# _Decision Memo_
# Desegregation Of Judicial System In D.C.

**CS112 - Professor Diamond, A.**

Ash Nguyen

2018/12/06

## 1 Summary

In this decision memo addressed to the policymakers that are in charge of the justice system reform in Washington, we replicated and extended on the original paper of Modak et al in 2016 about vicarious effect of experiences with the justice system. At least in the D.C. area, we came to the conclusion by different statistical methods that analyzed a survey data in 2012 that secondhand effect for experiences of the justice system did exist, meaning that individuals that had bad experiences with either the police or courts were likely to share these experiences with their acquaintances in their immediate networks, thus propagating their personal judgments and affect their acquaintances' judgments. Moreover, these secondhand effect seem to negatively affect African Americans and other minority groups such as Latinos or Asians compared to White Americans. Policymakers should keep these information in mind when reforming the judicial system to properly seal the racial divide between different race/ethnicity in D.C. More study are needed to generalized these results in different regions of the US.

## 2 Problem & Objective

Americans' views on the judicial system and the related governmental bodies vary widely based on many factors, one of which is the racial and ethnicity segregation that continues to exists despite much effort spent in the government sector to conciliate the broad gap created by historical circumstances. African Americans and other minority groups in America face disadvantages in many aspects compared to White Americans, evident by horrific events that question the treatment of the judicial system toward the different sides of the racial line such as the shooting of Michael Brown in Ferguson, Missouri. Clearly, the government must look into the causes of these issues to effectively desegregate the racial line all across the US.

There are evidences suggesting the negative view about the justice system of African Americans contributes to the unfair/unduly treatment they receive with the police and the courts, and these treatments in turn conduce the negativity toward the justice system of this population, creating a terrible reinforcement loop (Tonry 2011). It is undeniable that the negativity toward the justice system for this population is because of multiple reasons, but some papers investigating the secondary effect of negative experiences by close acquaintances do indicate the role of vicarious-based perceived injustice: people tend to view the justice system more negatively just because they heard about negative experiences with the justice system (Rosenbaum et al 2005, Weitzer & Tuch 2006). This is an important issue for at least two reasons:

(1) Communities tend to revolve around networks of closed acquaintances, and these networks are quite homogeneous in term of racial background. For example, a recent survey (Mondak et al 2016) suggests that 74.5% of network members of Black respondents living in the D.C. area are also African Americans while African Americans only constitute 4% of the D.C. population; and the statistics is a staggering 84.9% for White Americans. If secondhand effect due to network encounters exists, it has to contribute greatly to the overall perceived justice/injustice of a population.

(2) Secondhand effect concerns policies greatly, and if such effect exists it provides actionable items for policymakers to mitigate racial disparities in the judicial system: desegregation can become an effective tool to ease the racial tension between minority groups and the police, as well as directly affect the mentioned reinforcement loop to create more trust in the justice system.

Because of the relevant of vicarious effect in perceived justice, in this memo we aim to replicate and extend on parts of the results by Peffley, Testa, Mondak and Hurwitz (2016) on the effect of the encounter networks on the view of the justice system of multiple populations in the D.C. area to provide policymakers with evidenced-based information about desegregation in Washington.

# 3 Data & Analysis

## 3.1 Data

The data used in the paper comprise of responses from an online survey about Justice in Washington State in 2012, administrated by YouGov. Among 1524 respondents, 40.1% was Whites, 21.0% was Asian/Pacific Islanders, 20.0% was Latinos and 18.9% was African Americans (according to self-identification). In the original paper (and in this memo for the subsequent extension), the authors excluded 10 respondents because of invalid ZIP codes, resulting in 1514 observations in total. The survey was quite extensive, asking for information about the respondent's extended network (up to 3 people) and the respondent themselves valences of perceived justice toward the police and the court.

The survey also provided the following control variables: biological sex, education level, age, party identification, income, religious preferences, marital status and employment status. All valences were coded in a scale from -3 (very unfair) to 3 (very fair and responsive).We believe the wording and formulation of the survey questions generally did not pose any confirmation bias or nudge the respondent toward any particular position. For more information regarding the wording/prompt of the survey, please visit the Harvard Dataverse replication material for the cited paper.

In the original paper, missing data is coded as NA if they are from a control variable, and as "neutral" (which is 0 in a scale from -3 to 3) if they are from a survey response on valence (for example, missing data in the valence of police network encounter are coded as 0). In the replication analysis, we used the original coding of data to ensure results replicability. In the extension analysis, we wrote a customized imputer for missing data coded as NA to replace them with the median of each subpopulation to ensure accuracy and representativeness. For example, median values of age is imputed for each of the population of Blacks, Whites, Latinos and Asians.

As with any data collected from self-reported survey, the generalization of the subsequent replication and extension to the other populations should be taken with caution.

