

> kable(married\_male)

|edu | freq|panel |

|:---------|----:|:-----|

|hypogamy | 516|1996 |

|homophily | 1190|1996 |

|hypergamy | 583|1996 |

|hypogamy | 266|2001 |

|homophily | 698|2001 |

|hypergamy | 424|2001 |

|hypogamy | 284|2004 |

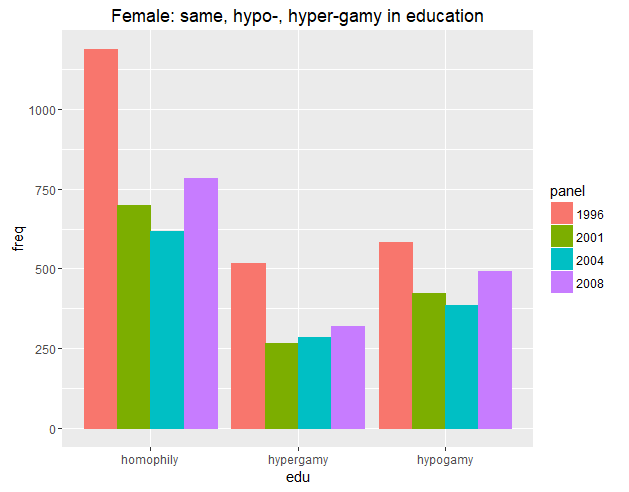
|homophily | 618|2004 |

|hypergamy | 387|2004 |

|hypogamy | 321|2008 |

|homophily | 783|2008 |

|hypergamy | 493|2008 |



> kable(married\_female)

|edu | freq|panel |

|:---------|----:|:-----|

|hypogamy | 583|1996 |

|homophily | 1190|1996 |

|hypergamy | 516|1996 |

|hypogamy | 424|2001 |

|homophily | 698|2001 |

|hypergamy | 266|2001 |

|hypogamy | 387|2004 |

|homophily | 618|2004 |

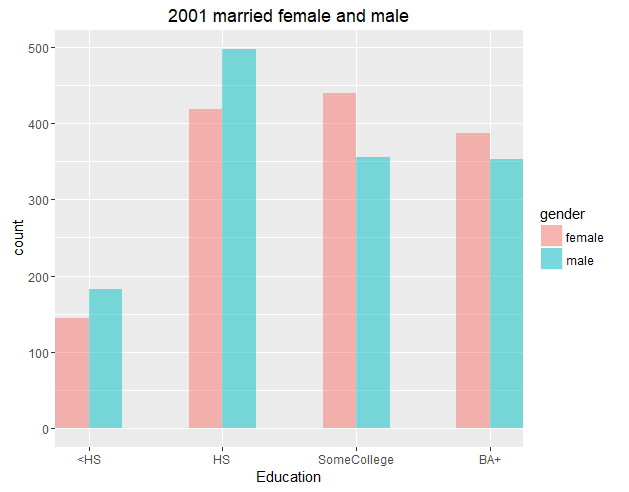
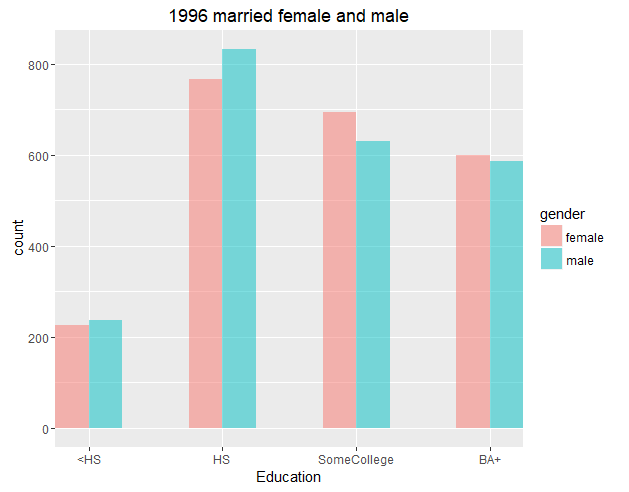
|hypergamy | 284|2004 |

