# Sullivan's Method with Covariates: A Bayesian Approach for Obtaining Interval Estimates of HLE for Specific Subpopulations

Scott M. Lynch and J. Scott Brown
Princeton University · Miami University

May 29, 2006

This research was supported by NICHD grant R03HD050374-01.

# Introduction

- Sullivan’s Method is most common method for estimating HLE/ALE, but method is limited
- Covariates can only be included via disaggregation/aggregation
- usually limited to very few covariates (mortality file)
- S.E.’s too small—prevalence-based method attempting to capture incidence process

# Cont’d

Goal: Describe method that:

Allows inclusion of covariates measured at different levels in health and mortality files

Allows construction of interval estimates

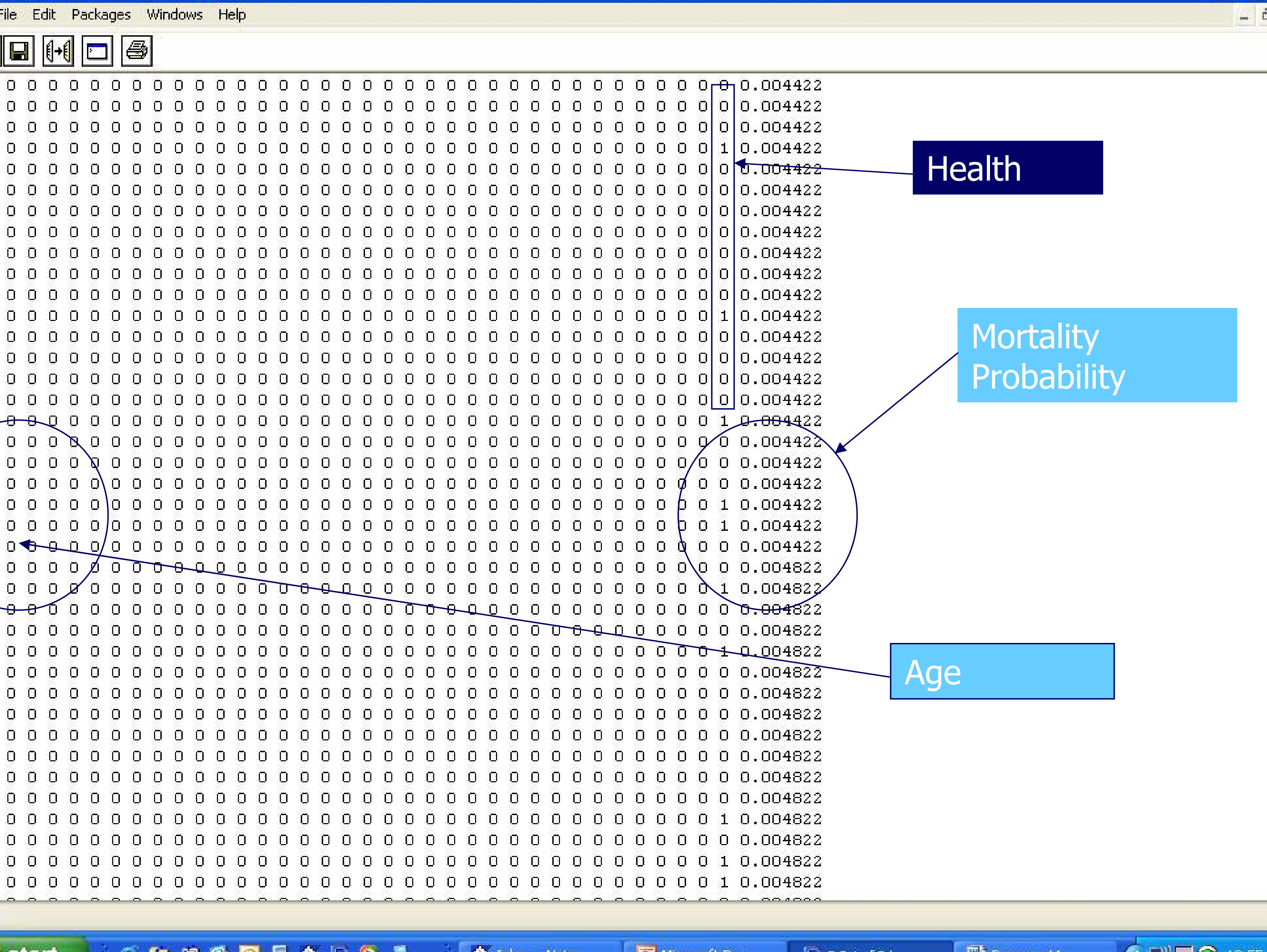
Question: To what extent does SES explain black-white disparities in HLE in the US, and how has this changed over time?

# A New Method

1. Structure Data appropriately
2. “hazard” model w/ Gibbs sampling
3. Generate distributions of transition probability matrices using the Gibbs parameter samples + covariate profile
4. Compute life tables for each transition probability matrix
5. Summarize results