## 3.2 Analysis

### 3.2.1 Replication

The following analysis attempt to replicate and provide more details on the methods used to produce "Table 2" and "Figure 1" of the original paper.

| | **No Control** | **With Control** | **With Control+Personal Exp.** |
|---|---|---|---|
| **Constant** | -0.62 | -0.76 | 0.18 |
| St.Error | (0.12) | (0.31) | (0.31) |
| **White** | 0.99 | 0.85 | 0.61 |
| St.Error | (0.14) | (0.16) | (0.16) |
| **Latino** | 1.03 | 0.43 | 0.32 |
| St.Error | (0.17) | (0.17) | (0.17) |
| **Asian** | 0.49 | 0.96 | 0.85 |
| St.Error | (0.16) | (0.18) | (0.18) |

Table 1: Valence of Police Encounter Networks

Table 1 and Table 2 present the replicated, abridged version of "Table 2" in the original paper. Using simple linear regression models (function lm() in R), the authors run three different models: No Control column represents the coefficients (and standard errors) for a linear regression model to predict the valence of police/court encounter of

| | **No Control** | **With Control** | **With Control+Personal Exp.** |
|---|---|---|---|
| **Constant** | -0.09 | -0.44 | -0.21 |
| St.Error | (0.12) | (0.32) | (0.32) |
| **White** | 0.68 | 0.50 | 0.37 |
| St.Error | (0.15) | (0.16) | (0.16) |
| **Latino** | 0.89 | 0.23 | 0.20 |
| St.Error | (0.17) | (0.18) | (0.17) |
| **Asian** | 0.27 | 0.77 | 0.66 |
| St.Error | (0.16) | (0.18) | (0.18) |

Table 2: Valence of Court Encounter Networks

the network of all respondents in table 1/table 2 respectively with binary variables indicating white, latino and asian as predictors; With Control tells the same values for the same models, adding controls predictors (such as age, education, sex, etc.); With Control+Personal Exp. is similar, with added control predictors of the respondents own personal experiences with the court system. Since Black is the excluded group, the coefficients are effective the differences between the valence in the predictor group and Blacks. For example, in table 1, the mean differences in Whites and Blacks for police encounter network is 0.99, with a standard error of 0.12. This is an interesting way to conduct difference-in-mean tests, but it is based on untestable assumptions. First, using linear model means the author assume linear relationship between the valence and each of the predictors, which is unlikely in reality. Second, controlling for covariates using a regression models assume no interaction between the covariates and the main predictors, which is also very unlikely, since for example the fact that a respondent is white is very likely to affect his/her employment status or income, as basics statistics shows correlations between these variables.

However, under the above assumptions, Table 1 and Table 2 do provide some information regarding the differences in the perceived justice of the networks of different populations, regarding the police and the court. Differences between Blacks and White are quite large, around 0.85 even after controlling for other covariates and 0.61 after controlling for personal experiences (for the police) and around 0.50 and 0.37 respectively after control for covariates and personal experiences (for the court). After controls, the differences between Latinos-Blacks is not nearly as much (0.43 compared to 0.85 for the police). Interestingly, the gaps between Asians-Blacks is comparable to the gaps between Whites-Blacks, indicating that the disparity between the Asian and Black population is also very staggering. Another interesting point is that control for covariates change the difference between Asians-Blacks and Latinos-Blacks dramatically, while they do not do much for difference between Whites-Blacks (it remains high even after controls). This might be explained by unobserved covariates that affect Latinos-Blacks and Asians-Blacks but not Whites. This problem will be addressed by sensitivity test in the extension.

### 3.2.2 Extension

As discussed, using regression models for difference in mean tests is interesting but is based on unlikely assumptions. On this section, weattempt to control for covariates using a more general approach: statistical matching. Then, to account for endogeneity, weuse the Rosenbaum's test for hidden bias.

First, using simple Student t-test for difference in means, we receive the results in Table 3 and Table 4. We will now focus our discussion on police perceived justice, because similar results are applicable for court system and will be presented in the appendix.

| | **White** | **Black** | **Latino** | **Asian** |
|---|---|---|---|---|
| **White** | 1 | 1.36e-10 | 5.36e-3 | 6.27e-1 |
| **Black** | 1.36e-10 | 1 | 1.67e-4 | 1.97e-8 |
| **Latino** | 5.36e-3 | 1.67e-4 | 1 | 3.85e-2 |
| **Asian** | 6.27e-17 | 1.97e-8 | 3.85e-2 | 1 |

Table 3: Student t-test Difference-in-mean p-scores, Police Encounter

| | **White** | **Black** | **Latino** | **Asian** |
|---|---|---|---|---|
| **White** | 0 | -0.57 | -0.20 | -0.03 |
| **Black** | 0.57 | 0 | 0.37 | 0.54 |
| **Latino** | 0.20 | -0.37 | 0 | 0.17 |
| **Asian** | 0.03 | -0.54 | -0.17 | 0 |

Table 4: Student t-test Difference-in-mean differences, Police Encounter

The simple t-test results in all statistical significant results, but we need to keep in mind that the population are very differences in all other covariates, and the differences here can be attributed to differences in other variables, rather than just differences in races/ethnicities. However, we do see similar results that concur with the previous replication when looking at Table 4. The valence of Whites-Blacks (0.57) is significantly different than of Latinos-Blacks (0.37), while Asians-Blacks is pretty similar (0.54). This suggests more sophisticated methods are required to link these differences to differences in races. We will now turn our discussion into matching techniques.

