Variation of mode choice based on time of the day

Group B **HE Jiahang (Nagoya University)**

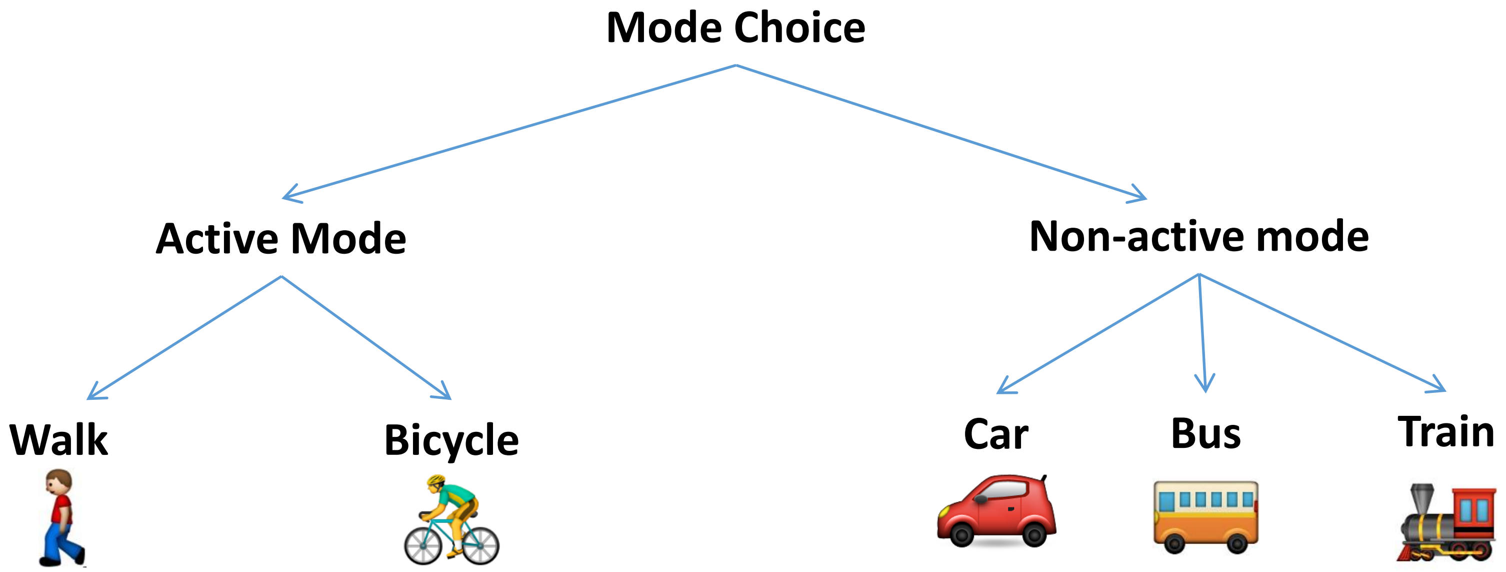
Su Mon Nwe (Nagoya University)

Aupal Mondal (IIT Bombay)

Karan Barpete (IIT Bombay)

Varun Varghese (IIT Bombay)

Primary structure of the Mode choice model



Why a Nested Logit Model based on Active and Non-active modes?