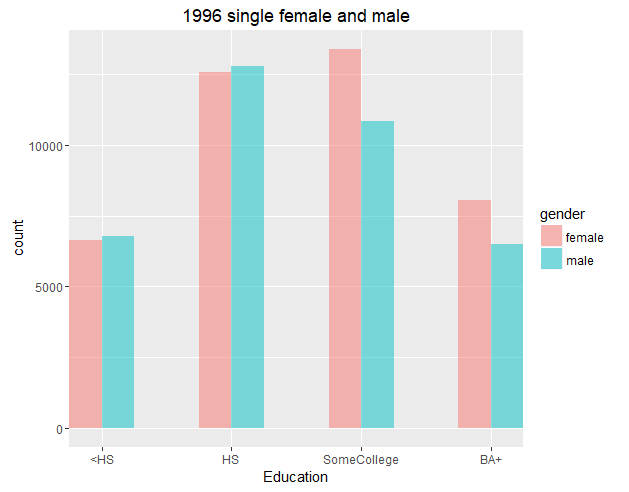
|hypogamy | 493|2008 |

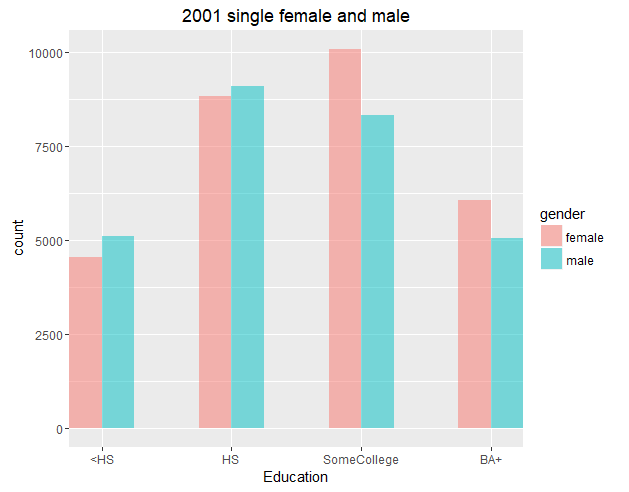
|homophily | 783|2008 |

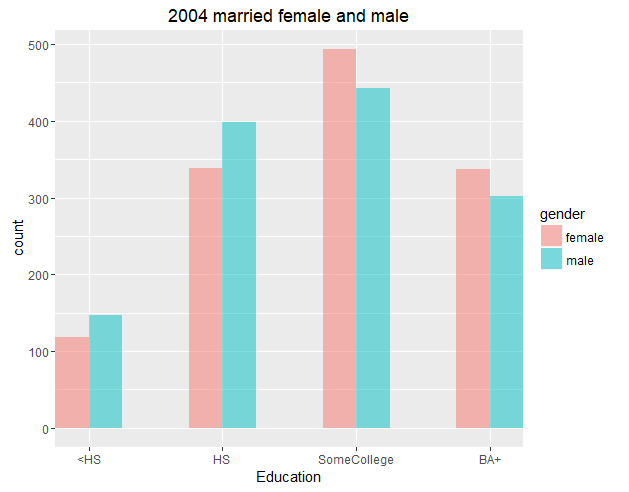
|hypergamy | 321|2008 |

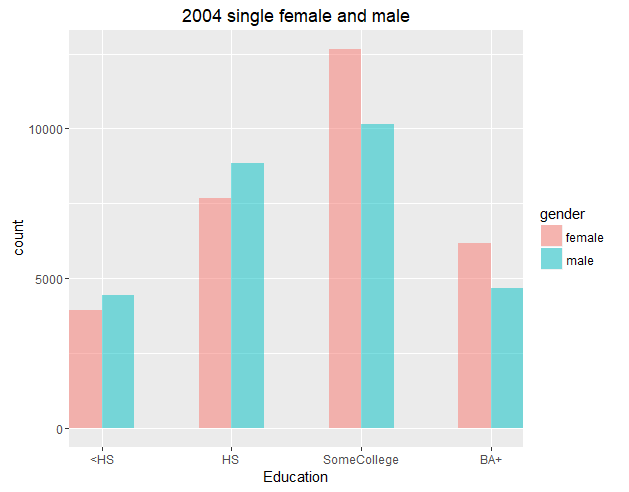
Note that hypergamy occurs more often in males than hypogamy. It’s the opposite for females.

