Problem A.1. Corremos la regresión discontinua

$$HomicideRate_m = \beta_0 + \beta_1 PANwin_m + \beta_2 PANwin_m \times VoteDif_m +$$

$$\beta_3(1 - PANwin_m) \times VoteDif_m + \epsilon_m$$

al hecerlo vemos que el estimador de β_1 equivale a 9.33 puntos porcentuales, que representa el efecto promedio de aumento en la tasa de homicidio en una municipalidad en que ganó el PAN.

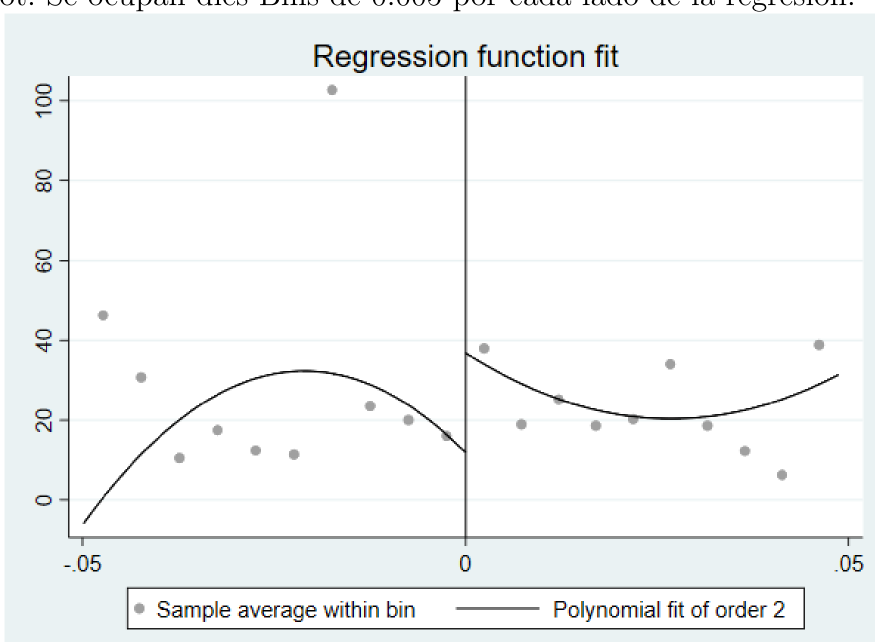
Problem A.2. Si en vez la función es de

$$HomicideRate_m = \beta_0 + \beta_1 PANwin_m + \beta_2 PANwin_m \times VoteDif_m + \beta_3 PANwin_m \times VoteDif_m^2 +$$

$$\beta_4(1 - PANwin_m) \times VoteDif_m + \beta_5(1 - PANwin_m) \times VoteDif_m^2 \epsilon_m$$

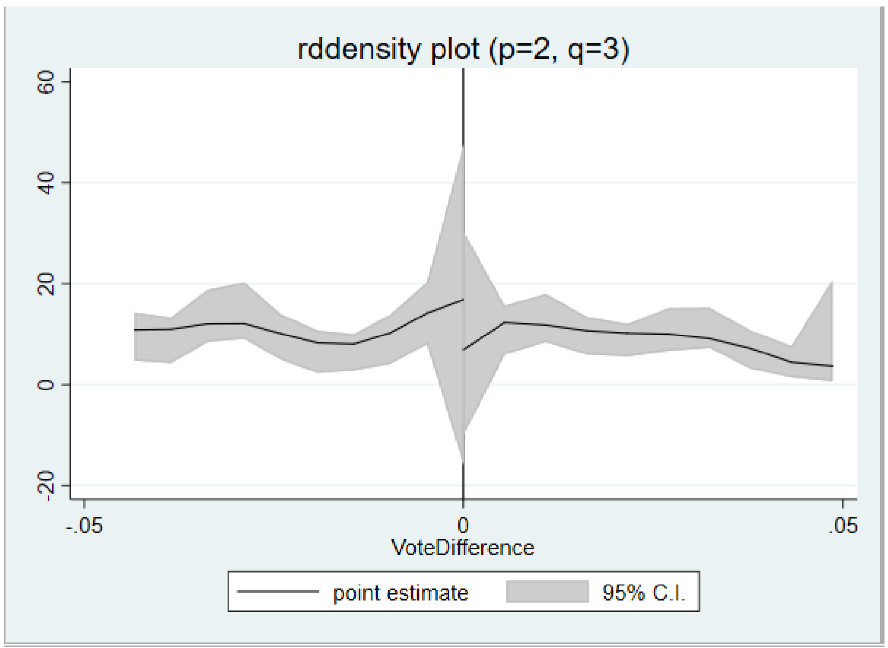
El efecto de la victoria del PAN en la tasa de delincuencia es de 10.97 puntos porcentuales, poco mas de punto y medio mas que si hacemos la regresión linealmente.

Problem A.3. Generamos un grafico similar al grafico 4B de Dell ocupando el comando RDplot. Se ocupan dies Bins de 0.005 por cada lado de la regresión.