- Travellers in general have a differentiated perception about modes that require physical activity (active modes) from the non-active modes.
- Characteristics of traveller may influence active and non-active mode choice .
- Time of the day (peak/off-peak) and the purpose of the trip may affect the choice of active or non-active model of an individual.

```r
  pa <- x[6]
  a1 <- x[7]
  a2 <- x[8]
  p1<-x[9]
  p2<-x[10]
  p3<-x[11]
  p4<-x[12]
  #a3 <- x[13]
 #a4 <- x[10]
    ## declare the log-likelihood variable, set value to 0
    LL = 0

 ## calculate the utility function: :introduce the desired explanatory variables in the function
    train   <- Data$ModeAvailableTrain*exp(d1*Data$TotalTimeTrain/100+p1*Data$Peak  +b1*matrix(1,nrow =hh,ncol=1))
    bus     <- Data$ModeAvailableBus   *exp(d1*Data$TotalTimeBus/100     +b2*matrix(1,nrow =hh,ncol=1))
    car     <- Data$ModeAvailableCar   *exp(d1*Data$TimeCar/100+p2*Data$Peak          +b3*matrix(1,nrow =hh,ncol=1))
    bike    <- Data$ModeAvailableBike *exp(d1*Data$TimeBike/100+p3*Data$Peak        +b4*matrix(1,nrow =hh,ncol=1))
    walk    <- Data$ModeAvailableWalk *exp(d1*Data$TimeWalk/100+p4*Data$Peak )
    ### Construct nests.First, calculate active or non-active choice probabilityP(LV1)
    ## setting of log-sum variables
   logsum.act   <- log( ( (bike +walk)!=0) * (bike + walk) + ((bike+walk)==0))
   logsum.nonact <- log( ( (car + bus+ train)!=0) * (car + bus+ train) + ((car + bus+ train)==0))

   ##Calculate active or non-active choice probability with these log-sum variables
   nume.act    <- (logsum.act != 0)*exp(pa*logsum.act+a2*Data$Leisure) + (logsum.act == 0)
   nume.nonact  <- (logsum.nonact != 0)*exp(pa*logsum.nonact + a1*Data$AgeFemale45) + (logsum.nonact == 0)

   deno <- nume.act + nume.nonact
   P.act <- nume.act / deno
   P.nonact <- nume.nonact / deno
```

Our original utility function of nested logit model.

Nested function

```
  ### Next, calculate conditional probability.P(LV2|LV1)
  ## conditional probability under active choice
 deno.act     <- bike + walk
 P.bike.act  <- bike / ((deno.act!=0)*deno.act + (deno.act==0))
 P.walk.act   <- walk   / ((deno.act!=0)*deno.act+ (deno.act==0))

## conditional probability non-active choice
 deno.nonact   <- car + bus + train
 P.car.nonact  <- car   / ((deno.nonact!=0)*deno.nonact + (deno.nonact==0))
 P.bus.nonact  <- bus  / ((deno.nonact!=0)*deno.nonact + (deno.nonact==0))
 P.train.nonact <- train / ((deno.nonact!=0)*deno.nonact + (deno.nonact==0))

 ### Finally, calculate joint probability. P(LV1, LV2) = P(LV2|LV1)*P(LV1)
   P.bike <- P.bike.act * P.act
   P.walk   <- P.walk.act * P.act
   P.car    <- P.car.nonact  * P.nonact
   P.bus <- P.bus.nonact * P.nonact
   P.train  <- P.train.nonact * P.nonact
   ## Avoid problems stemming from choice probabilities becoming zero.
   P.train <- (P.train!=0)*P.train + (P.train==0)
   P.bus    <- (P.bus!=0)  *P.bus    + (P.bus==0)
   P.car    <- (P.car!=0)  *P.car    + (P.car==0)
   P.bike   <- (P.bike!=0) *P.bike   + (P.bike==0)
  P.walk   <- (P.walk!=0) *P.walk   + (P.walk==0)
```

```
     ## Choice results
    Ctrain    <- Data$MainModeENG =="Rail"
   Cbus      <- Data$MainModeENG =="Bus"
   Ccar      <- Data$MainModeENG =="Car"
   Cbike     <- Data$MainModeENG =="Bicycle"
  Cwalk     <- Data$MainModeENG =="Walk"
    ## Calculate the Log-likelihood function
  LL <- colSums(Ctrain*log(P.train) + Cbus*log(P.bus) + Ccar  *log(P.car)   + Cbike  *log(P.bike) +Cwalk *log(P.walk))

   }
      ##### Maximize the Log-likelihood function#####

          ##Parameter optimization
          res <- optim(b0,fr, upper=c(Inf, Inf, Inf, Inf, Inf, Inf, Inf), method = "L-BFGS-B", hessian = TRUE, control=

           ## Parmeter estimation?Hessian matrix calculation
             b    <- res$par
            hhh <- res$hessian

               ## Calculate the t-statistic
                tval <- b/sqrt(-diag(solve(hhh)))

                  ## L(0), Log-Likelihood when all parameters are 0
                L0 <- fr(b0)
               ## LL, maximium likelihood
                 LL <- res$value
```