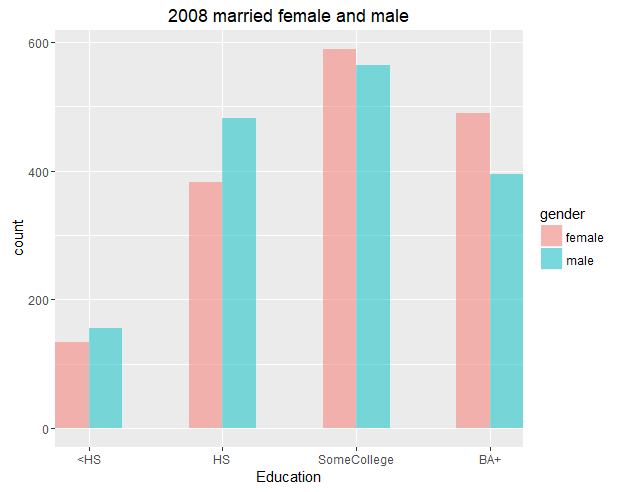


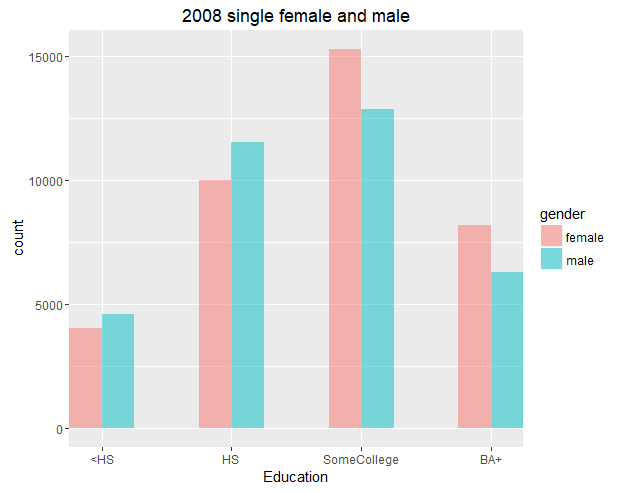


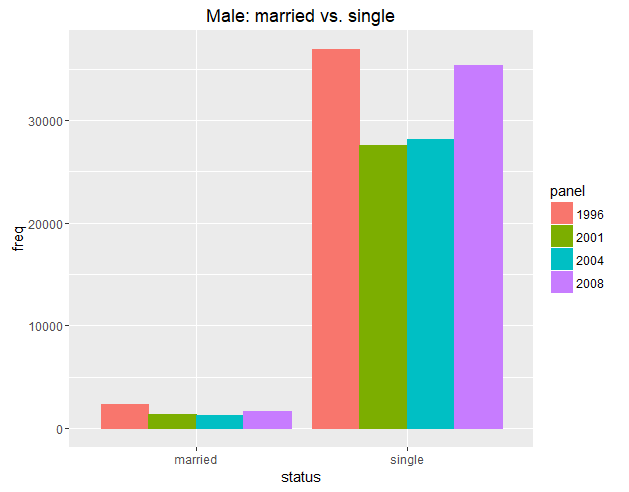
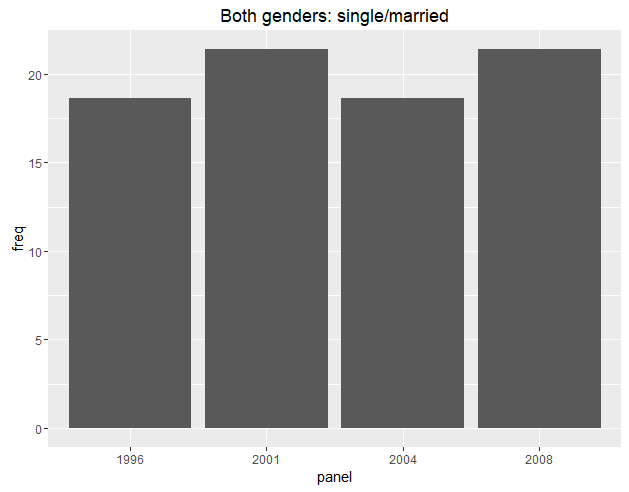
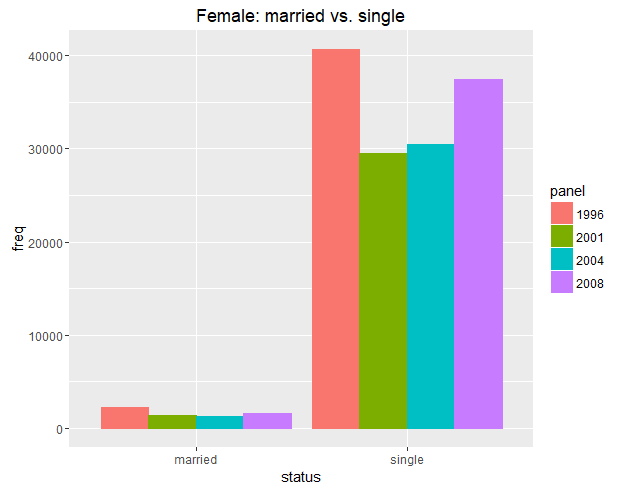
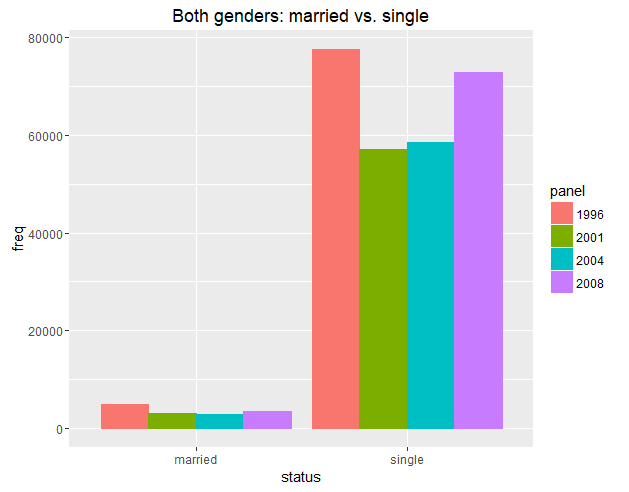