Under the assumption of no hidden bias (which will be addressed later using sensitivity test), we will perform matching. Differences for police encounter using different matching methods: propensity scores, Mahalanobis distance and genetic matching are compared in Table 5. In these matching, we recoded the data and assume treatment effect as differences for two populations in each case (for example, coding White==1 and Black==0 enables us to match for other covariates to obtain populations that are similar in those covariates, then take the differences in the outcome of interests: valence of police encounter).

| | Propensity Scores | Mahalanobis | Genetic |
|---|---|---|---|
| **Whites-Blacks** | -0.76 | -0.62 | -0.68 |
| St.Error | (0.17) | (0.14) | (0.15) |
| **Whites-Latinos** | -0.35 | -0.22 | -0.27 |
| St.Error | (0.16) | (0.13) | (0.15) |
| **Whites-Asians** | -0.04 | -0.05 | -0.06 |
| St.Error | (0.15) | (0.13) | (0.14) |

Table 5: Valence of Police Encounter Networks For Matched Population

As Table 5 shows, the differences in Whites-Blacks, Whites-Latinos and Whites-Asians is quite similar to Table 4. However, we tend to believe the matching results more because we effectively controlled for all observed covariates in the matching case. Of all matching technique, genetic matching seems to achieve the best balance across all observed covariates, and it tells us the differences between the police encounter valence is different by around 0.68, which is quite significant compared to other populations. We therefore have reasons to believe the differences is significant, because different methods of difference-in-mean tests indicate the same results.

The sensitivity analysis to hidden bias is performed by using Rosenbaum $\Gamma$ method, with psens() function in the rbounds package. Results are summarized in Table 6.

| $\Gamma$ | P-scores Upper p-value | Mahalanobis Upper p-value | Genetic Upper p-value |
|---|---|---|---|
| 1.0 | 0.0000 | 0.0000 | 0.0000 |
| 1.1 | 0.0000 | 0.0000 | 0.0000 |
| 1.2 | 0.0000 | 0.0000 | 0.0000 |
| 1.3 | 0.0000 | 0.0000 | 0.0000 |
| 1.4 | 0.0000 | 0.0000 | 0.0000 |
| 1.5 | 0.0000 | 0.0002 | 0.0000 |
| 1.6 | 0.0000 | 0.0025 | 0.0006 |
| 1.7 | 0.0000 | 0.0174 | 0.0059 |
| 1.8 | 0.0002 | 0.0737 | 0.0315 |
| 1.9 | 0.0020 | 0.2048 | 0.1083 |
| 2.0 | 0.0125 | 0.4079 | 0.2596 |

Table 6: Sensitivity To Hidden Bias For Police Valence Of Different Matching Methods

Table 6 indicates that even only around $\Gamma = 1.8 - 1.9$, hidden bias was able to significantly change our conclusions. This shows that the results we have here is very sensitive to hidden bias, and this is to be expected because differences in races create a plethora of confoundedness in the populations, and it is unlikely that our observed covariates accounted for all of these hidden bias. This indicates that the replication, and the extension, validity can be come into question and the differences in the valence of police encounter can be attributed to other unobserved factors. However, there are some weights in the conclusions of the original paper because we tried three different methods that led to the same conclusion: pseudo-difference-in-mean by regression analysis (replication); simple t-test and statistical matching (extension). The conclusion, which is African Americans and other minority groups have different views of the justice system than White Americans in the D.C. area, and these views are generally more negative, is therefore quite credible. Of course, with this level of hidden bias sensitivity, any definite policy should be proceeded with much caution.

Another interesting point that the sensitivity analysis reveals was that propensity scores matching actually produces the best guard against hidden bias, among the three. This can be explained by the fact that propensity scores were built by a logistic model that did account for a constant term, which partially attributed to hidden bias that other observed covariates could not account for. Genetic matching, which was run for a population of 200 over 40 generations, did a good job of balancing the observed covariates, but since there is no mechanism to account for unconfoundedness, it is slightly more sensitive to hidden bias.