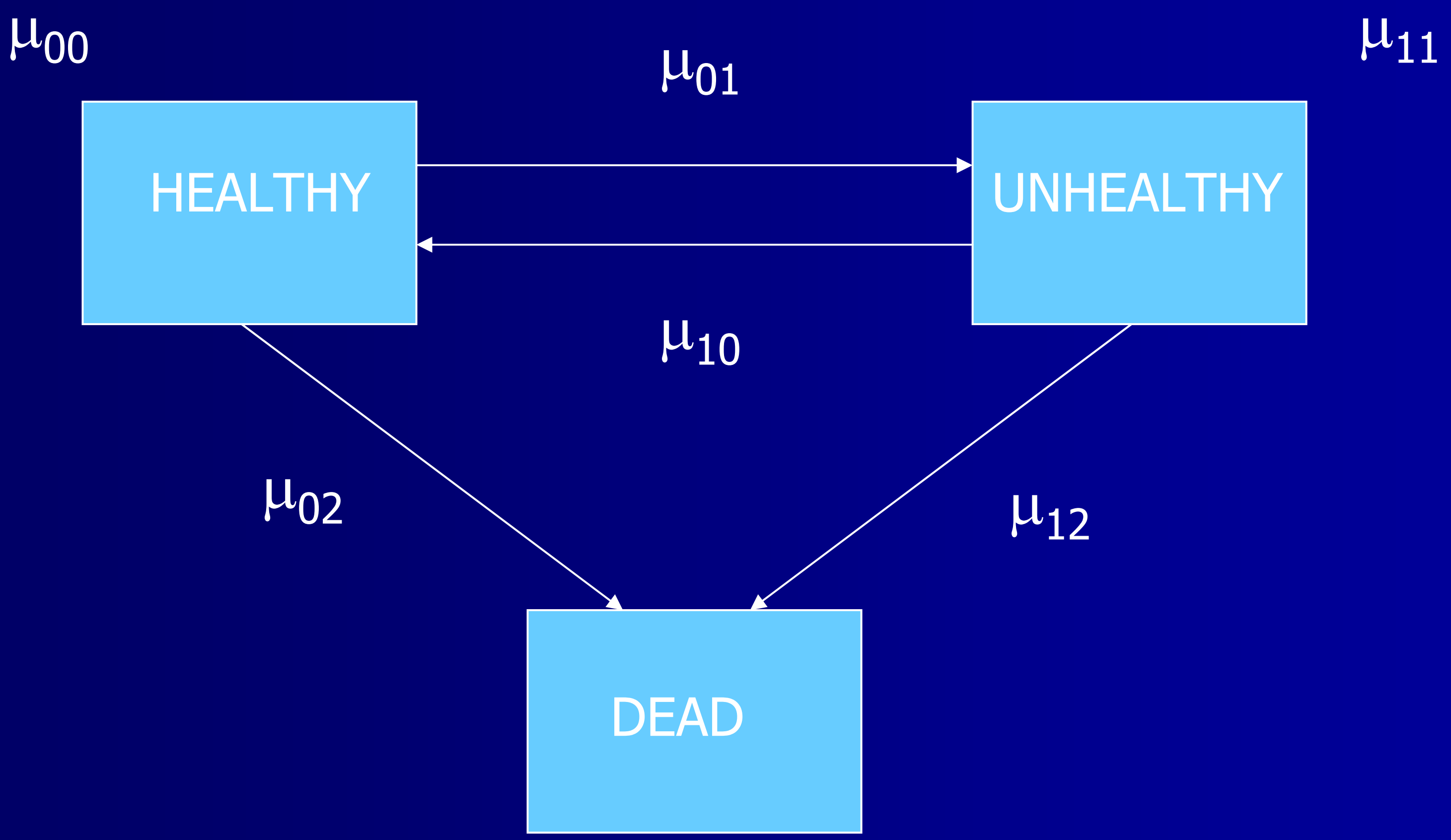
# Data for "Hazard" Model

- Data are cross-sectional mortality data measured on covariates X(m) and individual-level health data measured on X(h).
  - Length(X(h))>Length(X(m)) [more refined covariate space for health file]
- Merge mortality probabilities by X(m) into health file—one-to-many merge



File Edit Packages Windows Help
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.004422
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.004422
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.004422
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0.004422
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.004422
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.004422
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.004422
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.004422
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.004422
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.004422
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.004422
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0.004422
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.004422
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.004422
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.004422
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.004422
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0.004422
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.004422
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.004422
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.004422
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0.004422
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0.004422
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.004422
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.004822
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0.004822
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.004822
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.004822
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0.004822
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.004822
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.004822
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.004822
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.004822
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.004822
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.004822
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0.004822
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.004822
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0.004822
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0.004822
Health
Mortality Probability
Age

# Multistate Transitions



# Bivariate 'Outcome' Space

| Death → <br> Health <br> ↓ | Alive (0) | Dead (1) |
|---|---|---|
| Healthy (0) | Healthy & Alive | Healthy & Dead |
| Unhealthy (1) | Unhealthy & Alive | Unhealthy & Dead |

# Bivariate Probit Model

- Hazard model for bivariate outcome state space is discrete time bivariate probit:

$$p(y=[h\ \ k])=\Phi_2(\tau_{1-h},t_{2-h};t_{1-k},t_{2-k})$$

- Where, h and k are 0,1 two-dimensional outcomes, $\tau_0$=-∞, $\tau_1$=Xβ, and $\tau_2$=∞ (in each dimension)
- Age is key covariate

File Edit Packages Windows Help

```
#FILL IN THIS INFORMATION
#note: error correlation assumed 0--sigma is identity

maxiter= 10000                #number of iterations
numx=2                        #number of x vars in y eq. (incl. intercept)
numxz=2                       #number of x vars in z eq.
numy=2                        #number of y vars--don't adjust
inf="c:\\bushsucks.dat"       #put the dataset names here
outf="c:\\bushsucks.out"

#DON'T ADJUST BELOW HERE
g<-read.table(inf,header=F)
mnorm<-function(covmat){return((rnorm(length(diag(covmat)),mean=0,sd=1))%*%chol(covmat))}
x<-as.matrix(g[,1:numx]); y<-as.matrix(g[,(numx+1):(numx+numy)])
vb<-solve(t(x)%*%x)
b<-matrix(0,numx,numy); z<-matrix(NA,length(x[,1]),numy)
z[,numy]<-qnorm(y[,2],mean=0,sd=1)

for(i in 1:maxiter){
u<-as.matrix(runif(length(x[,1]),min=0,max=1))
xb=as.matrix((x%*%b)[,1])
z[,1]=qnorm(y[,1]*u + u*(-1)^y[,1]*pnorm(0,mean=xb,sd=x[,1])
          + y[,1]*pnorm(0,mean=xb,sd=x[,1]),mean=xb,sd=x[,1])
b<-vb%*%(t(x)%*%z) + cbind(t(mnorm(vb)),t(mnorm(vb)))
if(numxz<numx){b[numxz+1:numx,numy]=0}

if(i%%(numx*numy)==0){print(c(i, "b0h=",round(b[1,1],digits=4), "b0m=",round(b[1,2],digits=4)))
write(c(i,0,c(t(b))),file=outf,append=T,ncolumns=numx*numy+2)}
}
```

