Discrimination

GV482 Problem Set - Game Theory

Dianyi Yang

2024-02-29

1 Preliminaries (not covered in the seminar)

(a) Explain briefly why the real effect of being Black on police violence is 2%.

Controlling for criminal records, the difference in police violence between the two groups is 2%.

(b) Show that the researcher overestimates the effect of being Black on police violence if $x_B < x_W$ (i.e., Whites would be discriminated at the level of stop) and underestimates it if $x_B > x_W$ (i.e., Blacks would be discriminated at the level of stop as it is most likely).

$$
\begin{aligned}
\widehat{ATE}_{\text{Black}} &= 4\% x_B + 12\%(1 - x_B) - 2\% x_W - 10\%(1 - x_W) \\
&= 0.12 - 0.08 x_B - 0.1 + 0.08 x_W \\
&= 0.02 + 0.08(x_W - x_B) \\
\text{Bias} &= 0.08(x_W - x_B)
\end{aligned}
$$

(c) Explain why we should think of the effects (rather than the effect) of race on police violence.

1. Differential stops
2. Heterogeneous effect of criminal record
3. Socio-economic disadvantages that led to higher crime rates.

(d) Suppose that the proportion of individuals stopped from each group is equal to their proportion in the population (that is, there is no discrimination at the level of stops). Explain why the estimated effect is higher than any real effect of being Black on political violence whenever $\alpha_W \geq \frac{1}{6} + \alpha_B \frac{5}{6}$.

$$
\begin{aligned}
\text{Bias} &= \widehat{ATE} - ATE \\
&= [7\%\alpha_B + 12\%(1-\alpha_B) - 4\%\alpha_W - 10\%(1-\alpha_W)] + [(7\% - 4\%)\frac{\alpha_W + \alpha_B}{2} - (12\% - 10\%)(\frac{1-\alpha_W}{} \\
&= [12\% - 5\%\alpha_B - 10\% + 6\%\alpha_W] - [\frac{3}{2}\%(\alpha_W + \alpha_B) - 2\% + 1\%(\alpha_W + \alpha_B)] \\
&= 3.5\%\alpha_W - 7.5\%\alpha_B > 0
\end{aligned}
$$

2 The model

Part A - Police behaviour

In this first part, we think to determine the police officer's behavior as a function of his information: $v^*(s^c_{I_g}, s^v_{I_g})$. To do so, we first consider what the signals tell P about the civilian's level of criminality and propensity for violence.

Q1

In this question, we seek to understand what the police officer can infer from the signals. To answer these questions, you need to use Bayes' rule.

(a) Show that the officer's posterior after signal $s^c_{I_g} = 1$ satisfies: $Pr(\theta^c_{I_g} = 1 | s^c_{I_g} = 1) = 1$.

$$
\begin{aligned}
Pr(\theta^c_{I_g} = 1 | s^c_{I_g} = 1) &= \frac{Pr(\theta^c_{I_g} = 1) Pr(s^c_{I_g} = 1 | \theta^c_{I_g} = 1)}{Pr(\theta^c_{I_g} = 1) Pr(s^c_{I_g} = 1 | \theta^c_{I_g} = 1) + Pr(\theta^c_{I_g} = -1) Pr(s^c_{I_g} = 1 | \theta^c_{I_g} = -1)} \\
&= \frac{\pi^c_g q^c}{\pi^c_g q^c + (1 - \pi^c_g) \times 0} \\
&= 1
\end{aligned}
$$

(b) Show that the officer's posterior after signal $s^c_{I_g} = -1$ satisfies: $Pr(\theta^c_{I_g} = 1 | s^c_{I_g} = -1) = 0$.

$$
\begin{aligned}
Pr(\theta^c_{I_g} = 1 | s^c_{I_g} = -1) &= \frac{Pr(\theta^c_{I_g} = 1) Pr(s^c_{I_g} = -1 | \theta^c_{I_g} = 1)}{Pr(\theta^c_{I_g} = 1) Pr(s^c_{I_g} = -1 | \theta^c_{I_g} = 1) + Pr(\theta^c_{I_g} = -1) Pr(s^c_{I_g} = -1 | \theta^c_{I_g} = -1)} \\
&= \frac{\pi^c_g \times 0}{\pi^c_g \times 0 + (1 - \pi^c_g) \times q^c} \\
&= 0
\end{aligned}
$$