## 4 Conclusion & Recommendations

The replication of the Modak et al paper and its subsequent extension did produce an unanimous agreement on the negativity of African Americans on the justice system (police and court to be specific) compared to White Americans. We also concluded that differences in these aspects exists between other minority groups, such as Latinos or Asians. The conclusions were slightly dampened by its sensitivity to hidden bias (which is to be expected, since we are dealing with a complex issue such as race/ethnicity that has many social consequences), but in the end is quite credible. The original and the conclusion of the replication and extension give hints to policymaker on diversifying the network of encounters, and be aware that the treatment of individuals in the justice system can have secondhand effect on their acquaintance network.

## References

[1] Rosenbaum, D. P., Amie, M. S., Costello, S. K., Hawkins, D. F., Ring, M. K. (2005). *Attitudes toward the Police: The Effects of Direct and Vicarious Experience.* Retrieved from:
http://www.heinonline.org.ccl.idm.oclc.org/HOL/Page?public=false&handle=hein.journals/policqurt8&page=343&collection=journals

[2] Weitzer, R., Tuch, S. A. (2006). *Race and Policing in America: Conflict and Reform.* Retrieved from:
https://www-journals-uchicago-edu.ccl.idm.oclc.org/doi/abs/10.1086/592520

[3] Tonry, M. (2011). *Punishing Race: A Continuing American Dilemma.* Retrieved from:
http://journals.sagepub.com.ccl.idm.oclc.org/doi/full/10.1177/0094306112438191q

[4] Mondak, J. J., Hurwitz, J., Peffley, M., Testa, P. (2016). *The Vicarious Bases of Perceived Injustice.* Retrieved from:
https://onlinelibrary-wiley-com.ccl.idm.oclc.org/doi/abs/10.1111/ajps.12297

## Appendix

All codes can be accessed here:

https://gist.github.com/AshNguyen/cf42f284f8c9972719fec9311a3ab178

### Replication Code (180 lines, R Studio 2018)

```r
###___Replication Of Paper___###

#Auxiliary t.test function
wtd.t.test.fn<-function(dv="pc_pol_unfair",g1=1,g2=2,...){
  if(dv=="pol"|dv=="ct"){
    # For comparisons of network components
    test<-with(df4[df4$race==g1|df4$race==g2,],
               wtd.t.test(get(dv)[race==g1],get(dv)[race==g2],
                          weight=weight[race==g1],weighty=weight[race==g2],
                          samedata=F))
    return(test)
  }else{
    test<-with(data[data$race==g1|data$race==g2,],
               wtd.t.test(get(dv)[race==g1],get(dv)[race==g2],
                          weight=weight[race==g1],
                          weighty=weight[race==g2],
                          samedata=F))
    return(test)}
}
#Auxilary prediction function (using predict() for previous linear models)
pred.maker<-function(mod,type="A",...){
  pred.df<-groupmeans.df
  pred.df$race<-factor(c("Asians","Blacks","Latinos","Whites"),
                       levels=c("Latinos","Asians","Whites","Blacks"))

  pred.df$white<-0
  pred.df$white[pred.df$race.f=="white"]<-1
  pred.df$latino<-0
  pred.df$latino[pred.df$race.f=="latino"]<-1
  pred.df$asian<-0
  pred.df$asian[pred.df$race.f=="asian"]<-1
  if(type=="A"){
    # Personal or Vicarious
    p<-pred.df
    p[,names(pred.df)[grepl("pc_",names(pred.df))]]<-0
    p[,names(pred.df)[grepl("neg|pos|neut",names(pred.df))]]<-0
    p$fit<-predict(mod,p,se.fit=T)$fit
    p$ll<-predict(mod,p,se.fit=T)$fit -1.96*predict(mod,p,se.fit=T)$se
    p$ul<-predict(mod,p,se.fit=T)$fit +1.96*predict(mod,p,se.fit=T)$se
  }
  if(type=="B"){
    p<-pred.df
    #pB[,names(pred.df)[grepl("pc_",names(pred.df))]]<-0
```

```
    p[,names(pred.df)[grepl("neg|pos|neut",names(pred.df))]]<-0
    p$fit<-predict(mod,p,se.fit=T)$fit
    p$ll<-predict(mod,p,se.fit=T)$fit -1.96*predict(mod,p,se.fit=T)$se
    p$ul<-predict(mod,p,se.fit=T)$fit +1.96*predict(mod,p,se.fit=T)$se
  }
  if(type=="C"){
    p<-pred.df
    p[,names(pred.df)[grepl("pc_",names(pred.df))]]<-0
    #pB[,names(pred.df)[grepl("neg|pos",names(pred.df))]]<-0
    p$fit<-predict(mod,p,se.fit=T)$fit
    p$ll<-predict(mod,p,se.fit=T)$fit -1.96*predict(mod,p,se.fit=T)$se
    p$ul<-predict(mod,p,se.fit=T)$fit +1.96*predict(mod,p,se.fit=T)$se
  }
  if(type=="D"){
    p<-pred.df
    p$fit<-predict(mod,p,se.fit=T)$fit
    p$ll<-predict(mod,p,se.fit=T)$fit -1.96*predict(mod,p,se.fit=T)$se
    p$ul<-predict(mod,p,se.fit=T)$fit +1.96*predict(mod,p,se.fit=T)$se

  }

  palette<-c("#999999","#CCCCCC","#FFFFFF","#333333")
  fig<-ggplot(p,aes(y=fit,x=race,fill=race))+geom_bar(stat="identity")+
    coord_flip()+
    xlab(NULL)+ylab(NULL)+ scale_y_continuous(limits=c(0,7),breaks=0:7,
    labels=0:7)+
    scale_fill_manual(values=palette)+ theme(legend.position="none")+
    geom_errorbar(aes(ymin=ll,ymax=ul),width=.3)
  #+theme(panel.background = element_blank())

  return(p[c(1,4,2,3),c("race.f","fit","ll","ul")])
}

#Auxiliary plot maker function for Figure 1
fig.maker<-function(mod,type="A",...){
  pred.df<-groupmeans.df
  pred.df$race<-factor(c("Asians","Blacks","Latinos","Whites"),
                        levels=c("Latinos","Asians","Whites","Blacks"))

  pred.df$white<-0
  pred.df$white[pred.df$race.f=="white"]<-1
  pred.df$latino<-0
  pred.df$latino[pred.df$race.f=="latino"]<-1
  pred.df$asian<-0
  pred.df$asian[pred.df$race.f=="asian"]<-1