We calculate the scale parameter at first

```
>                    ##### Output #####
>                    print(res)
$par
 [1]   0.37973780  -1.94622631  -1.31251851  -1.08916068 -11.12017329   1.01370529   0.67255019   0.92477451   0.05081488
[10]  -0.84076453  -0.91399755  -0.38458997

$value
[1] -1260.269

$counts
function gradient 
      66       66 

$convergence
[1] 0

$message
[1] "CONVERGENCE: REL_REDUCTION_OF_F <= FACTR*EPSMCH"
```

And we got the answer of scale parameter as 1.01, which is very close to 1.

```r
   ## Choice results
  Ctrain    <- Data$MainModeENG =="Rail"
 Cbus      <- Data$MainModeENG =="Bus"
 Ccar      <- Data$MainModeENG =="Car"
 Cbike     <- Data$MainModeENG =="Bicycle"
Cwalk     <- Data$MainModeENG =="Walk"
  ## Calculate the Log-likelihood function
LL <- colSums(Ctrain*log(P.train) + Cbus*log(P.bus) + Ccar  *log(P.car)   + Cbike  *log(P.bike) +Cwalk *log(P.walk))

 }
     ##### Maximize the Log-likelihood function#####

         ##Parameter optimization
         res <- optim(b0,fr, upper=c(Inf, Inf, Inf, Inf, Inf, 1, Inf), method = "L-BFGS-B", hessian = TRUE, control=list(fr

          ## Parmeter estimation?Hessian matrix calculation
            b    <- res$par
          hhh <- res$hessian

              ## Calculate the t-statistic
               tval <- b/sqrt(-diag(solve(hhh)))

                 ## L(0), Log-Likelihood when all parameters are 0
               L0 <- fr(b0)
              ## LL, maximium likelihood
                LL <- res$value
```

So we set the scale parameter as 1, then the model became a multinomial logit model.

```
>                  ##### Output #####
>                  print(res)
$par
 [1]   0.37973780  -1.94622631  -1.31251851  -1.08916068 -11.12017329   1.01370529   0.67255019   0.92477451   0.05081488
[10]  -0.84076453  -0.91399755  -0.38458997

$value
[1] -1260.269

$counts
function gradient
      66       66
```

Nested logit model

Multinomial logit model

```
>                  ##### Output #####
>                  print(res)
$par
 [1]   0.38159200  -1.95083329  -1.32217701  -1.09506010 -11.17634759   1.00000000   0.67432473   0.92125786   0.05730248
[10]  -0.83645881  -0.91138658  -0.38181228

$value
[1] -1260.282

$counts
function gradient
      48       48

$convergence
[1] 0

$message
[1] "CONVERGENCE: REL_REDUCTION_OF_F <= FACTR*EPSMCH"
```

```
>                   ## L(0)
>                   print(L0)
[1] -2158.061

>                   ## LL
>                   print(LL)
[1] -1260.269

>                   ##rho-square
>                   print((L0-LL)/L0)
[1] 0.416018

>                   ## adjusted rho-square
>                   print((L0-(LL-length(b)))/L0)
[1] 0.4104575
```

Original nested logit model

```
>                   ## L(0)
>                   print(L0)
[1] -2158.061

>                   ## LL
>                   print(LL)
[1] -1260.282

>                   ##rho-square
>                   print((L0-LL)/L0)
[1] 0.4160121

>                   ## adjusted rho-square
>                   print((L0-(LL-length(b)))/L0)
[1] 0.4104516
```