(c) Show that the officer's posterior after signal $s^c_{I_g} = 0$ satisfies: $Pr(\theta^c_{I_g} = 1 | s^c_{I_g} = 0) = \pi^c_g$.

$$
\begin{aligned}
Pr(\theta^c_{I_g} = 1 | s^c_{I_g} = 0) &= \frac{Pr(\theta^c_{I_g} = 1) Pr(s^c_{I_g} = 0 | \theta^c_{I_g} = 1)}{Pr(\theta^c_{I_g} = 1) Pr(s^c_{I_g} = 0 | \theta^c_{I_g} = 1) + Pr(\theta^c_{I_g} = -1) Pr(s^c_{I_g} = 0 | \theta^c_{I_g} = -1)} \\
&= \frac{\pi^c_g \times (1 - q^c)}{\pi^c_g \times (1 - q^c) + (1 - \pi^c_g) \times (1 - q^c)} \\
&= \frac{\pi^c_g}{\pi^c_g + (1 - \pi^c_g)} \\
&= \pi^c_g
\end{aligned}
$$

(d) Show that the police officer's posterior regarding the propensity for violence ($\theta^v_{I_g}$) satisfies:

$$
\begin{aligned}
Pr(\theta^v_{I_g} = 1 | s^v_{I_g} = 1) &= 1 \\
Pr(\theta^v_{I_g} = 1 | s^v_{I_g} = 0) &= \pi^v_g \\
Pr(\theta^v_{I_g} = 1 | s^v_{I_g} = -1) &= 0
\end{aligned}
$$

Same as above.

Q2

(a) Explain briefly, but carefully why: $v^*(s^c_{I_g}, s^v_{I_g}) = E_{\theta^c_{I_g}, \theta^v_{I_g}}(\alpha \theta^c_{I_g} + \theta^v_{I_g} | s^c_{I_g}, s^v_{I_g}) + b^v_g$.

Comment: The expression $E_{\theta^c_{I_g}, \theta^v_{I_g}}$ means that we are taking the expectations over the random variables $\theta^c_{I_g}, \theta^v_{I_g}$, which are the unknown from P's perspective.

Take FOC of $E_{\theta^c_{I_g}, \theta^v_{I_g}}(U_P)$ with respect to v:

$$
\begin{aligned}
E_{\theta^c_{I_g}, \theta^v_{I_g}}[-2(v - (\alpha \theta^c_{I_g} + \theta^v_{I_g} + b^v_g)) | s^c_{I_g}, s^v_{I_g}] &= 0 \\
E_{\theta^c_{I_g}, \theta^v_{I_g}}[v - (\alpha \theta^c_{I_g} + \theta^v_{I_g} + b^v_g) | s^c_{I_g}, s^v_{I_g}] &= 0 \\
v^* &= E_{\theta^c_{I_g}, \theta^v_{I_g}}(\alpha \theta^c_{I_g} + \theta^v_{I_g} + b^v_g | s^c_{I_g}, s^v_{I_g}) \\
v^* &= E_{\theta^c_{I_g}, \theta^v_{I_g}}(\alpha \theta^c_{I_g} + \theta^v_{I_g} | s^c_{I_g}, s^v_{I_g}) + b^v_g
\end{aligned}
$$

(b) Explain briefly, but carefully why we can rewrite: $v^*(s^c_{I_g}, s^v_{I_g}) = \alpha E_{\theta^c_{I_g},\theta^v_{I_g}}(\theta^c_{I_g}|s^c_{I_g}) + E_{\theta^c_{I_g},\theta^v_{I_g}}(\theta^v_{I_g}|s^v_{I_g}) + b^v_g$.

$$
\begin{aligned}
v^* &= E_{\theta^c_{I_g},\theta^v_{I_g}}(\alpha\theta^c_{I_g} + \theta^v_{I_g}|s^c_{I_g}, s^v_{I_g}) + b^v_g \\
&= \alpha E_{\theta^c_{I_g},\theta^v_{I_g}}(\theta^c_{I_g}|s^c_{I_g}, s^v_{I_g}) + E_{\theta^c_{I_g},\theta^v_{I_g}}(\theta^v_{I_g}|s^c_{I_g}, s^v_{I_g}) + b^v_g \\
&= \alpha E_{\theta^c_{I_g},\theta^v_{I_g}}(\theta^c_{I_g}|s^c_{I_g}) + E_{\theta^c_{I_g},\theta^v_{I_g}}(\theta^v_{I_g}|s^v_{I_g}) + b^v_g
\end{aligned}
$$

