Effects of Job Training for Disadvantaged Workers

IDS702 Team Project Group Yellow

Duke

Data & Modeling Process

Data Set

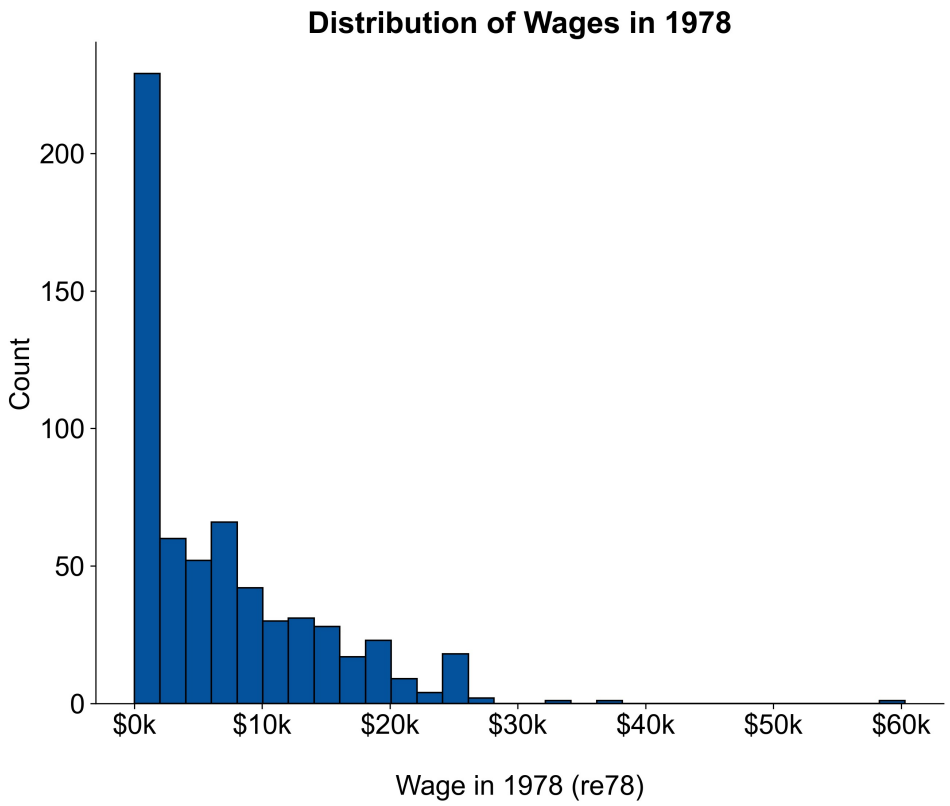
- Collected by Robert LaLonde [1] in the 1970s
- Studies effect of job training for disadvantaged workers on earnings
- Recorded information on demographics, education status and prior wages
- Only male participants
- Paid participation