start | Inbox - Netsc... | Microsoft Pow... | RGui - [C:\Doc... | Search Rutger... | 12:20 AM

# Hazard Model: Gibbs Sampler

1. Simulate Latent Data, $Z \mid \beta$

   a. $Z(h) \sim N[\ X(h)^T \beta(h)\ ,\ 1\ ]$

   b. $Z(d) = \Phi^{-1}(\text{mortality prob})$

2. Simulate $\beta \mid Z$

   $$\beta \sim N[\ (X^T X)^{-1}(X^T Z)\ ,\ (X^T X)^{-1}\ ]$$

   (Note: b for unbalanced covariates set to 0)

3. Repeat

# Life Table Construction

1. Select Covariate Profile *(X)* and compute p(dead), p(unhealthy), and p(healthy) at age x $(\forall x)$

$$p(d)=\Phi(X'\beta(d))$$

$$p(u)=(1-p(d))\Phi(X'\beta(h))$$

$$p(h)=1-(p(d)+p(u))$$

# Table Construction, cont'd

2. Now we have states at start and end of each age interval, but this doesn't give us *transition* probabilities!

   We must estimate them…

# Ecological Inference 1

| | | | |
|---|---|---|---|
| | | | R $[p_h(a)]$ |
| | | | 1-R $[p_{uh}(a)]$ |
| C $[P_h(a+1)]$ | 1-C-M $[p_{uh}(a+1)]$ | M $[P_d(a+1)]$ | 1 |

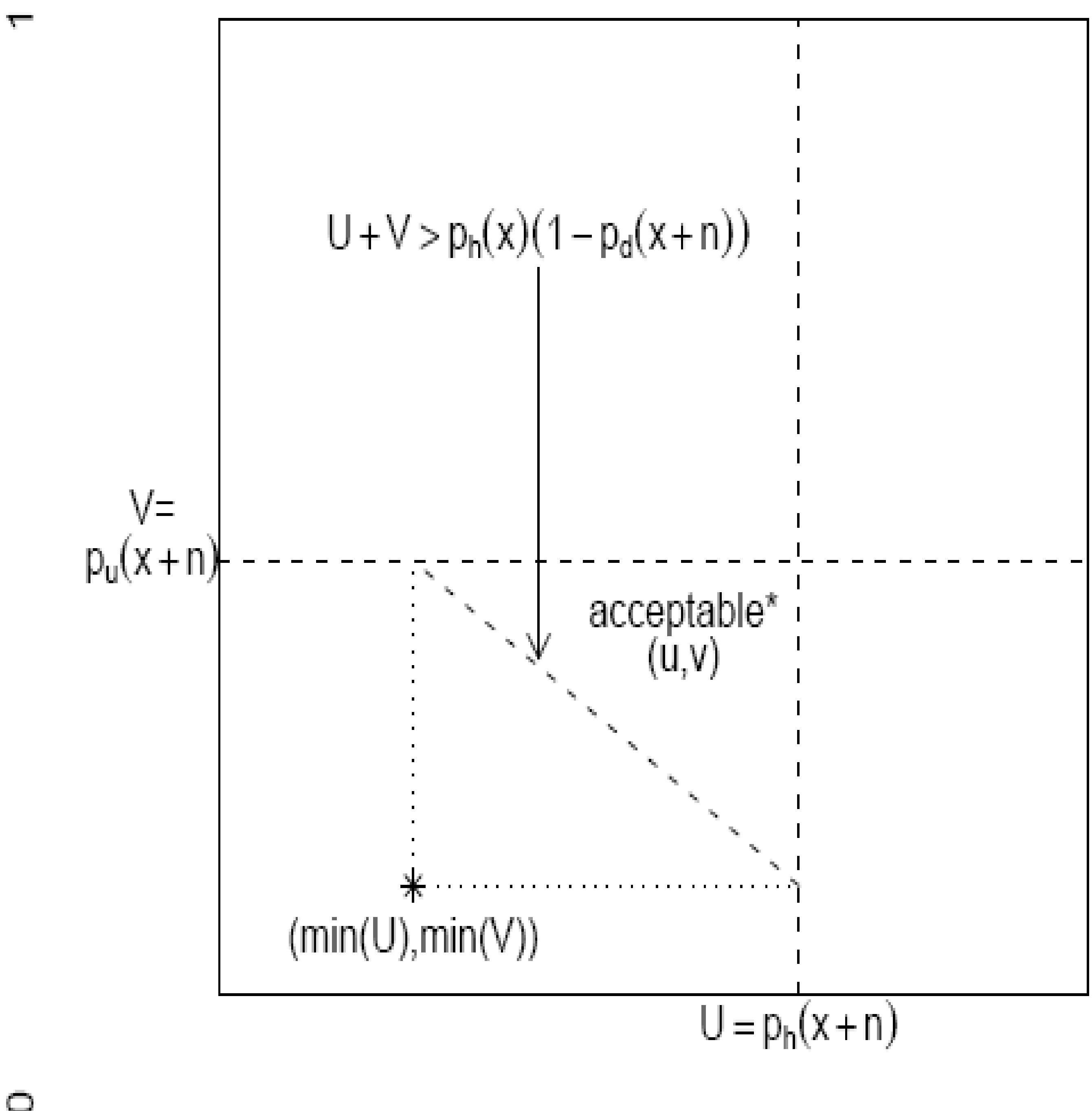
# Ecological Inference 2

| | | | |
|---|---|---|---|
| X | Y | R-(X+Y) | **R [$p_h(a)$]** |
| C-X | 1-C-M-Y | M-R+X+Y | **1-R [$p_{uh}(a)$]** |
| **C [$P_h(a+1)$]** | **1-C-M [$p_{uh}(a+1)$]** | **M [$P_d(a+1)$]** | **1** |