  if(type=="A"){
    # Personal or Vicarious
    p<-pred.df
    p[,names(pred.df)[grepl("pc_",names(pred.df))]]<-0
    p[,names(pred.df)[grepl("neg|pos|neut",names(pred.df))]]<-0
    p$fit<-predict(mod,p,se.fit=T)$fit
    p$ll<-predict(mod,p,se.fit=T)$fit -1.96*predict(mod,p,se.fit=T)$se
    p$ul<-predict(mod,p,se.fit=T)$fit +1.96*predict(mod,p,se.fit=T)$se
  }
```

```
  if(type=="B"){
    p<-pred.df
    #pB[,names(pred.df)[grepl("pc_",names(pred.df))]]<-0
    p[,names(pred.df)[grepl("neg|pos|neut",names(pred.df))]]<-0
    p$fit<-predict(mod,p,se.fit=T)$fit
    p$ll<-predict(mod,p,se.fit=T)$fit -1.96*predict(mod,p,se.fit=T)$se
    p$ul<-predict(mod,p,se.fit=T)$fit +1.96*predict(mod,p,se.fit=T)$se
  }
  if(type=="C"){
    p<-pred.df
    p[,names(pred.df)[grepl("pc_",names(pred.df))]]<-0
    #pB[,names(pred.df)[grepl("neg|pos",names(pred.df))]]<-0
    p$fit<-predict(mod,p,se.fit=T)$fit
    p$ll<-predict(mod,p,se.fit=T)$fit -1.96*predict(mod,p,se.fit=T)$se
    p$ul<-predict(mod,p,se.fit=T)$fit +1.96*predict(mod,p,se.fit=T)$se
  }
  if(type=="D"){
    p<-pred.df
    p$fit<-predict(mod,p,se.fit=T)$fit
    p$ll<-predict(mod,p,se.fit=T)$fit -1.96*predict(mod,p,se.fit=T)$se
    p$ul<-predict(mod,p,se.fit=T)$fit +1.96*predict(mod,p,se.fit=T)$se

  }

 #palette<-c("#999999","#CCCCCC","#FFFFFF","#333333")
  fig<-ggplot(p,aes(y=fit,x=race,col=race))+geom_point(size=2)+coord_flip()
  +
    xlab(NULL)+ylab(NULL)+ ylim(3.5,6.5)+
    scale_color_grey(start = 0.8, end = 0.3)+    theme_bw()+theme(legend.
    position="none")+
    geom_errorbar(aes(ymin=ll,ymax=ul),width=.3)


  print(p[,c("race.f","fit","ll","ll")])
  return(fig)
}

#Set working directory and load/clean data
setwd("/Users/ash/Downloads/ashk")
load("ajpsdata.rda")
data<-data[data$validzip==1,] #excluding invalid ZIP codes
pacman::p_load("knitr","weights","car","ggplot2","grid","plyr",
               "gridExtra","stargazer","xtable")


#Replication Table 2: Simple lm() for vicarious police/court encounter
'
Police
'
# Baseline
m2p.1.mn<-lm(polnet_mn_nz~white+asian+latino,data,weights=weight)
summary(m2p.1.mn)
# Controls
m2p.2.mn<-lm(polnet_mn_nz~white+asian+latino+
```

```
                   age+female01+education+married+fullem+
                   relig_imp+church_atd+pid+pidDKNA01
                +income_with_PNS+PNS_income01+Pop4Q+disc_count,
                data, weights=weight)
summary(m2p.2.mn)
# Personal Experience
m2p.3.mn<-lm(polnet_mn_nz~white+asian+latino+
                   age+female01+education+married+fullem+
                   relig_imp+church_atd+
                   pid+pidDKNA01+income_with_PNS+PNS_income01
                +Pop4Q+disc_count
                +pc_pol_rude+pc_pol_unfair, data, weights=weight)
summary(m2p.3.mn)
'
Court
'
# Baseline
m2c.1.mn<-lm(ctnet_mn_nz~white+asian+latino, data, weights=weight)
summary(m2c.1.mn)
# Controls
m2c.2.mn<-lm(ctnet_mn_nz~white+asian+latino+
                   age+female01+education+married+fullem+
                   relig_imp+church_atd+pid+pidDKNA01+
                   income_with_PNS+PNS_income01+
                   Pop4Q+disc_count, data, weights=weight)
summary(m2c.2.mn)
# Personal Experience
m2c.3.mn<-lm(ctnet_mn_nz~white+asian+latino+
                   age+female01+education+married+fullem+relig_imp+church_atd+
                   pid+pidDKNA01+income_with_PNS+PNS_income01
                +Pop4Q+disc_count
                +pc_ct_rude+pc_ct_unfair, data, weights=weight)
summary(m2c.3.mn)
```