Multinomial logit model

Nested logit model

```
>                 ##estimated parameter values
>                 print(b)
 [1]   0.37973780  -1.94622631  -1.31251851  -1.08916068 -11.12017329   1.01370529   0.67255019   0.92477451   0.05081488
[10]  -0.84076453  -0.91399755  -0.38458997
```

```
>                 ## t-statistic
>                 print(tval)
 [1]   2.405447  -6.366029  -8.154827  -7.590359 -17.256611  11.604624   2.170456   2.318519   0.131675  -2.238590
[11]  -2.374214  -1.003533
```

Peak or Non peak explanatory variable

According to our model, it means if the travel time increase, the possibility of choosing the alternative is decreasing.

Multinomial logit model

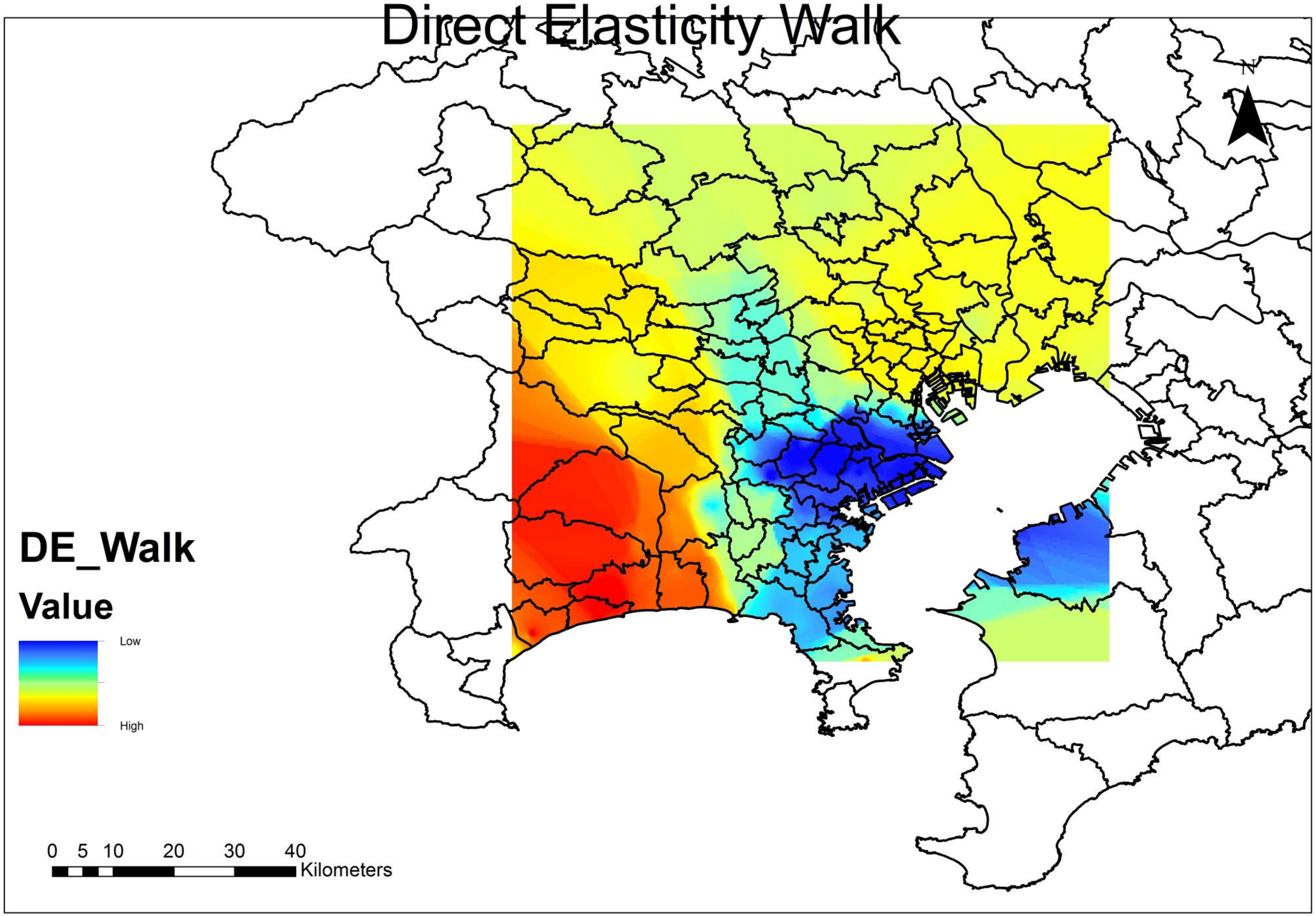
```
>                 ##estimated parameter values
>                 print(b)
 [1]   0.38159200  -1.95083329  -1.32217701  -1.09506010 -11.17634759   1.00000000   0.67432473   0.92125786   0.05730248
[10]  -0.83645881  -0.91138658  -0.38181228
```

```
>                 ## t-statistic
>                 print(tval)
 [1]   2.4021272  -6.3675545  -8.1721911  -7.6274878 -17.3352552  11.6569103   2.1793043   2.3165880   0.1482927
[10]  -2.2247825  -2.3608487  -0.9931024
```