# Ecol. Inference 2, cont'd

- $X+Y \leq R$
- $X \leq C$
- $Y \leq 1-C-M$
- $X+Y \geq M-R$
- $1-X/R<(1-C-M-Y)/(1-R)$ (embeddability)
- $X+Y>R(1-M)$ (mortality constraint)

| | | | |
|---|---|---|---|
| X | Y | R-(X+Y) | **R** **[$p_h(a)$]** |
| C-X | 1-C-M-Y | M-R+X+Y | **1-R** **[$p_{uh}(a)$]** |
| **C** **[$P_h(a+1)$]** | **1-C-M** **[$p_{uh}(a+1)$]** | **M** **[$P_d(a+1)$]** | **1** |


1
U + V > p_h(x)(1 − p_d(x + n))
V=
p_u(x + n)
acceptable*
(u,v)
(min(U),min(V))
U = p_h(x + n)
0

# Table Construction, cont'd

3. Must transform matrix of transition probabilities *(P)* into matrix of hazards, M:

$$\begin{bmatrix} \sum \mu_{h.} & -\mu_{hu} & -\mu_{hd} \\ -\mu_{uh} & \sum \mu_{u.} & -\mu_{ud} \\ 0 & 0 & 0 \end{bmatrix}$$

# Table Construction, cont'd

In a continuous time Markov process, *P=exp(M)*, so *M=ln(P)*. This can be done using the infinite series representation of the ln() function:

$$\ln(P) = \sum_{i=1}^{\infty} \frac{(-1)^{i-1}(P-I)^{i}}{i}$$

# Table Construction, cont'd

4. Given M, we need to compute l(x) :

   l(x+n)=l(x)exp{-nM(x)}, where

$$\exp\left\{-nM(x)\right\} = I + \sum_{i=1}^{\infty} \frac{(-1)^i n^i M^i(x)}{i!}$$

# Table Construction, cont'd

5. Then, compute L(x) (= ∫ l(x)):

$$L(x) = nl(x)\left[ I + \sum_{i=1}^{\infty} \frac{(-1)^i n^i M^i(x)}{(i+1)!} \right]$$