Modeling Process

Duke

EDA & Experiments

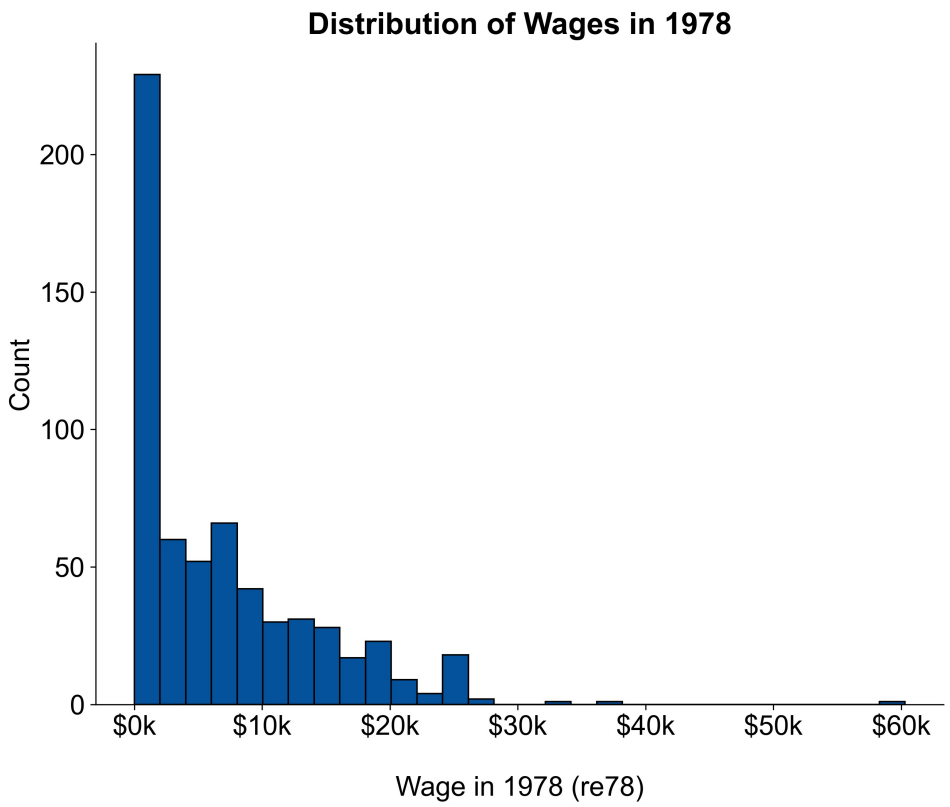
Univariate



Zero inflation → model $re78 - re74$ instead

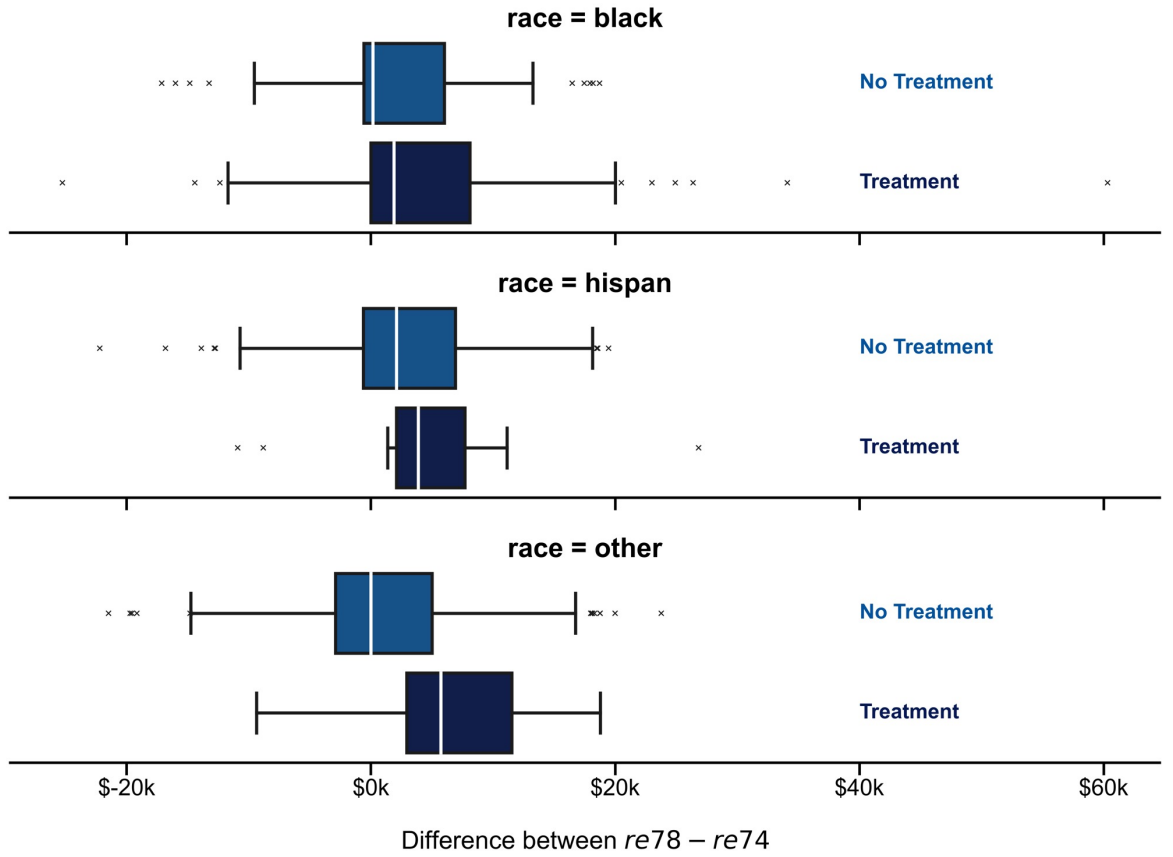
EDA & Experiments

Univariate



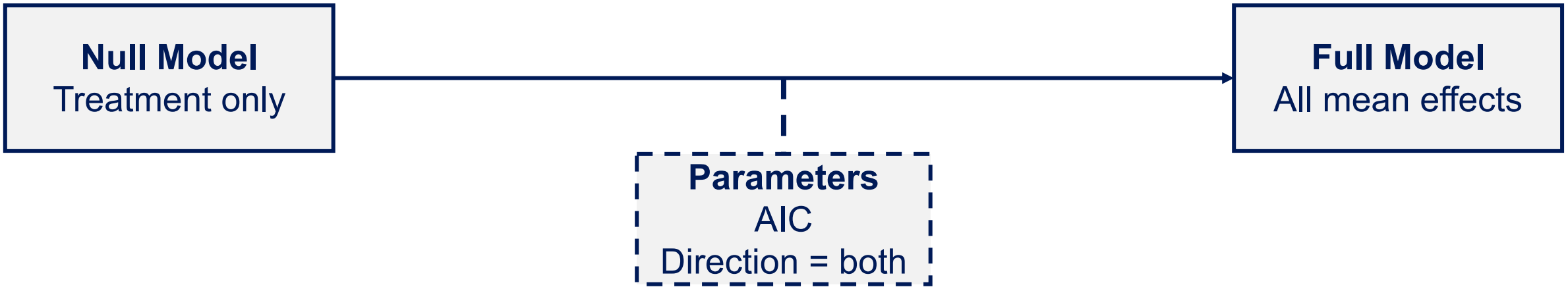
Zero inflation → model $re78 - re74$ instead

Interactions



No interaction between treatment and race

Stepwise Regression



Resulting Model

$$\hat{y} = \widehat{\beta_0} + \widehat{\beta_1} \cdot treat + \widehat{\beta_2} \cdot age + \widehat{\beta_3} \cdot married$$

	Estimate	Std. Error	t value	$Pr(>\|t\|)$	
(Intercept)	4933.27	974.81	5.06	0.00	***
