089
AGE         -3.749941e-02 3.016851e-03 -12.4299863 131963.1081
FEMALE       9.967850e-02 1.585490e-02   6.2869210 127425.0757
ATYPE2       3.065392e-01 1.937845e-02  15.8185626    639.9450
ATYPE3       3.772982e-01 2.468176e-02  15.2865201   1700.4235
NSEASON1    -7.167735e-02 2.377789e-02  -3.0144540    575.4299
NSEASON2    -5.323242e-02 2.433102e-02  -2.1878418    952.7860
NSEASON3    -6.745087e-03 2.461794e-02  -0.2739908    352.6997
LOS         -3.237787e-03 7.064497e-04  -4.5831818  36101.4192
TOTCHG      -4.339639e-06 5.321358e-07  -8.1551359   9218.3157
RACE2       -8.378841e-02 2.347370e-02  -3.5694588    621.0766
RACE3       -1.437074e-01 2.800739e-02  -5.1310544    296.0473
RACE4       -1.039486e-01 3.726879e-02  -2.7891605    487.0386
AWEEKEND    -2.083815e-02 1.963036e-02  -1.0615267 133735.5810
NDX         -1.061728e-01 3.726475e-03 -28.4914839 130256.5639
                Pr(>|t|)  missing          fmi
(Intercept) 0.000000e+00       NA 0.0252771848
AGE         0.000000e+00        0 0.0011595282
FEMALE      3.248677e-10        0 0.0019303935
ATYPE2      0.000000e+00       NA 0.1182741633
ATYPE3      0.000000e+00       NA 0.0722584557
NSEASON1    2.687842e-03       NA 0.1247599341
NSEASON2    2.892344e-02       NA 0.0968126464
NSEASON3    7.842521e-01       NA 0.1594965626
LOS         4.594857e-06        0 0.0134789554
TOTCHG      4.440892e-16     5018 0.0301271524
RACE2       3.853390e-04       NA 0.1200662185
RACE3       5.222446e-07       NA 0.1741291410
RACE4       5.491480e-03       NA 0.1356565986
AWEEKEND    2.884525e-01        0 0.0006851676
NDX         0.000000e+00        0 0.0014892814
```

Figure A.16: Description of code to fit missing data models within SAS PROC/MI using two step imputation

```
proc mi nimpute=20 data = tc out = tci;
    class aweekend female ;
    var los age ndx routine aweekend female totchg;
    monotone reg;
run;
proc print data = tci (obs = 10); run;
data back1; set tci rem1 (in = dups);
    impnum = 20;
    if dups then do i = 1 to impnum;
         _imputation_ = i; output;
    end;
    else output;
run;

data patts; set back1;
misspatt = ((newatype eq .) + 1) + 10 * ((race eq .) + 1) + 100 * ((nseason eq .) +1);
run;

proc freq data = patts;
    tables misspatt;
run;

proc sort data = patts; by _imputation_; run;
data k.patts; set patts; run;

proc mi nimpute=1 data = patts out = i_srt;
    where misspatt = 111 or misspatt = 112 or misspatt = 122;
    by _imputation_;
    class routine aweekend female nseason race newatype;
    var los age ndx totchg routine aweekend female nseason race newatype;
    monotone propensity (/ngroups = 20) ;
run;
proc mi nimpute=1 data = patts out = i_tsr;
    where misspatt = 111 or misspatt = 121 or misspatt = 221;
    by _imputation_;
    class routine aweekend female newatype nseason race;
    var los age ndx totchg routine aweekend female newatype nseason race;
    monotone discrim ;
run;
proc mi nimpute=1 data = patts out = i_rts;
    where misspatt = 111 or misspatt = 211 or misspatt = 212;
    by _imputation_;
    class aweekend female race newatype nseason;
    var los age ndx routine aweekend female totchg race newatype nseason;
    monotone discrim ;
run;

data mis; set i_srt i_tsr i_rts;
    where misspatt ne 111;
run;

data k.whole; set patts (where = (misspatt eq 111)) mis; run;

proc sort data = k.whole; by _imputation_; run;

proc logistic data = k.whole descending ;
    by _imputation_;
    class routine race aweekend nseason female newatype/param = glm;
    model routine = age race aweekend nseason female newatype los totchg ndx/covb;
    ods output ParameterEstimates=kidsparms CovB=kidsscovb;
run;

proc print data = kidsscovb (obs = 2); run;
proc print data = kidsparms (obs = 20); run;