independence of signals

(c) Using your answer to Q1, explain briefly why

$$
\begin{aligned}
E_{\theta^v_{I_g}}(\theta^v_{I_g}|s^v_{I_g} = 1) &= 1 \\
E_{\theta^v_{I_g}}(\theta^v_{I_g}|s^v_{I_g} = 0) &= 2\pi^c_g - 1 \\
E_{\theta^v_{I_g}}(\theta^v_{I_g}|s^v_{I_g} = -1) &= -1 \\
E_{\theta^c_{I_g}}(\theta^c_{I_g}|s^c_{I_g} = 1) &= 1 \\
E_{\theta^c_{I_g}}(\theta^c_{I_g}|s^c_{I_g} = 0) &= 2\pi^v_g - 1 \\
E_{\theta^c_{I_g}}(\theta^c_{I_g}|s^c_{I_g} = -1) &= -1
\end{aligned}
$$

Most are uncontroversial as θ can only take 0 and 1/-1 given the signal. The $2\pi^v_g - 1$ term comes from:

$$
\begin{aligned}
E_{\theta^c_{I_g}}(\theta^c_{I_g}|s^c_{I_g} = 0) &= 1Pr(\theta^c_{I_g} = 1|s^c_{I_g} = 0) + (-1)Pr(\theta^c_{I_g} = -1|s^c_{I_g} = 0) \\
&= \pi^c_g - (1 - \pi^c_g) \\
&= 2\pi^c_g - 1
\end{aligned}
$$

Part B - The decision whether to stop the civilian

Q3

We can now turn to the decision whether to stop the civilian. We denote $A_g(s^c_{I_g})$ P's expected utility from stopping I_g after receiving signal $s^c_{I_g}$ (remember that $s^v_{I_g}$ is received after stopping so cannot be condition his decision on it at this stage). We will compute it in this question. Note that this question is algebra intensive. It is not too hard, but it is quite tedious. If you hate algebra, you may want to skip it, though it is always good practice (for your exams, for your future life) to try to answer this sort of questions.

(a) Show that $-(v^*(s^c_{I_g}, s^v_{I_g}) - b^v_g - \theta^v_{I_g} - \alpha\theta^c_{I_g})^2 = -(E(\theta^v_{I_g}|s^v_{I_g}) - \theta^v_{I_g})^2 - \alpha^2(E(\theta^c_{I_g}|s^c_{I_g}) - \theta^c_{I_g})^2 - 2\alpha(E(\theta^v_{I_g}|s^v_{I_g}) - \theta^v_{I_g})(E(\theta^c_{I_g}|s^c_{I_g}) - \theta^c_{I_g})$.

$$
\begin{aligned}
& -(v^*(s^c_{I_g}, s^v_{I_g}) - b^v_g - \theta^v_{I_g} - \alpha\theta^c_{I_g})^2 \\
=& -((\alpha E(\theta^c_{I_g}|c^v_{I_g}) + E(\theta^v_{I_g}|c^v_{I_g}) + b^v_g) - b^v_g - \theta^v_{I_g} - \alpha\theta^c_{I_g})^2 \\
=& -(\alpha(E(\theta^c_{I_g}|c^v_{I_g}) - \theta^c_{I_g}) + E(\theta^v_{I_g}|c^v_{I_g}) - \theta^v_{I_g})^2 \\
=& -(E(\theta^v_{I_g}|s^v_{I_g}) - \theta^v_{I_g})^2 - \alpha^2(E(\theta^c_{I_g}|s^c_{I_g}) - \theta^c_{I_g})^2 - 2\alpha(E(\theta^v_{I_g}|s^v_{I_g}) - \theta^v_{I_g})(E(\theta^c_{I_g}|s^c_{I_g}) - \theta^c_{I_g})
\end{aligned}
$$

(b) Show that $E_{\theta^c_{I_g},\theta^v_{I_g}}[-(E(\theta^v_{I_g}|s^v_{I_g}) - \theta^v_{I_g})^2 - \alpha^2(E(\theta^c_{I_g}|s^c_{I_g}) - \theta^c_{I_g})^2 - 2\alpha(E(\theta^v_{I_g}|s^v_{I_g}) - \theta^v_{I_g})(E(\theta^c_{I_g}|s^c_{I_g}) - \theta^c_{I_g})|\theta^c_{I_g}, \theta^v_{I_g}] = -Var(\theta^v_{I_g}|s^v_{I_g}) - \alpha^2 Var(\theta^c_{I_g}|s^c_{I_g})$.