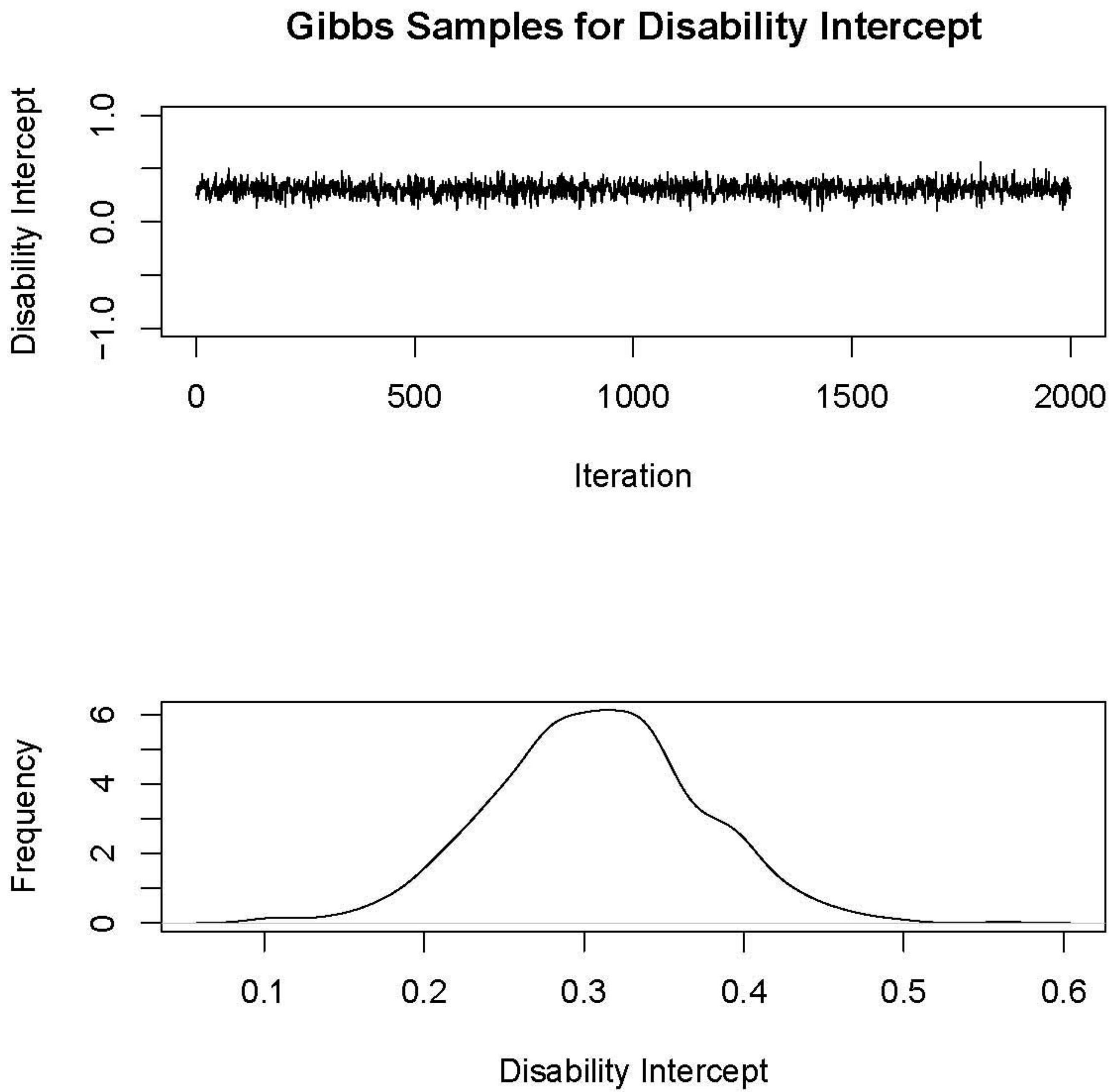
# Table Construction, cont'd

6. Finally, compute state expectancies (note: these are all diagonal matrices):

$$e(x)=(\Sigma L(x))\ l(x)^{-1}$$

# Example (Role of SES in Explaining black-white disparities in HLE)

- NHIS 1982-2002 data + NCHS mortality data
- NHIS data: age, male, black, south, education, income, health (dichotomous)
- NCHS data: age, male, black, mortality probability


Gibbs Samples for Disability Intercept
Disability Intercept
1.0
0.0
-1.0
0
500
1000
1500
2000
Iteration
Frequency
6
4
2
0
0.1
0.2
0.3
0.4
0.5
0.6
Disability Intercept