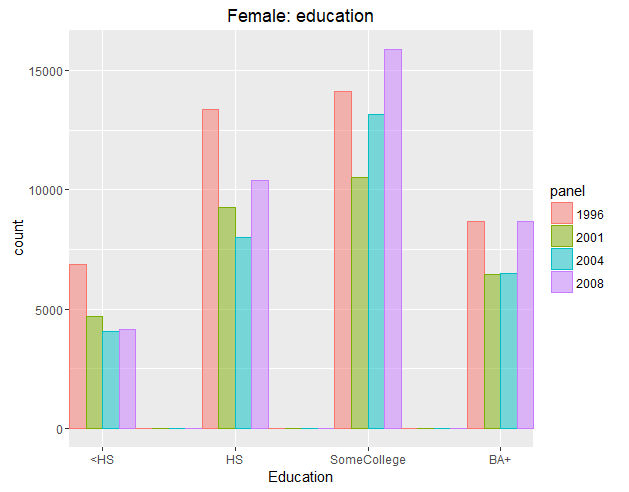


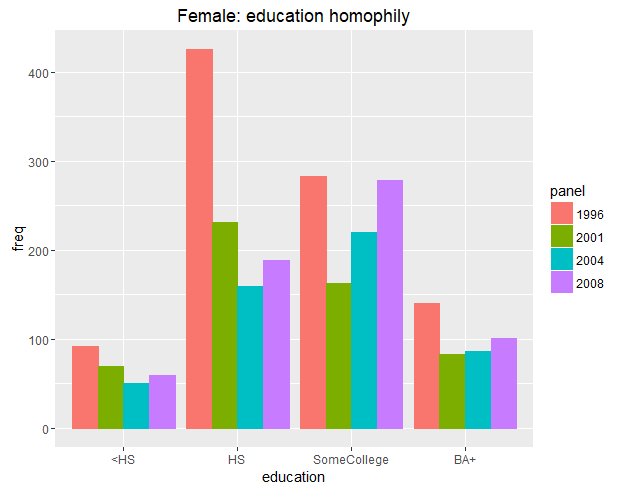


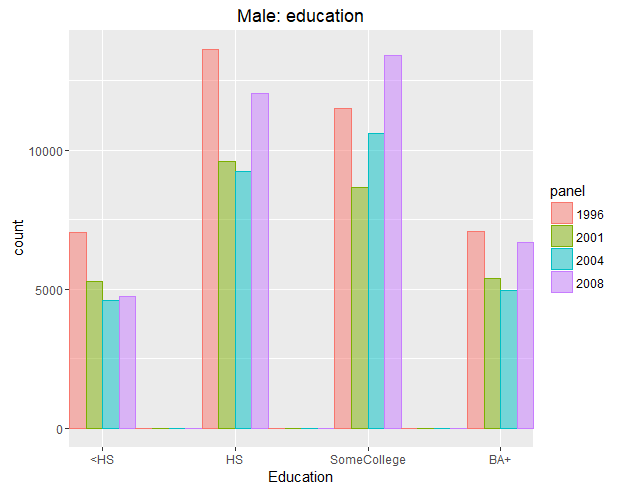


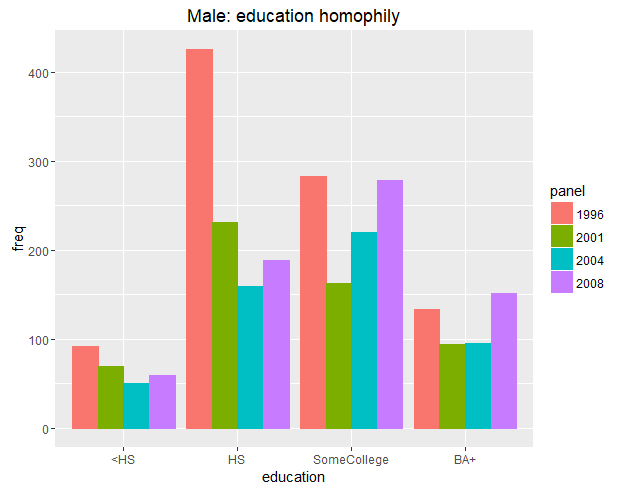












| | 1996| 2001| 2004| 2008|

|:-------------------------|----------:|----------:|----------:|----------:|

|b1.intercept | -1.0577143| -1.1088192| -1.1148591| -1.1042116|

|b1nodematch.educlevel\_t.1 | 0.0149054| 0.1693907| 0.1564603| 0.2715590|

|b1nodematch.educlevel\_t.2 | 0.2804184| 0.2388275| 0.1460179| 0.0708076|

|b1nodematch.educlevel\_t.3 | 0.0652752| 0.0449458| -0.0227423| -0.0139418|

|b1nodematch.educlevel\_t.4 | 0.8182828| 0.7112963| 0.6577622| 0.6282845|

|b1absdiff.educlevel\_t.1 | 0.5462528| 0.4811792| 0.3958748| 0.3562347|

|b1greaterthan.educlevel\_t | -1.1251202| -1.2845949| -1.0543652| -1.1177411|

|b2.intercept | -1.0577143| -1.1088192| -1.1148591| -1.1042116|

|b2nodematch.educlevel\_t.1 | 0.0149054| 0.1693907| 0.1564603| 0.2715590|

|b2nodematch.educlevel\_t.2 | 0.2804184| 0.2388275| 0.1460179| 0.0708076|

|b2nodematch.educlevel\_t.3 | 0.0652752| 0.0449458| -0.0227423| -0.0139418|

|b2nodematch.educlevel\_t.4 | 0.8182828| 0.7112963| 0.6577622| 0.6282845|

|b2absdiff.educlevel\_t.1 | 0.5462528| 0.4811792| 0.3958748| 0.3562347|

|b2greaterthan.educlevel\_t | -1.1114760| -0.9886846| -0.9979918| -0.9431799|