Problem A.4. El estimador de regresion discontinua es inherentemente local, pues mide la diferencia solo despues de un evento especifico. Por esto, las estimaciones de RD deben fijar un area pequeña al rededor del evento para juzgar las diferencias a ambos lados. En este caso, como queremos ver carreras politicas con resultados muy cercanos, se fija el intervalo en 0.05 pues significa que en cada observacion, el PAN obtuvo entre 45 % y 55 % de los votos.

Problem A.5. El test de McCrary consiste en un analisis de la densidad de la variable de corrida al rededor del punto de quiebre. Si se existiese alguna forma de manipulación exogena de los datos, este metodo nos lo mostraría. En este caso, el p-valor de la hipotesis .ªmbos lados del punto de quiebre tienen igual comportamiento.ᵉˢ de 77 % por lo que no podemos rechazarla y es creible que no hubo manipulación previa. La existencia de esta manipulacion es improbable dada la naturaleza de los datos pues implicaría que alguien es capaz de seleccionar los margenes de vistoria de los candidatos de las elecciones a nivel municipal.

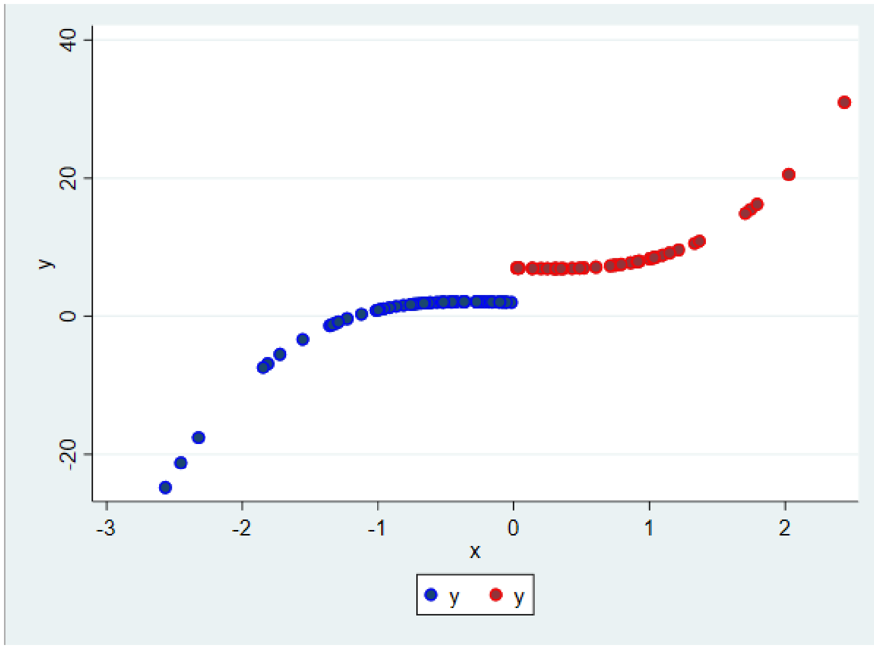


Problem B.1. Generamos la 100 observaciones que sigan la función cúbica

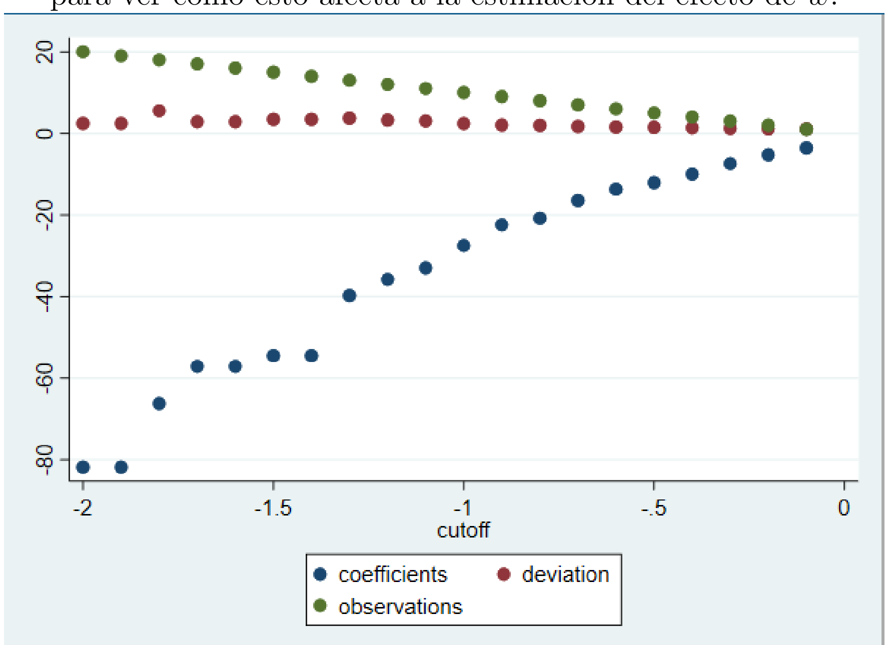
$$y = 2 - 0{,}5x + 0{,}1x^2 + 1{,}7x^3 + 5w + \epsilon$$

con x una variabale aleatoria de comportamiento normal estandar.