treat1	2085.12	719.90	2.90	0.00	**
age	-93.59	34.43	-2.72	0.01	**
married1	-1842.51	719.56	-2.56	0.01	*

Duke

Testing Interactions

Testing stepwise model against….

- Stepwise + $treat \times black$
- Stepwise + $treat \times hispan$
- Stepwise + $treat \times married$
- Stepwise + $treat \times age$

	Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
1	610	3.715e+10				
2	609	3.663e+10	1	5.178e+8	8.61	0.0035

F Test for $treat \times age$ interaction

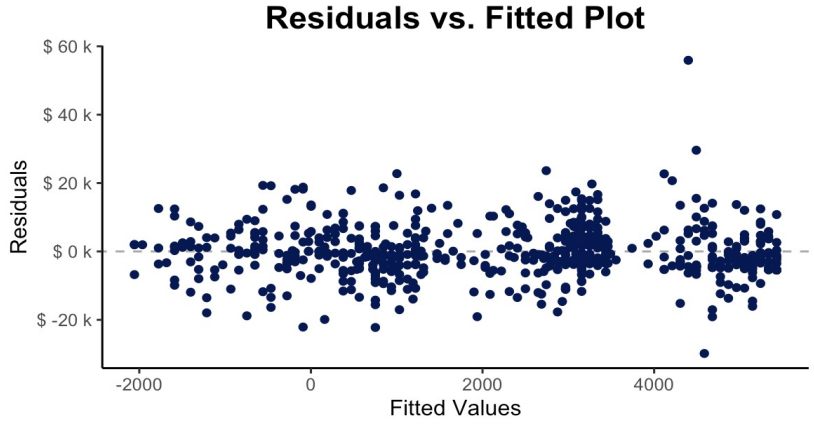
The treatment effect differs by age group, but not by any of the other demographics