|ExpUtil.b1.1 | 0.0366710| 0.0317510| 0.0317599| 0.0316256|

|ExpUtil.b1.2 | 0.0611804| 0.0492391| 0.0438271| 0.0398927|

|ExpUtil.b1.3 | 0.0491126| 0.0427746| 0.0389531| 0.0400172|

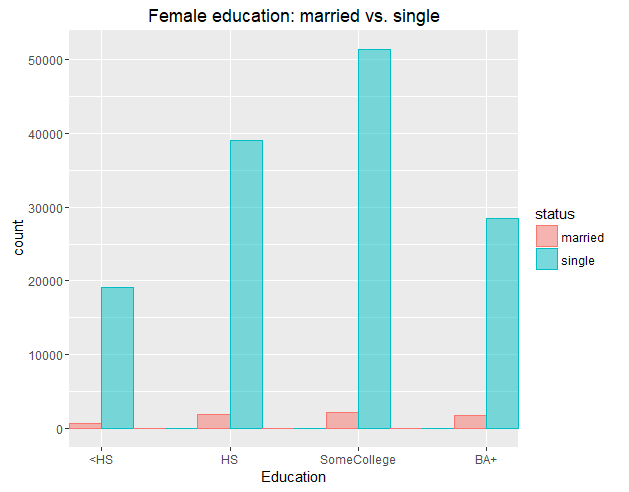
|ExpUtil.b1.4 | 0.0741079| 0.0625082| 0.0537639| 0.0558845|

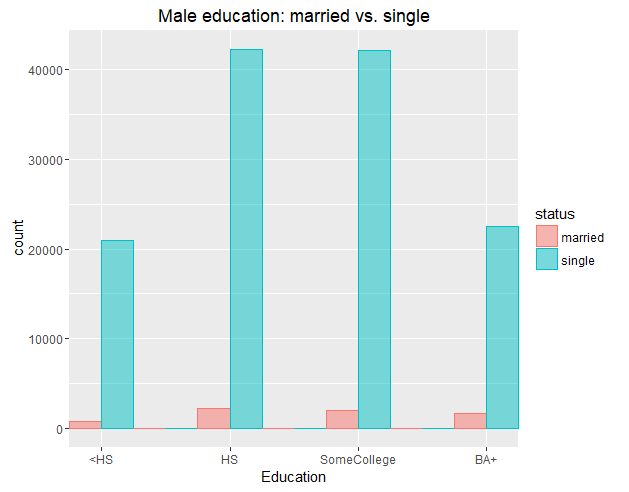
|ExpUtil.b2.1 | 0.0385051| 0.0381429| 0.0326585| 0.0346878|

|ExpUtil.b2.2 | 0.0649741| 0.0537960| 0.0455148| 0.0426637|

|ExpUtil.b2.3 | 0.0550820| 0.0447094| 0.0429068| 0.0424293|

|ExpUtil.b2.4 | 0.0890121| 0.0657174| 0.0646953| 0.0630287|





In the above plots, breaking out by panel year still looks about the same.

pmfj = check\_CP\_latent(ff, out$solution, mu, Xdata, Zdata, symmetric)

> print("estimated joint probabilities")

[1] "estimated joint probabilities"

> print(pmfj$pmfj\_est)

[,1] [,2] [,3] [,4] [,5]

[1,] 0.0016488875 0.002932406 0.0008361947 0.0004995955 0.08067288

[2,] 0.0029790076 0.010069841 0.0047297973 0.0009477297 0.15303597

[3,] 0.0010667963 0.006021361 0.0059454296 0.0030173990 0.16340772

[4,] 0.0006407468 0.001212915 0.0030750496 0.0096195209 0.09814711

[5,] 0.0822632446 0.155721943 0.1324043739 0.0791067420 0.00000000

> print("observed joint probabilities")

[1] "observed joint probabilities"

> print(pmfj$pmfj\_obs)

[,1] [,2] [,3] [,4] [,5]