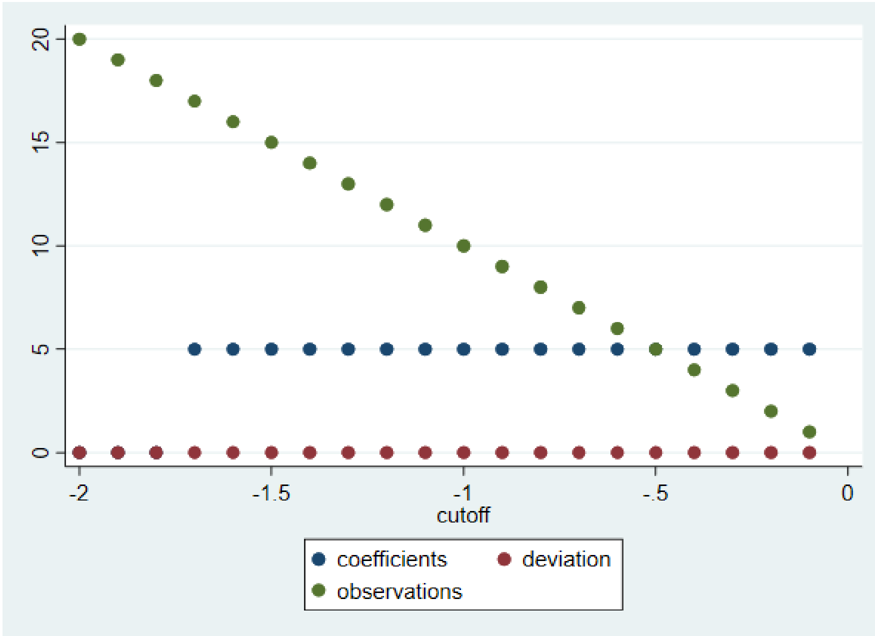
Problem B.2. Graficamos x e y:



Problem B.3. Corremos 20 regresiones discontinuas con intervalos cada vez mas chicos para ver como esto afecta a la estimacion del efecto de w.



Conocemos el comportamiento real de x e y por lo que sabemos que el efecto de w es de 5. Por esto, es claro que existe un problema de medicion pues entre menor es el intervalo, mas cercano se vuelve este a cero sin en ningun momento permanecer en su valor real. Por esto, corremos el mismo modelo pero asumiendo una forma cubica en ambos lados:



En esta la varianza se mantiene en cero y el estimador llega y se queda en 5 hasta que la disminucion del intervalo hace imposible la identificación.

Problem B.4. La estimación anterior hace ver que al hacer una estimacion lineal local, el intervalo de corte es altamente influyente en que tan sesgado sea el estimador. Por esto, es necesario encontrar un grado de amplitud apropiado en cada caso. Imbens y Kalyanaraman(2012) plantean un estimador de la amplitud h en 2012 igual a

$$h_{IK_{n,o}} = [\frac{\hat{V_{IK}}}{2(p+1)\hat{B^2_{IK,p}} + \hat{R^2_{IK,p}}}]^{\frac{1}{(2p+3)}} \times n^{\frac{-1}{(2p+3)}}$$

donde $\hat{V_{IK}}$ es la varianza estimada, $B^2_{IK,p}$ es el sesgo estimado generado por el uso de la ecuacion de grado p y $R^2_{IK,p}$ es un termino de control que evita que cuando el sesgo disminuya el intervalo se dispare. La diferencia entre este estimador y el estimador de Calonico, Cattaneo, and Titiunik es que en este:

$$h_{IK_{n,o}} = [\frac{\hat{V_{CCT}}}{2(p+1)\hat{B^2_{CCT,p}} + \hat{R^2_{CCT,p}}}]^{\frac{1}{(2p+3)}} \times n^{\frac{-1}{(2p+3)}}$$

La varianza y el sesgo son estimados de una manera distinta, mas consistente.

La siguiente tabla muestra distintos valores de corte, tanto a la derecha como izquierda, del punto x=0.

Outcome: y. Running variable: x.

	BW est. (h)		BW bias (b)	
Method	Left of c	Right of c	Left of c	Right of c
mserd	0.173	0.173	0.307	0.307
msetwo	0.163	0.223	0.325	0.309
msesum	0.195	0.195	0.318	0.318
msecomb1	0.173	0.173	0.307	0.307
msecomb2	0.173	0.195	0.318	0.309
cerrd	0.137	0.137	0.307	0.307
certwo	0.129	0.177	0.325	0.309
cersum	0.155	0.155	0.318	0.318
cercomb1	0.137	0.137	0.307	0.307
cercomb2	0.137	0.155	0.318	0.309

De estos valores, mserd es el ancho estimado de Calonico, Cattaneo, and Titiunik, $h_{CCT} = 0{,}173$. Curiosamente, varios de los metodos de estimacion de intervalos optimos tienen su corte por encima de 1.7, el valor en que la forma cubica se indetermina.