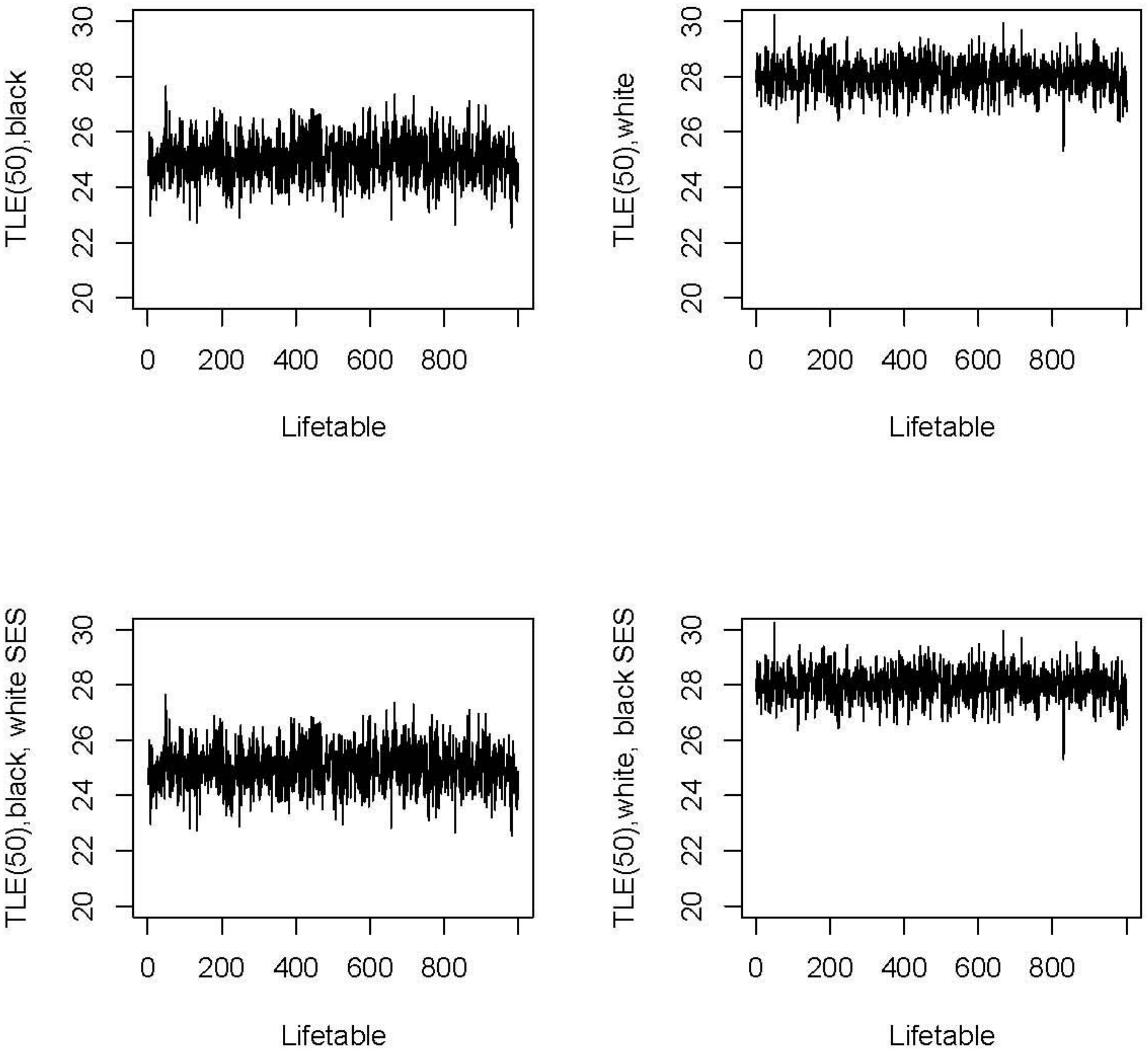
# Model Results: Health

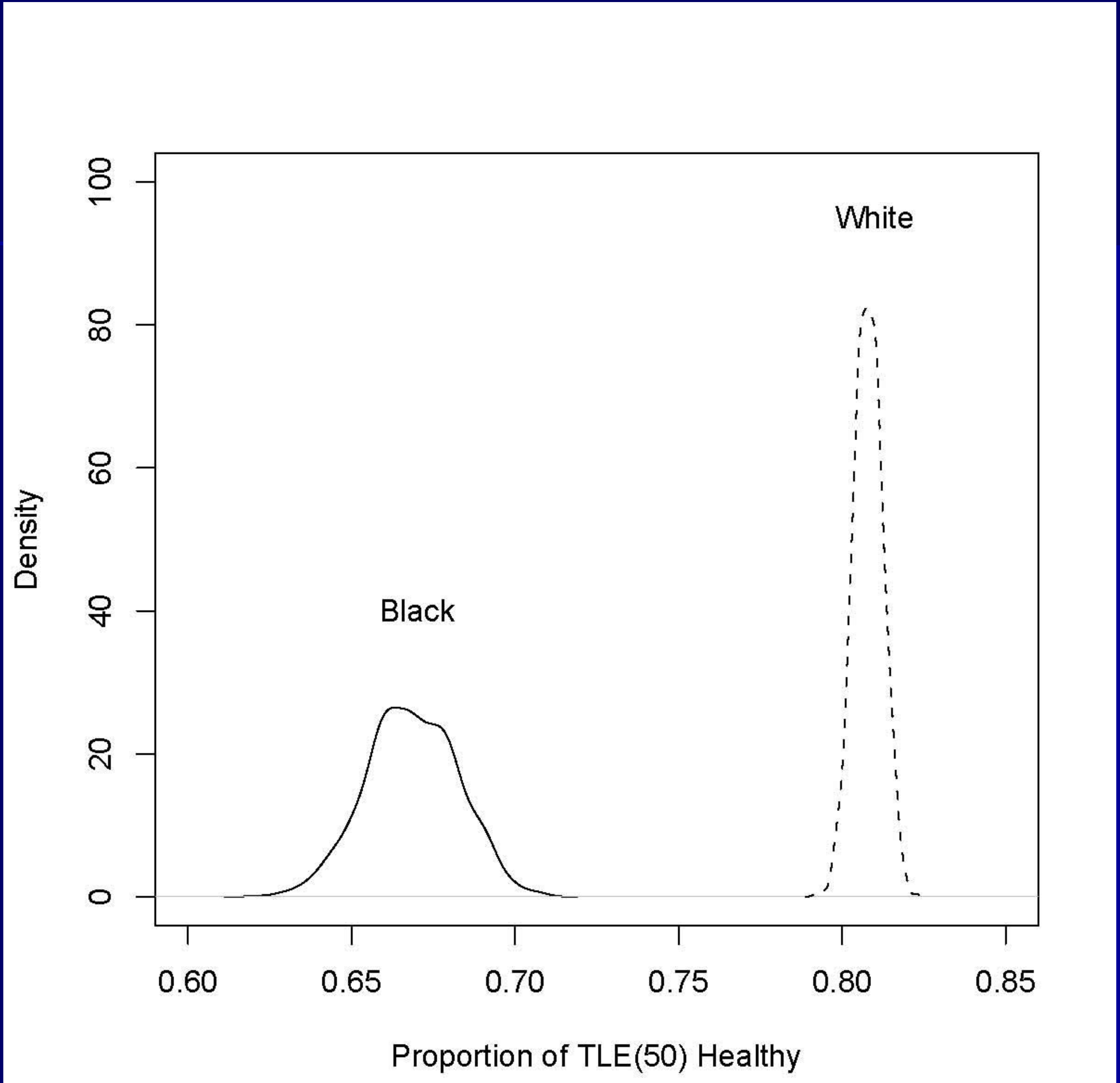
| Variable | MLE | Gibbs p.m. |
|---|---|---|
| Intercept | .306(.067) | .308(.065) |
| Age | .0038(.001) | .0038(.001) |
| Male | .098(.029) | .097(.029) |
| Black | .243(.041) | .242(.042) |
| South | .149(.029) | .150(.029) |
| Education | -.066(.005) | -.066(.005) |
| Income | -.013(.0007) | -.013(.0007) |