data kp2; set kidsparms; if df ne 0;
if classval0 ne '' then variable = compress(variable||classval0);
run;

proc print data = kp2 (obs = 20); run;

proc mianalyze parms=kp2
    covb(effectvar=stacking)=kidsscovb;
    modeleffects Intercept  age race1 race2 race3 aweekend0 nseason1
    nseason2 nseason3 female0 newatype1 newatype2 los totchg ndx;
run;
```

Figure A.17: Output from SAS PROC MI using two stage imputation

The MIANALYZE Procedure Model Information

PARMS Data Set	WORK.KP2
COVB Data Set	WORK.KIDSSCO
Number of Imputations	20

Multiple Imputation Variance Information

Parameter	Variance Between	Variance Within	Variance Total	DF	Relative Increase in Var.	Fraction Missing Info.	Relative Efficiency
Intercept	0.000461	0.005616	0.006100	3014.1	0.086243	0.080006	0.996016
age	1.3318326E-8	0.000010474	0.000010488	1.07E7	0.001335	0.001334	0.999933
race1	0.000348	0.001466	0.001832	476.59	0.249477	0.203003	0.989952
race2	0.000407	0.001793	0.002220	512.86	0.238353	0.195607	0.990314
race3	0.000479	0.002157	0.002660	531.39	0.233185	0.192126	0.990485
aweekend0	4.4236506E-8	0.000449	0.000449	1.78E9	0.000103	0.000103	0.999995
nseason1	0.000104	0.000587	0.000697	770.49	0.186287	0.159213	0.992102
nseason2	0.000098719	0.000538	0.000642	728.92	0.192535	0.163741	0.991879
nseason3	0.000055963	0.000589	0.000647	2305.4	0.099847	0.091571	0.995442
female0	5.4249308E-8	0.000291	0.000291	4.95E8	0.000196	0.000196	0.999990
newatype1	0.000000140	0.000655	0.000655	3.76E8	0.000225	0.000225	0.999989
newatype2	0.000000124	0.000773	0.000773	6.75E8	0.000168	0.000168	0.999992
los	4.6712575E-9	0.000000602	0.000000607	291101	0.008145	0.008086	0.999596
totchg	5.184886E-15	4.031789E-13	4.086231E-13	107039	0.013503	0.013342	0.999333
ndx	1.5385562E-8	0.000016604	0.000016620	2.01E7	0.000973	0.000972	0.999951

Multiple Imputation Parameter Estimates

Parameter	Estimate	Std Error	95% Confidence Limits		DF	Minimum	Maximum
Intercept	3.104038	0.078103	2.95090	3.25718	3014.1	3.073192	3.152175
age	-0.036523	0.003239	-0.04287	-0.03018	1.07E7	-0.036751	-0.036304
race1	0.110903	0.042798	0.02681	0.19500	476.59	0.073480	0.137640
race2	0.042818	0.047118	-0.04975	0.13539	512.86	-0.004068	0.071938
race3	-0.070233	0.051579	-0.17156	0.03109	531.39	-0.130730	-0.043424
aweekend0	0.036344	0.021197	-0.00520	0.07789	1.78E9	0.036047	0.037051
nseason1	-0.011170	0.026393	-0.06298	0.04064	770.49	-0.029910	0.007638
nseason2	-0.072749	0.025338	-0.12249	-0.02300	728.92	-0.091803	-0.057172
nseason3	-0.052218	0.025442	-0.10211	-0.00233	2305.4	-0.064840	-0.039615
female0	-0.118519	0.017052	-0.15194	-0.08510	4.95E8	-0.119095	-0.118150
newatype1	-0.378938	0.025594	-0.42910	-0.32877	3.76E8	-0.379799	-0.378393
newatype2	-0.066875	0.027804	-0.12137	-0.01238	6.75E8	-0.067662	-0.066197
los	-0.003867	0.000779	-0.00539	-0.00234	291101	-0.003961	-0.003719
totchg	-0.000003171	0.000000639	-0.00000	-0.00000	107039	-0.000003	-0.000003
ndx	-0.111112	0.004077	-0.11910	-0.10312	2.01E7	-0.111329	-0.110907

Figure A.18: Description of code to fit models within S-Plus missing data library