$$
\begin{aligned}
& E_{\theta^c_{I_g},\theta^v_{I_g}}[-(E(\theta^v_{I_g}|s^v_{I_g}) - \theta^v_{I_g})^2 - \alpha^2(E(\theta^c_{I_g}|s^c_{I_g}) - \theta^c_{I_g})^2 - 2\alpha(E(\theta^v_{I_g}|s^v_{I_g}) - \theta^v_{I_g})(E(\theta^c_{I_g}|s^c_{I_g}) - \theta^c_{I_g})|\theta^c_{I_g}, \theta^v_{I_g}] \\
=& -Var(\theta^v_{I_g}|s^v_{I_g}) - \alpha^2 Var(\theta^c_{I_g}|s^c_{I_g}) - 2\alpha(0)(0) \\
=& -Var(\theta^v_{I_g}|s^v_{I_g}) - \alpha^2 Var(\theta^c_{I_g}|s^c_{I_g})
\end{aligned}
$$

(c) Show that

$Var(\theta^v_{I_g}|s^v_{I_g} = 1) = 0$

Because $P(\theta^v_{I_g} = 1|s^v_{I_g} = 1) = 1$.

$Var(\theta^v_{I_g}|s^v_{I_g} = 0) = 4\pi^v_g(1 - \pi^v_g)$

$$
\begin{aligned}
& Var(\theta^v_{I_g}|s^v_{I_g}=0) \\
=& E[\theta^{v2}_{I_g}|s^v_{I_g}=0] - E[\theta^v_{I_g}|s^v_{I_g}=0]^2 \\
=& [1Pr(\theta^{v2}=1_{I_g}|s^v_{I_g}=0) + 1Pr(\theta^{v2}_{I_g}=1|s^v_{I_g}=0)] - (2\pi^v_g-1)^2 \\
=& 1-(4\pi^{v2}_g - 4\pi^v_g + 1) \\
=& 4\pi^v_g(\pi^v_g - 1)
\end{aligned}
$$

$Var(\theta^v_{I_g}|s^v_{I_g}=-1)=0$

Because $P(\theta^v_{I_g}=-1|s^v_{I_g}=-1)=1$.

Part C - The researchers' problem

Q5

In this question, we compare the average violence against members of group M and members of group m, which we denote $\bar{V_M}$ and $\bar{V_m}$, respectively. We, first, need to compute these expected level of violence.

(a) We start with a civilian from the majority group. Recall that a citizen from group M is stopped only if $s^c_{I_M}=1$. Show that

$$\bar{V_M} = \alpha + (2\pi^v_M - 1)$$

$$
\begin{aligned}
\bar{V_M} &= Pr(s^v_M=1)(\alpha+1+b^v_M) + Pr(s^v_M=0)(\alpha+2\pi^v_M-1+b^v_g) + Pr(s^v_M=-1)(\alpha-1+b^v_g) \\
&= \pi^v_M q^v(\alpha+1+b^v_M) + (\pi^v_M(1-q^v)+(1-\pi^v_M)(1-q^v))(\alpha+2\pi^v_M-1+b^v_M) + (1-\pi^v_M)q^v(\alpha-1+b^v_g \\
&= \pi^v_M q^v(\alpha+1+b^v_M) + (1-q^v)(\alpha+2\pi^v_M-1+b^v_M) + (1-\pi^v_M)q^v(\alpha-1+b^v_M) \\
&= \alpha + \pi^v_M q^v + (1-q^v)(2\pi^v_M-1) - (1-\pi^v_M)q^v + b^v_M \\
&= \alpha + \pi^v_M q^v + 2\pi^v_M - 1 - 2\pi^v_M q^v + q^v - q^v + \pi^v_M q^v + b^v_M \\
&= \alpha + 2\pi^v_M - 1 + b^v_M \\
&= \alpha + 2\pi^v_M - 1
\end{aligned}
$$

(b) We now turn to $\bar{V_m}$, which will prove slightly more complicated. Recall that a member of group m is stopped whenever she produces signal $s^c_{I_m}=1$ or $s^c_{I_m}=0$. Explain why the proportion of minority individuals with signal $s^c_{I_m}=1$ among those arrested is: $\frac{\pi^c_m q^c}{\pi^c_m q^c + 1 - q^c}$.

Assume only those give a signal of $s^v_{I_g}$ are arrested.

$$\text{prop} =$$