# Model Results: Death

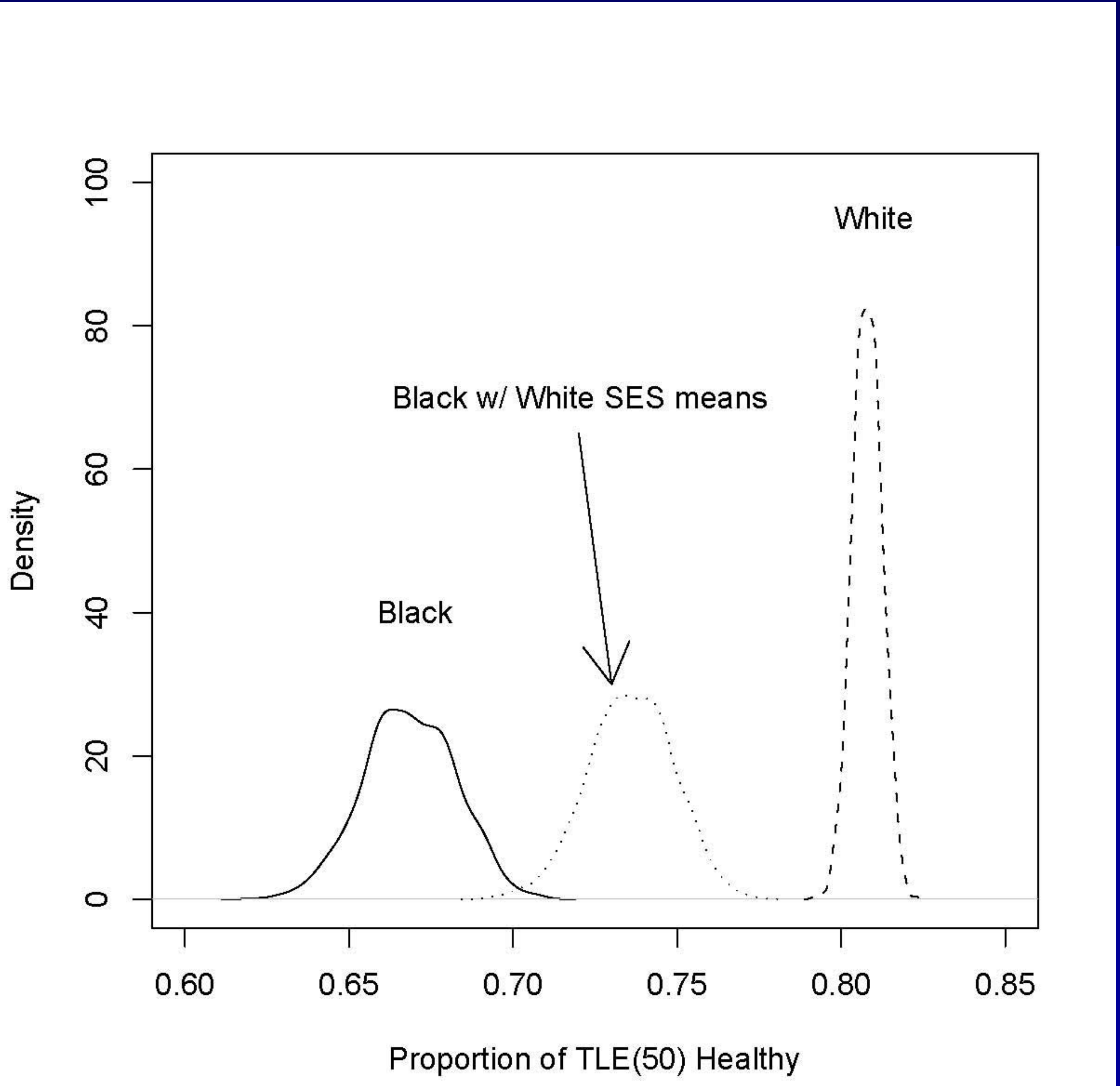
| Var | MLE | Gibbs p.m. |
|---|---|---|
| Intercept | -2.82(.021) | -2.83(.046) |
| Age | .038(.001) | .038(.001) |
| Male | .188(.02) | .188(.02) |
| Black | .181(.03) | .182(.03) |

Note: ML s.e.'s divided by MSE for comparison


TLE(50), black
TLE(50), white
TLE(50), black, white SES
TLE(50), white, black SES
20
22
24
26
28
30
0
200
400
600
800
Lifetable



White
Black
Density
100
80
60
40
20
0
0.60
0.65
0.70
0.75
0.80
0.85
Proportion of TLE(50) Healthy