```
library(missing)
options(contrasts=c("contr.treatment", "contr.poly"))
margins.form <- ~ ATYPE + NSEASON + RACE + AWEEKEND + ROUTINE + FEMALE +
    ATYPE*NSEASON + ATYPE+RACE + ATYPE*AWEEKEND + ATYPE*ROUTINE + ATYPE*FEMALE +
    NSEASON*RACE + NSEASON*AWEEKEND + NSEASON*ROUTINE + NSEASON*FEMALE +
    RACE*AWEEKEND + RACE*ROUTINE + RACE*FEMALE +
    AWEEKEND*ROUTINE + AWEEKEND*FEMALE + ROUTINE*FEMALE
kidfact <- data.frame(kidsplus$age,as.factor(kidsplus$atype),
   as.factor(kidsplus$nseason),as.factor(kidsplus$female),
   kidsplus$los,as.factor(kidsplus$routine),
   kidsplus$totchg,as.factor(kidsplus$race),
   as.factor(kidsplus$aweekend),kidsplus$ndx)
names(kidfact) <- c("AGE","ATYPE","NSEASON","FEMALE","LOS","ROUTINE","TOTCHG",
    "RACE","AWEEKEND","NDX")
ccglm <- glm(ROUTINE ~ AGE + ATYPE + AWEEKEND + FEMALE + LOS + NDX +
   RACE + TOTCHG + NSEASON, na.action=na.omit,
   family=binomial, data=kidfact)
kidcgm <- preCgm(kidfact)
kid.em <- emCgm(kidcgm, margins=margins.form, design = margins.form, prior=1.05)
dataDepend <- dataDepPrior(kidcgm, nPriorObs=50, algorithm="da")
impout <- impCgm(kid.em,nimpute=10)
impout1 <- miSubscript(impout, 1)
impglm <- miEval(glm(ROUTINE ~ AGE + ATYPE + AWEEKEND + FEMALE + LOS +
    NDX + RACE + TOTCHG + NSEASON,
    family=binomial, data=impout, subset =
    !is.infinite(impout1$TOTCHG)))
sumfit <- miEval(summary(impglm))
coefs <- miEval(coef(sumfit))
results <- miMeanSE(miEval(coefs[, 1]), miEval(coefs[, 2]))
```

Figure A.19: Description of output from S-Plus missing data library

```
> results
$est:
 (Intercept)         AGE    ATYPE2    ATYPE3    AWEEKEND     FEMALE
    2.734479 -0.03740613 0.3137568 0.3822975 -0.01846629 0.09821091
          LOS        NDX       RACE2      RACE3      RACE4        TOTCHG
 -0.003195076 -0.1059842 -0.07862305 -0.1493073 -0.1029841 -4.403978e-06
    NSEASON2   NSEASON3     NSEASON4
 -0.07259462 -0.0558042 -0.009378634

$std.err:
 (Intercept)         AGE     ATYPE2     ATYPE3   AWEEKEND     FEMALE
  0.05553486 0.003019782 0.02081542 0.02508927 0.01964768 0.01585714
          LOS         NDX      RACE2      RACE3      RACE4       TOTCHG
 0.0007103084 0.003731087 0.02261825 0.02705698 0.03702192 5.385703e-07
   NSEASON2   NSEASON3  NSEASON4
 0.02408424 0.02415537 0.0246976

$df:
 (Intercept) AGE ATYPE2 ATYPE3 AWEEKEND FEMALE LOS NDX RACE2 RACE3 RACE4
          NA  NA     NA     NA       NA     NA  NA  NA    NA    NA    NA
 TOTCHG NSEASON2 NSEASON3 NSEASON4
     NA       NA       NA       NA

$m:
[1] 10

$r:
 (Intercept)         AGE    ATYPE2    ATYPE3    AWEEKEND      FEMALE       LOS
   0.0188569 0.002802194 0.3168717 0.1091718 0.001285257 0.001589901 0.0211495
         NDX      RACE2     RACE3     RACE4     TOTCHG  NSEASON2   NSEASON3
 0.003671249 0.05358623 0.1336732 0.1393954 0.05036733 0.1706005 0.08983401
  NSEASON4
 0.1968896

$fminf:
 (Intercept)         AGE    ATYPE2    ATYPE3    AWEEKEND      FEMALE       LOS
  0.01858261 0.002796094 0.2502102 0.1003611 0.001283973 0.001587936 0.0208048
         NDX      RACE2     RACE3     RACE4     TOTCHG  NSEASON2   NSEASON3
 0.003660782 0.05140594 0.1206243 0.1252462 0.04843821 0.1497413 0.08381139
  NSEASON4
 0.1694804
```