Duke

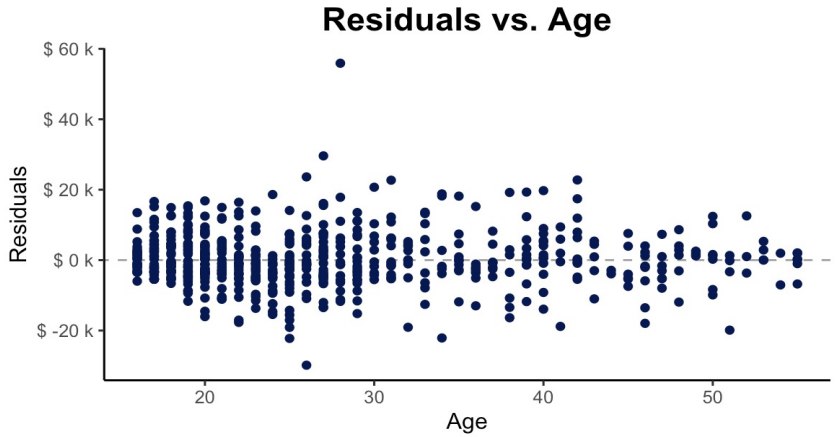
Model Assessment

$$\hat{y} = \widehat{\beta_0} + \widehat{\beta_1} \cdot treat + \widehat{\beta_2} \cdot age + \widehat{\beta_3} \cdot married + \widehat{\beta_4} \cdot treat \times age$$