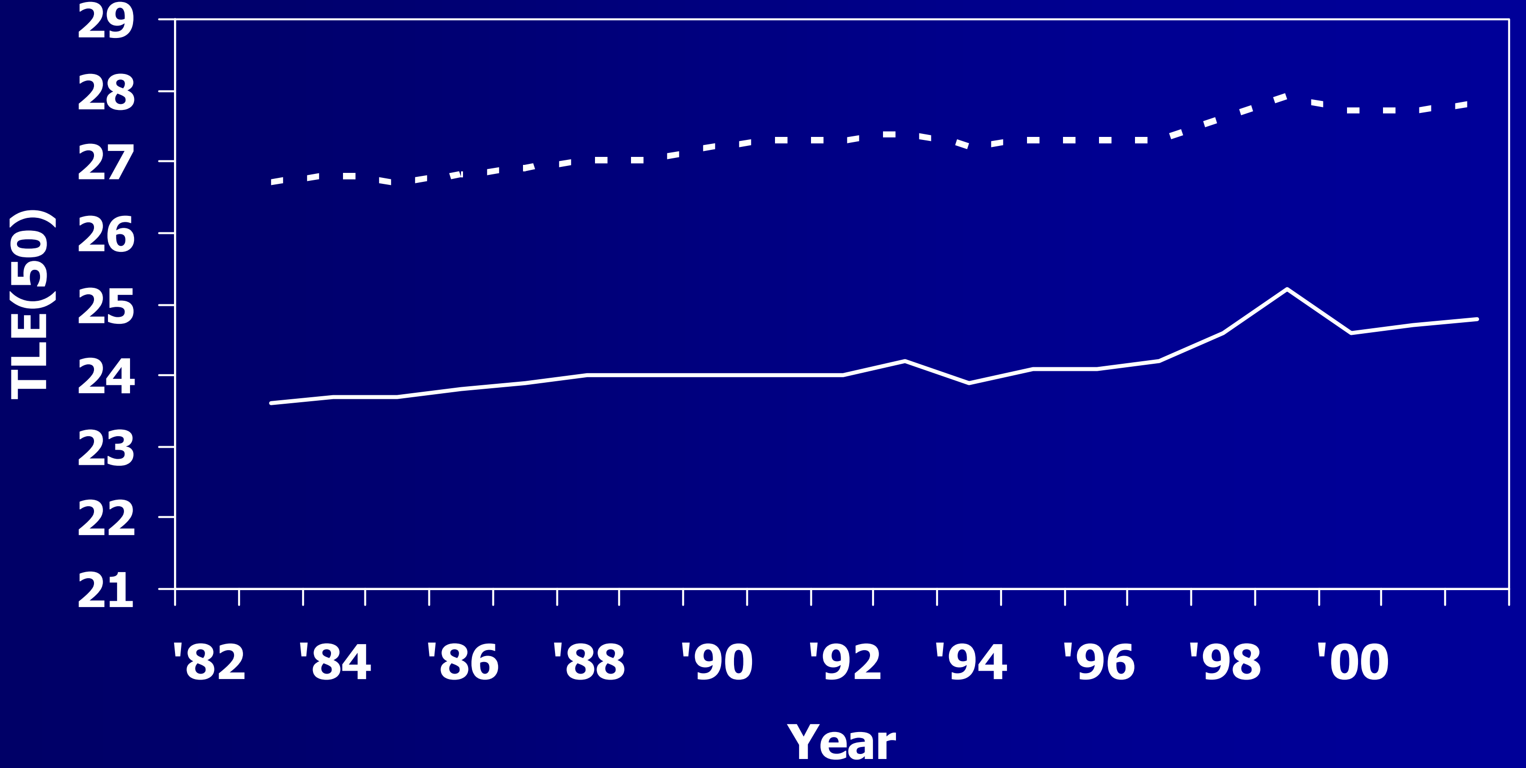


White
Black w/ White SES means
Black
Density
0
20
40
60
80
100
0.60
0.65
0.70
0.75
0.80
0.85
Proportion of TLE(50) Healthy

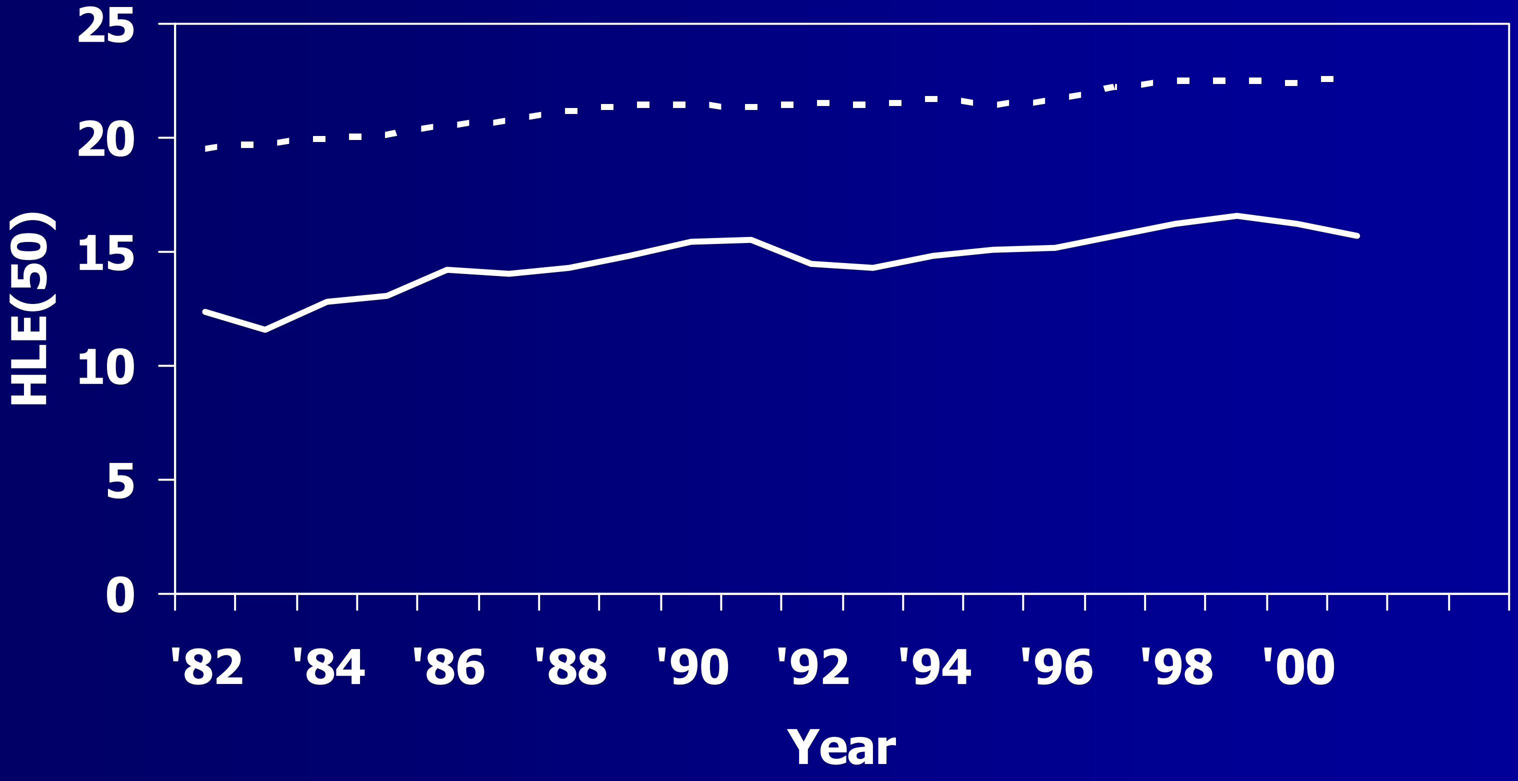
# 2002 Result

- % of black-white difference explained by SES is 48.8%
- s.e. of this difference is .045,
- Empirical interval is [.41,.58]
- NOT same result for standard model of health:  b=.41 without SES, b=.24 with (40.6% reduction)

# Total Life at 50, '82-'01 (white=dashed line)



# Healthy Life at 50, ’82-’01



# Proportion of Race Difference in HLE% Explained by SES

# Proportion of Race Difference in HLE% Explained by SES