[1,] 0.0022372997 0.002529121 0.0005836434 0.0001702293 0.08087109

[2,] 0.0024561659 0.010359670 0.0046205102 0.0012402422 0.15306048

[3,] 0.0009727390 0.005617568 0.0068821284 0.0034045865 0.16299458

[4,] 0.0001215924 0.001775249 0.0032343571 0.0094598867 0.09812505

[5,] 0.0825369033 0.155699035 0.1320371586 0.0790107244 0.00000000

> # KL-divergence = sum(p\*log(p/q))

> kl = -sum(pmfj$pmfj\_est\*log(pmfj$pmfj\_est/pmfj$pmfj\_obs), na.rm=T)

> print("KL-divergence: p = 1996 est. preference, q = 1996 observed")

[1] "KL-divergence: p = 1996 est. preference, q = 1996 observed"

> print(kl)

[1] -0.001231639

**Something is wrong…it’s not supposed to be negative**

The other direction:

> kl = -sum(pmfj$pmfj\_obs\*log(pmfj$pmfj\_obs/pmfj$pmfj\_est), na.rm=T)

> print("KL-divergence: q = 1996 est. preference, p = 1996 observed")

[1] "KL-divergence: q = 1996 est. preference, p = 1996 observed"

> print(kl)

[1] -0.0009552019

> # compare counterfactual joint probabilities with observed

> pmfjc = create\_counterfactual\_distri(ff, out$solution, mu2008, Xdata, Zdata, Xdata2008, Zdata2008, symmetric)

> print("counterfactual joint probabilities")

[1] "counterfactual joint probabilities"

> print(pmfjc$pmfj\_est)

[,1] [,2] [,3] [,4] [,5]

[1,] 0.0007280203 0.001689164 0.0006377624 0.0003084626 0.05266375

[2,] 0.0016942601 0.007471820 0.0046467582 0.0007537450 0.12868671

[3,] 0.0008802883 0.006482363 0.0084747277 0.0034818308 0.19936445

[4,] 0.0004686050 0.001157300 0.0038848265 0.0098379668 0.10612792

[5,] 0.0602111678 0.148701828 0.1674066177 0.0809685181 0.00000000

> print("observed joint probabilities")

[1] "observed joint probabilities"

> print(pmfjc$pmfj\_obs)

[,1] [,2] [,3] [,4] [,5]

[1,] 0.0015527133 0.001315859 0.0005526607 0.0001052687 0.05283173

[2,] 0.0013158587 0.004947629 0.0029738407 0.0008421496 0.13152008

[3,] 0.0010526870 0.004500237 0.0073161745 0.0026580346 0.20139218

[4,] 0.0001842202 0.001921154 0.0040002105 0.0067898310 0.10754513

[5,] 0.0604768672 0.152021159 0.1692062740 0.0829780515 0.00000000

> # KL-divergence = sum(p\*log(p/q))

> klc = -sum(pmfjc$pmfj\_est\*log(pmfjc$pmfj\_est/pmfjc$pmfj\_obs), na.rm=T)

> print("KL-divergence: p = counterfactual 2008 with 1996 preference, q = 2008 observed")

[1] "KL-divergence: p = counterfactual 2008 with 1996 preference, q = 2008 observed"

> print(klc)