**Extension Code (247 lines, R Studio 2018)**

```
###___Extension Of Paper___###

#Auxiliary imputer function for numeric NAs
num_imputer<-function(X){
  for (j in c(1:ncol(X))){
    if (class(X[,j])=='numeric'){
      im=median(X[,j],na.rm=TRUE)
      for (i in c(1:length(X[,j]))){
        if (is.na(X[i,j])==TRUE){
          X[i,j]=im
        }
      }
    }
  }
  return(X)
}

#Load/clean data and relevant packages
```

```
load("/Users/ash/Downloads/ashk/ajpsdata.rda")
library(Matching)


#Simple t-test between different populations for negative police/court experiences
'
Police
'
coef=matrix(data=0,nrow=4,ncol=4)
colnames(coef)=c('White','Black','Latino','Asia')
rownames(coef)=c('White','Black','Latino','Asia')
diff=matrix(data=0,nrow=4,ncol=4)
colnames(diff)=c('White','Black','Latino','Asia')
rownames(diff)=c('White','Black','Latino','Asia')
for (i in c(1,2,3,4)){
  for (j in c(1,2,3,4)){
    result=wtd.t.test.fn("pc_pol_unfair",i,j)
    coef[i,j]=result$coefficients[3]
    diff[i,j]=result$additional[1]
  }
}
'
Court
'
coef=matrix(data=0,nrow=4,ncol=4)
colnames(coef)=c('White','Black','Latino','Asia')
rownames(coef)=c('White','Black','Latino','Asia')
diff=matrix(data=0,nrow=4,ncol=4)
colnames(diff)=c('White','Black','Latino','Asia')
rownames(diff)=c('White','Black','Latino','Asia')
for (i in c(1,2,3,4)){
  for (j in c(1,2,3,4)){
    result=wtd.t.test.fn("pc_ct_unfair",i,j)
    coef[i,j]=result$coefficients[3]
    diff[i,j]=result$additional[1]
  }
}

#Using matching to build controlled data
pop_white=subset(data,white==1)
pop_white$treat=1
pop_white<-num_imputer(pop_white)
pop_black=subset(data,black==1)
pop_black$treat=0
pop_black<-num_imputer(pop_black)
pop_latino=subset(data,latino==1)
pop_latino$treat=0
pop_latino<-num_imputer(pop_latino)
pop_asian=subset(data,asian==1)
pop_asian$treat=0
pop_asian<-num_imputer(pop_asian)

pop_w1b0=rbind(pop_white,pop_black)
```

```
pop_w1l0=rbind(pop_white,pop_latino)
pop_w1a0=rbind(pop_white,pop_asian)

#Matching using propensity scores; Mahalanobis distance & GenMatch() [white versus black]
'
Police
'
pmodel=glm(treat~education+age+female01+pid+pidDKNA01+income_with_PNS+PNS_income01
            +relig_imp+church_atd+married+fullem,data=pop_w1b0,family=binomial)
pscores=predict(pmodel,type='response')
Y=pop_w1b0$polnet_mn
X=cbind(pop_w1b0$age,pop_w1b0$education,pop_w1b0$female01,pop_w1b0$pid,
          pop_w1b0$pidDKNA01,pop_w1b0$income_with_PNS,pop_w1b0$PNS_income01,
          pop_w1b0$relig_imp,pop_w1b0$church_atd,pop_w1b0$married,
          pop_w1b0$fullem)
Tr=pop_w1b0$treat
matched_p=Match(Y=Y,X=pscores,Tr=Tr,estimand='ATE')
matched_M=Match(Y=Y,X=X,Tr=Tr,estimand='ATE',Weight=2)
w<-GenMatch(Tr=Tr,X=X,estimand='ATE',pop.size = 200,
              max.generations=40,wait.generations=3,
              hard.generation.limit=TRUE)
matched_G=Match(Y=Y,X=X,Tr=Tr,estimand='ATE',Weight.matrix=w)
MatchBalance(treat~education+age+female01+pid+pidDKNA01+income_with_PNS+PNS_income01
              +relig_imp+church_atd+married+fullem,data=pop_w1b0,match.out=matched_G,nboots=500)
cat('Estimate difference for propensity score matching:', matched_p$est)
cat('Estimate difference for Mahalanobis matching:', matched_M$est)
cat('Estimate difference for genetic matching:', matched_G$est)
"
Sensity to hidden bias
"
psens(matched_p,Gamma=2,GammaInc=0.1)
psens(matched_M,Gamma=2,GammaInc=0.1)
psens(matched_G,Gamma=2,GammaInc=0.1)
'
Court
'
pmodel=glm(treat~education+age+female01+pid+pidDKNA01+income_with_PNS+PNS_income01
            +relig_imp+church_atd+married+fullem,data=pop_w1b0,family=binomial)
pscores=predict(pmodel,type='response')
Y=pop_w1b0$ctnet_mn
X=cbind(pop_w1b0$age,pop_w1b0$education,pop_w1b0$female01,pop_w1b0$pid,
          pop_w1b0$pidDKNA01,pop_w1b0$income_with_PNS,pop_w1b0$PNS_income01,
          pop_w1b0$relig_imp,pop_w1b0$church_atd,pop_w1b0$married,
          pop_w1b0$fullem)
Tr=pop_w1b0$treat
matched_p=Match(Y=Y,X=pscores,Tr=Tr,estimand='ATE')
matched_M=Match(Y=Y,X=X,Tr=Tr,estimand='ATE',Weight=2)
```

```
w<-GenMatch(Tr=Tr,X=X,estimand='ATE')
matched_G=Match(Y=Y,X=X,Tr=Tr,estimand='ATE',Weight.matrix=w)
MatchBalance(treat~education+age+female01+pid+pidDKNA01+income_with_PNS+PNS_income01
             +relig_imp+church_atd+married+fullem,data=pop_w1b0,match.out=matched_G,nboots=500)
cat('Estimate difference for propensity score matching:', matched_p$est)
cat('Estimate difference for Mahalanobis matching:', matched_M$est)
cat('Estimate difference for genetic matching:', matched_G$est)
"
Sensity to hidden bias
"
psens(matched_p)
psens(matched_M)
psens(matched_G)

#Matching using propensity scores; Mahalanobis distance & GenMatch() [white versus latino]
'
Police
'
pmodel=glm(treat~education+age+female01+pid+pidDKNA01+income_with_PNS+PNS_income01
           +relig_imp+church_atd+married+fullem,data=pop_w1l0,family=binomial)