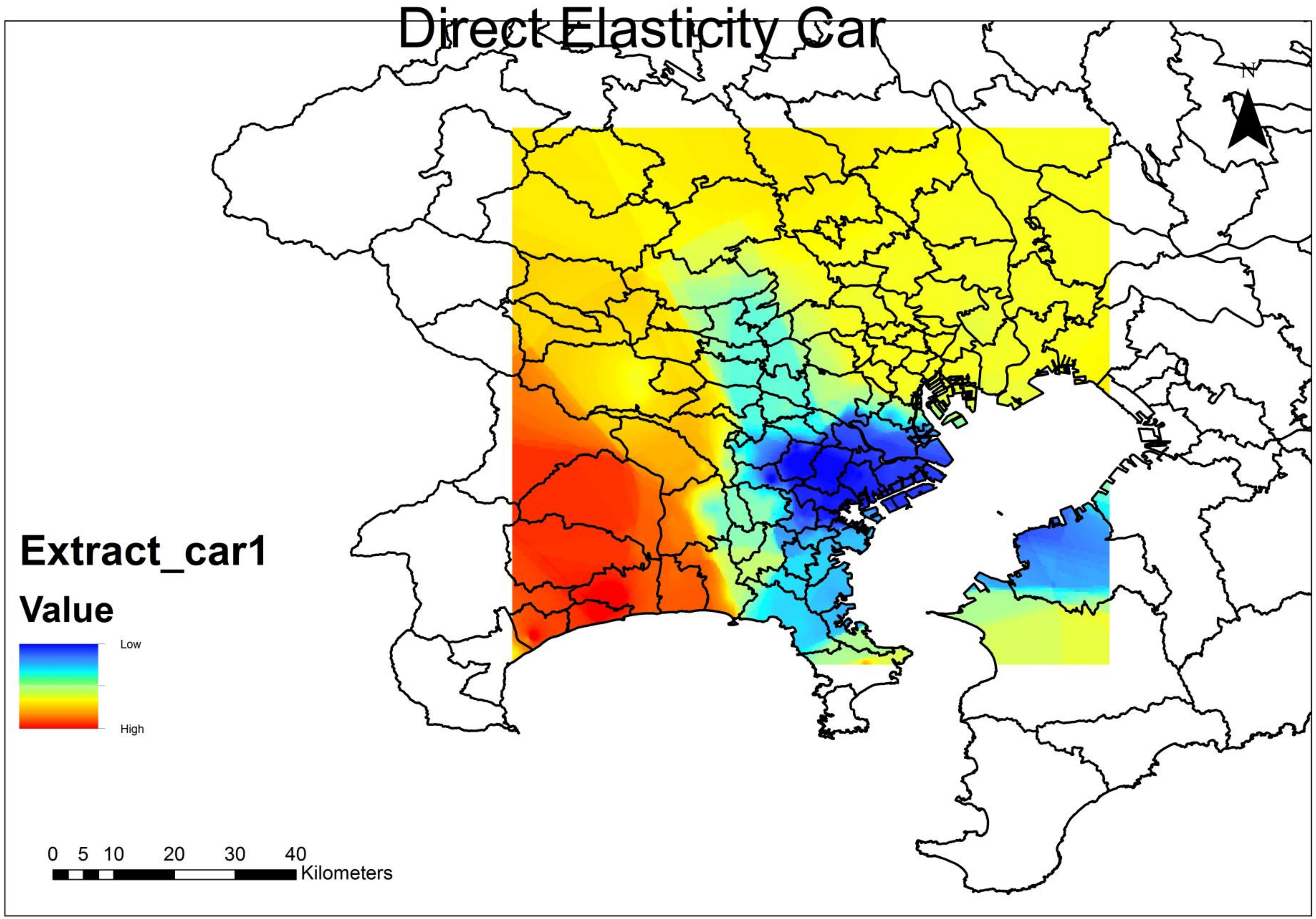
Elasticity based on distance

Distance		DE_bike_time	DE_walk_time	DE_car_time	DE_bus_time	DE_train_time
Less than 1km	Mean	-.42	-.63	-.56	-.12	-.74
	N	334.00	334.00	334.00	334.00	334.00
	Std. Deviation	.17	.28	.30	.16	1.06
1km-3km	Mean	-1.18	-1.97	-1.13	-.41	-1.87
	N	342.00	342.00	342.00	342.00	342.00
	Std. Deviation	.36	.66	.35	.56	1.07
3km-5km	Mean	-2.63	-4.22	-1.87	-.41	-2.61
	N	115.00	115.00	115.00	115.00	115.00
	Std. Deviation	.48	.89	.47	.90	1.20
5km-10km	Mean	-5.07	-7.03	-2.73	-.07	-2.41
	N	120.00	120.00	120.00	120.00	120.00
	Std. Deviation	1.02	1.31	.56	.48	1.04
Gretear than 10km	Mean	-14.62	-12.41	-6.41	-.08	-6.09
	N	611.00	611.00	611.00	611.00	611.00
	Std. Deviation	6.29	2.63	2.43	.46	3.00
Total	Mean	-6.83	-6.44	-3.31	-.19	-3.42
	N	1522.00	1522.00	1522.00	1522.00	1522.00
	Std. Deviation	7.63	5.45	3.04	.51	3.07

Cross Elasticity					
Wavg_bike	**Wavg_walk**	**Wavg_car**	**Wavg_bus**	**Wavg_train**	**Distance**
0.06	0.71	0.21	0.17	1.95	a) less than 1 km
0.11	1.19	0.15	0.17	2.2	b) 1km-3km
0.22	2.85	0.19	0.12	2.88	c) 3km-5km
0.33	6.42	0.14	0.15	1.9	d) 5km-10km
0.35	26.95	0.21	0.08	2.82	e) Greater than 10km



Direct Elasticity Walk
N
DE_Walk
Value
Low
High
0 5 10 20 30 40
Kilometers



Direct Elasticity Car
N
Extract_car1
Value
Low
High
0 5 10 20 30 40
Kilometers

Conclusion

- MNL structure most appropriate for the model
- Variation in elasticities plotted based on distance and spatially
- Elasticities increase with increasing distance for all modes.
- Spatial variation in elasticity for modes can be used for targeted policies in designated areas.

Thank You