[1] 0.0001470872

# reducing singles to 0.05

| | 1996| 2001| 2004| 2008|

|:-------------------------|----------:|----------:|----------:|----------:|

|b1.intercept | 2.8944757| 2.8752936| 3.0453375| 2.9410035|

|b1nodematch.educlevel\_t.1 | 0.9181226| 0.8874575| 0.8156821| 0.9767097|

|b1nodematch.educlevel\_t.2 | 0.6608916| 0.7881092| 0.6747625| 0.5348664|

|b1nodematch.educlevel\_t.3 | 0.6362321| 0.6165017| 0.6169504| 0.6407969|

|b1nodematch.educlevel\_t.4 | 1.1413819| 0.9578104| 0.9901566| 0.9601215|

|b1absdiff.educlevel\_t.1 | 0.5397426| 0.5232098| 0.4098541| 0.4157508|

|b1greaterthan.educlevel\_t | -0.2336312| -0.3724102| -0.1037268| -0.2002352|

|b2.intercept | 2.8944757| 2.8752936| 3.0453377| 2.9410035|

|b2nodematch.educlevel\_t.1 | 0.9181227| 0.8874575| 0.8156820| 0.9767097|

|b2nodematch.educlevel\_t.2 | 0.6608916| 0.7881091| 0.6747623| 0.5348664|

|b2nodematch.educlevel\_t.3 | 0.6362321| 0.6165017| 0.6169505| 0.6407970|

|b2nodematch.educlevel\_t.4 | 1.1413819| 0.9578104| 0.9901565| 0.9601215|

|b2absdiff.educlevel\_t.1 | 0.5397427| 0.5232098| 0.4098541| 0.4157509|

|b2greaterthan.educlevel\_t | -0.2285213| -0.0021751| 0.0515128| 0.0287441|

|ExpUtil.b1.1 | 16.4696349| 14.4519070| 17.5772189| 16.1598389|

|ExpUtil.b1.2 | 19.8825752| 20.5451535| 22.5133764| 17.5742331|

|ExpUtil.b1.3 | 19.2211211| 18.9777876| 22.8836551| 21.3882451|

|ExpUtil.b1.4 | 23.4824064| 22.5687014| 25.6878851| 22.6841444|

|ExpUtil.b2.1 | 15.4177667| 16.9873481| 17.6833623| 17.1966161|

|ExpUtil.b2.2 | 19.1200775| 21.5564979| 22.5015753| 17.8782188|

|ExpUtil.b2.3 | 19.6642140| 18.8072479| 23.3724555| 21.2263005|

|ExpUtil.b2.4 | 25.6265910| 19.7097996| 25.9687504| 23.7630971|

|  |
| --- |
| # reducing singles to 0.025  | | 1996| 2001| 2004| 2008|  |:-------------------------|----------:|----------:|----------:|----------:|  |b1.intercept | 2.8174458| 2.3830356| 3.1662275| 3.3047685|  |b1nodematch.educlevel\_t.1 | 0.9632770| 0.8711797| 0.8828231| 1.0422874|  |b1nodematch.educlevel\_t.2 | 0.8269403| 0.7517252| 0.6969595| 0.7097327|  |b1nodematch.educlevel\_t.3 | 0.6140762| 0.5374827| 0.6426121| 0.7059074|  |b1nodematch.educlevel\_t.4 | 0.9549546| 0.8668768| 0.9404029| 0.8635971|  |b1absdiff.educlevel\_t.1 | 0.5814699| 0.5591399| 0.4403457| 0.4300273|  |b1greaterthan.educlevel\_t | -0.3038387| -0.5004367| -0.0856192| -0.1353230|  |b2.intercept | 2.8174458| 2.3830356| 3.1662275| 3.3047685|  |b2nodematch.educlevel\_t.1 | 0.9632770| 0.8711797| 0.8828231| 1.0422874|  |b2nodematch.educlevel\_t.2 | 0.8269403| 0.7517252| 0.6969595| 0.7097327|  |b2nodematch.educlevel\_t.3 | 0.6140762| 0.5374827| 0.6426121| 0.7059074|  |b2nodematch.educlevel\_t.4 | 0.9549546| 0.8668768| 0.9404029| 0.8635971|  |b2absdiff.educlevel\_t.1 | 0.5814699| 0.5591399| 0.4403458| 0.4300273|  |b2greaterthan.educlevel\_t | -0.2379634| -0.1437920| 0.0890492| 0.1185668|  |ExpUtil.b1.1 | 14.9127889| 8.4813013| 20.6466964| 23.9791259|  |ExpUtil.b1.2 | 20.8785632| 11.9815632| 26.0603963| 29.0748267|  |ExpUtil.b1.3 | 17.7464401| 10.7970392| 27.1003055| 33.0280381|  |ExpUtil.b1.4 | 19.4410405| 12.4199672| 28.6764433| 31.9154091|  |ExpUtil.b2.1 | 14.6942203| 9.7755862| 21.1282243| 26.3865924|  |ExpUtil.b2.2 | 20.4384951| 12.5715978| 26.5238449| 29.4878567|  |ExpUtil.b2.3 | 17.9797935| 10.7029772| 27.4929499| 32.9352634|  |ExpUtil.b2.4 | 19.5975476| 10.9312592| 28.2010840| 31.2976835| |
|  |
| |  | | --- | |  | |