pscores=predict(pmodel,type='response')
Y=pop_w1l0$polnet_mn
X=cbind(pop_w1l0$age,pop_w1l0$education,pop_w1l0$female01,pop_w1l0$pid,
        pop_w1l0$pidDKNA01,pop_w1l0$income_with_PNS,pop_w1l0$PNS_income01,
        pop_w1l0$relig_imp,pop_w1l0$church_atd,pop_w1l0$married,
        pop_w1l0$fullem)
Tr=pop_w1l0$treat
matched_p=Match(Y=Y,X=pscores,Tr=Tr,estimand='ATE')
matched_M=Match(Y=Y,X=X,Tr=Tr,estimand='ATE',Weight=2)
w<-GenMatch(Tr=Tr,X=X,estimand='ATE',pop.size = 200,
            max.generations=40,wait.generations=3,
            hard.generation.limit=TRUE)
matched_G=Match(Y=Y,X=X,Tr=Tr,estimand='ATE',Weight.matrix=w)
MatchBalance(treat~education+age+female01+pid+pidDKNA01+income_with_PNS+PNS_income01
             +relig_imp+church_atd+married+fullem,data=pop_w1l0,match.out=matched_G,nboots=500)
cat('Estimate difference for propensity score matching:', matched_p$est)
cat('Estimate difference for Mahalanobis matching:', matched_M$est)
cat('Estimate difference for genetic matching:', matched_G$est)
"
Sensity to hidden bias
"
psens(matched_p)
psens(matched_M)
psens(matched_G)
'
Court
'
```

```r
pmodel=glm(treat~education+age+female01+pid+pidDKNA01+income_with_PNS+PNS_income01
          +relig_imp+church_atd+married+fullem,data=pop_w1l0,family=binomial)
pscores=predict(pmodel,type='response')
Y=pop_w1l0$ctnet_mn
X=cbind(pop_w1l0$age,pop_w1l0$education,pop_w1l0$female01,pop_w1l0$pid,
        pop_w1l0$pidDKNA01,pop_w1l0$income_with_PNS,pop_w1l0$PNS_income01,
        pop_w1l0$relig_imp,pop_w1l0$church_atd,pop_w1l0$married,
        pop_w1l0$fullem)
Tr=pop_w1l0$treat
matched_p=Match(Y=Y,X=pscores,Tr=Tr,estimand='ATE')
matched_M=Match(Y=Y,X=X,Tr=Tr,estimand='ATE',Weight=2)
w<-GenMatch(Tr=Tr,X=X,estimand='ATE')
matched_G=Match(Y=Y,X=X,Tr=Tr,estimand='ATE',Weight.matrix=w)
MatchBalance(treat~education+age+female01+pid+pidDKNA01+income_with_PNS+PNS_income01
            +relig_imp+church_atd+married+fullem,data=pop_w1l0,match.out=matched_G,nboots=500)
cat('Estimate difference for propensity score matching:', matched_p$est)
cat('Estimate difference for Mahalanobis matching:', matched_M$est)
cat('Estimate difference for genetic matching:', matched_G$est)
"
Sensity to hidden bias
"
psens(matched_p)
psens(matched_M)
psens(matched_G)

#Matching using propensity scores; Mahalanobis distance & GenMatch() [white versus asian]
'
Police
'
pmodel=glm(treat~education+age+female01+pid+pidDKNA01+income_with_PNS+PNS_income01
          +relig_imp+church_atd+married+fullem,data=pop_w1a0,family=binomial)
pscores=predict(pmodel,type='response')
Y=pop_w1a0$polnet_mn
X=cbind(pop_w1a0$age,pop_w1a0$education,pop_w1a0$female01,pop_w1a0$pid,
        pop_w1a0$pidDKNA01,pop_w1a0$income_with_PNS,pop_w1a0$PNS_income01,
        pop_w1a0$relig_imp,pop_w1a0$church_atd,pop_w1a0$married,
        pop_w1a0$fullem)
Tr=pop_w1a0$treat
matched_p=Match(Y=Y,X=pscores,Tr=Tr,estimand='ATE')
matched_M=Match(Y=Y,X=X,Tr=Tr,estimand='ATE',Weight=2)
w<-GenMatch(Tr=Tr,X=X,estimand='ATE',pop.size = 200,
            max.generations=40,wait.generations=3,
            hard.generation.limit=TRUE)
matched_G=Match(Y=Y,X=X,Tr=Tr,estimand='ATE',Weight.matrix=w)
MatchBalance(treat~education+age+female01+pid+pidDKNA01+income_with_PNS+PNS_income01
            +relig_imp+church_atd+married+fullem,data=pop_w1a0,match.out=
```

```
matched_G, nboots=500)
cat('Estimate difference for propensity score matching:', matched_p$est)
cat('Estimate difference for Mahalanobis matching:', matched_M$est)
cat('Estimate difference for genetic matching:', matched_G$se)
"
Sensity to hidden bias
"
psens(matched_p)
psens(matched_M)
psens(matched_G)
'
Court
'
pmodel=glm(treat~education+age+female01+pid+pidDKNA01+income_with_PNS+PNS_income01
        +relig_imp+church_atd+married+fullem,data=pop_w1a0,family=binomial)
pscores=predict(pmodel,type='response')
Y=pop_w1a0$ctnet_mn
X=cbind(pop_w1a0$age,pop_w1a0$education,pop_w1a0$female01,pop_w1a0$pid,
        pop_w1a0$pidDKNA01,pop_w1a0$income_with_PNS,pop_w1a0$PNS_income01,
        pop_w1a0$relig_imp,pop_w1a0$church_atd,pop_w1a0$married,
        pop_w1a0$fullem)
Tr=pop_w1a0$treat
matched_p=Match(Y=Y,X=pscores,Tr=Tr,estimand='ATE')
matched_M=Match(Y=Y,X=X,Tr=Tr,estimand='ATE',Weight=2)
w<-GenMatch(Tr=Tr,X=X,estimand='ATE')
matched_G=Match(Y=Y,X=X,Tr=Tr,estimand='ATE',Weight.matrix=w)
MatchBalance(treat~education+age+female01+pid+pidDKNA01+income_with_PNS+PNS_income01
          +relig_imp+church_atd+married+fullem,data=pop_w1a0,match.out=matched_G, nboots=500)
cat('Estimate difference for propensity score matching:', matched_p$est)
cat('Estimate difference for Mahalanobis matching:', matched_M$est)
cat('Estimate difference for genetic matching:', matched_G$est)
"
Sensity to hidden bias
"
psens(matched_p)
psens(matched_M)
psens(matched_G)
```