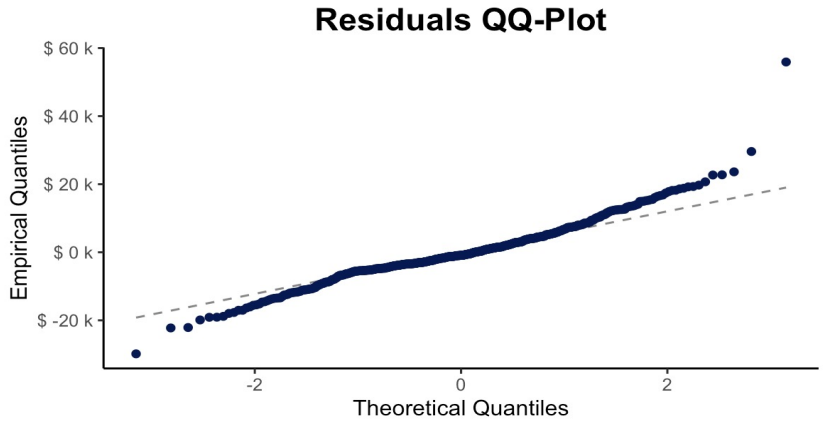
Equal Variance & Independence



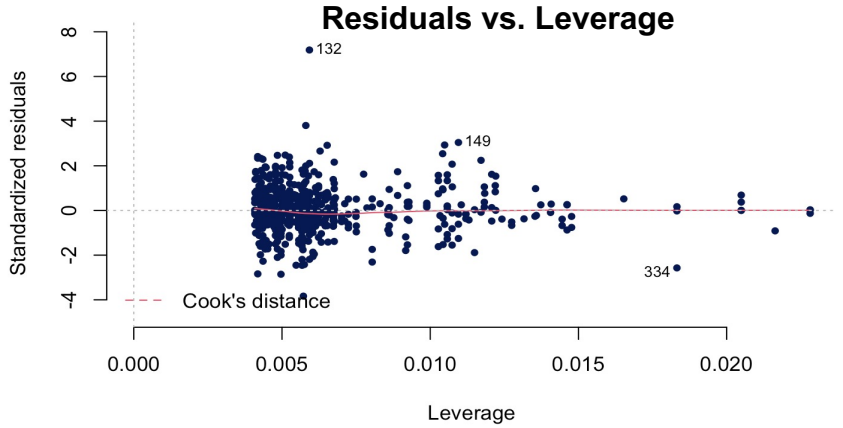
Linearity



Normality



Outliers



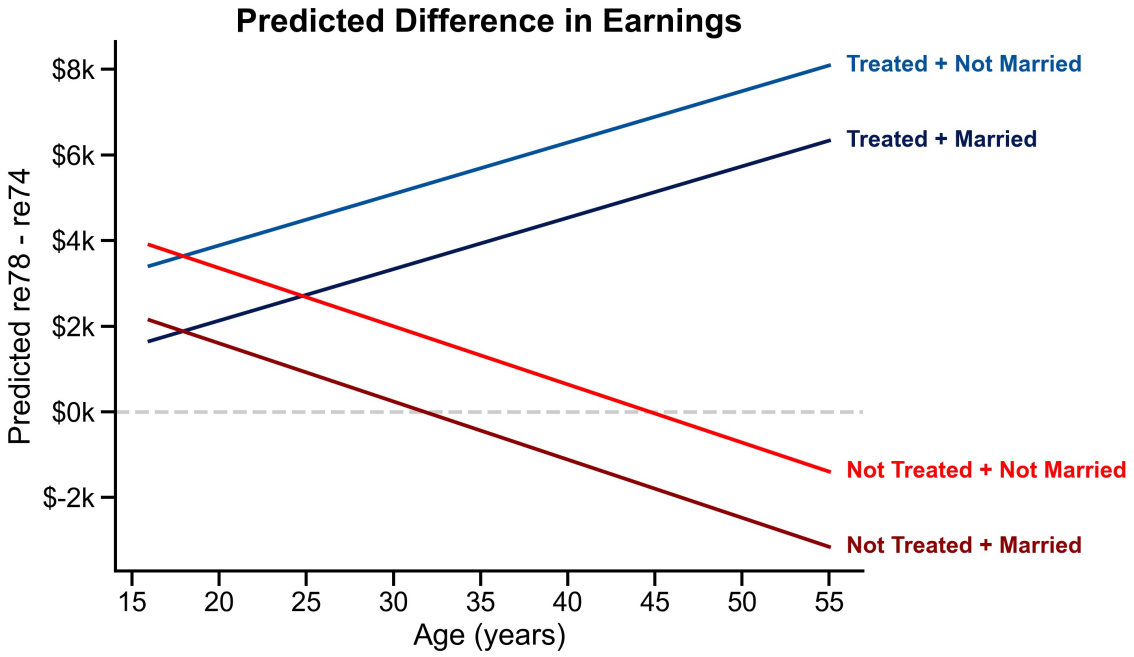
Interpretation

Model

	Estimate	Std. Error	t value	$Pr(>\|t\|)$	
(Intercept)	6072.02	1043.64	5.82	0.00	***
treat1	-4586.31	2383.74	-1.92	0.05	.
age	-135.79	37.11	-3.66	0.00	***
married1	-1756.52	715.72	-2.45	0.01	*
treat1:age	255.88	87.21	2.93	0.00	**

Confidence Intervals (95%)

	Lower Bound	Upper Bound
(Intercept)	4022.44	8121.60
treat1	-9267.66	95.03
age	-208.68	-62.90
married1	-3162.09	-350.94
treat1:age	84.61	427.15



Limitations

- Explained variance: $R^2 \approx 0.074$
- Omitted variable bias
- Strong assignment bias!

	black	hispan	other
No Treatment	20%	14%	66%
Treatment	84%	6%	10%

References

1. Lalonde, R. J. (1986), Evaluating the econometric evaluations of training programs with experimental data, *The American Economic Review*, 76, 604 - 620.

Duke