# reducing singles to 0.01

| | 1996| 2001| 2004| 2008|

|:-------------------------|----------:|----------:|----------:|----------:|

|b1.intercept | 4.0791194| 3.0578763| 3.4153799| 3.3912561|

|b1nodematch.educlevel\_t.1 | 1.1095668| 1.0093616| 0.9603488| 1.0007420|

|b1nodematch.educlevel\_t.2 | 0.8897209| 0.8313392| 0.7484703| 0.7224350|

|b1nodematch.educlevel\_t.3 | 0.8293695| 0.6829953| 0.6827009| 0.7462405|

|b1nodematch.educlevel\_t.4 | 1.1315124| 0.9508140| 0.9974678| 0.8903915|

|b1absdiff.educlevel\_t.1 | 0.5657488| 0.5353324| 0.4515912| 0.4206608|

|b1greaterthan.educlevel\_t | 0.0222564| -0.4047700| -0.0829459| -0.1025409|

|b2.intercept | 4.0791189| 3.0578763| 3.4153798| 3.3912561|

|b2nodematch.educlevel\_t.1 | 1.1095668| 1.0093616| 0.9603488| 1.0007420|

|b2nodematch.educlevel\_t.2 | 0.8897208| 0.8313392| 0.7484703| 0.7224350|

|b2nodematch.educlevel\_t.3 | 0.8293695| 0.6829953| 0.6827009| 0.7462406|

|b2nodematch.educlevel\_t.4 | 1.1315124| 0.9508140| 0.9974678| 0.8903915|

|b2absdiff.educlevel\_t.1 | 0.5657489| 0.5353324| 0.4515912| 0.4206608|

|b2greaterthan.educlevel\_t | 0.0966938| -0.0118638| 0.1093380| 0.1339878|

|ExpUtil.b1.1 | 62.4702253| 18.7291084| 27.6617234| 25.9775846|

|ExpUtil.b1.2 | 80.6946352| 25.4139826| 34.7581042| 32.3505608|

|ExpUtil.b1.3 | 78.3104338| 23.5717840| 35.8547588| 36.8465479|

|ExpUtil.b1.4 | 83.4189417| 26.7880380| 38.5293574| 34.9176939|

|ExpUtil.b2.1 | 63.2536204| 21.6701625| 28.3380957| 27.8618501|

|ExpUtil.b2.2 | 79.9346354| 26.5695854| 35.2938846| 32.0925139|

|ExpUtil.b2.3 | 79.0068696| 23.7324895| 36.3899105| 36.9710014|

|ExpUtil.b2.4 | 83.2450958| 23.3567420| 37.8828566| 35.3869819|

Reduced singles to 0.01, 1996 only:

> pmfj = check\_CP\_latent(ff, out$solution, mu, Xdata, Zdata, symmetric)

> print("estimated joint probabilities")

[1] "estimated joint probabilities"

> print(pmfj$pmfj\_est)

[,1] [,2] [,3] [,4] [,5]

[1,] 0.0412387010 0.039627445 0.009724375 0.008652498 0.0007942477

[2,] 0.0395098906 0.191201047 0.077808391 0.022330737 0.0020498283

[3,] 0.0118719732 0.102637400 0.120104394 0.064495077 0.0019095796

[4,] 0.0096239250 0.026836863 0.063300010 0.158528018 0.0015479862

[5,] 0.0008081346 0.002253529 0.001714478 0.001525499 0.0000000000

> print("observed joint probabilities")

[1] "observed joint probabilities"

> print(pmfj$pmfj\_obs)

[,1] [,2] [,3] [,4] [,5]

[1,] 0.0398096062 0.045002164 0.0103851147 0.003028992 0.0012981393

[2,] 0.0437040242 0.184335785 0.0822154911 0.022068369 0.0012981393

[3,] 0.0173085244 0.099956729 0.1224578105 0.060579836 0.0012981393

[4,] 0.0021635656 0.031588057 0.0575508438 0.168325400 0.0008654262

[5,] 0.0004327131 0.001947209 0.0008654262 0.001514496 0.0000000000

>

> # KL-divergence = sum(p\*log(p/q))

> kl = -sum(pmfj$pmfj\_est\*log(pmfj$pmfj\_est/pmfj$pmfj\_obs), na.rm=T)

> print("KL-divergence: p = 1996 est. preference, q = 1996 observed")

[1] "KL-divergence: p = 1996 est. preference, q = 1996 observed"

> print(kl)

[1] -0.01449855

>

> # compare counterfactual joint probabilities with observed

> pmfjc = create\_counterfactual\_distri(ff, out$solution, mu2008, Xdata, Zdata, Xdata2008, Zdata2008, symmetric)

> print("counterfactual joint probabilities")

[1] "counterfactual joint probabilities"

> print(pmfjc$pmfj\_est)

[,1] [,2] [,3] [,4] [,5]

[1,] 0.0228617123 0.01919580 0.007277352 0.004851277 0.0006646207

[2,] 0.0188597208 0.07974915 0.050137628 0.010780608 0.0014769338

[3,] 0.0095682410 0.07228051 0.130670036 0.052571032 0.0023230635

[4,] 0.0074607588 0.01817895 0.066243378 0.124293250 0.0018113901

[5,] 0.0007675461 0.00187021 0.002198167 0.001465357 0.0000000000

> print("observed joint probabilities")

[1] "observed joint probabilities"

> print(pmfjc$pmfj\_obs)

[,1] [,2] [,3] [,4] [,5]

[1,] 0.0366004963 0.031017370 0.0130272953 0.0024813896 0.0006203474

[2,] 0.0310173697 0.116625310 0.0700992556 0.0198511166 0.0018610422

[3,] 0.0248138958 0.106079404 0.1724565757 0.0626550868 0.0012406948

[4,] 0.0043424318 0.045285360 0.0942928040 0.1600496278 0.0009305211

[5,] 0.0009305211 0.001861042 0.0009305211 0.0009305211 0.0000000000

>

> # KL-divergence = sum(p\*log(p/q))

> klc = -sum(pmfjc$pmfj\_est\*log(pmfjc$pmfj\_est/pmfjc$pmfj\_obs), na.rm=T)

> print("KL-divergence: p = counterfactual 2008 with 1996 preference, q = 2008 observed")

[1] "KL-divergence: p = counterfactual 2008 with 1996 preference, q = 2008 observed"

> print(klc)

[1] 0.2289481

>