Measuring Effectiveness of ACME Manufacturing's Career 2030 Training Program

Team 3: Ella Lee, Valerie Chan, Irene Wang, Bradley Ge, Jim Tiao

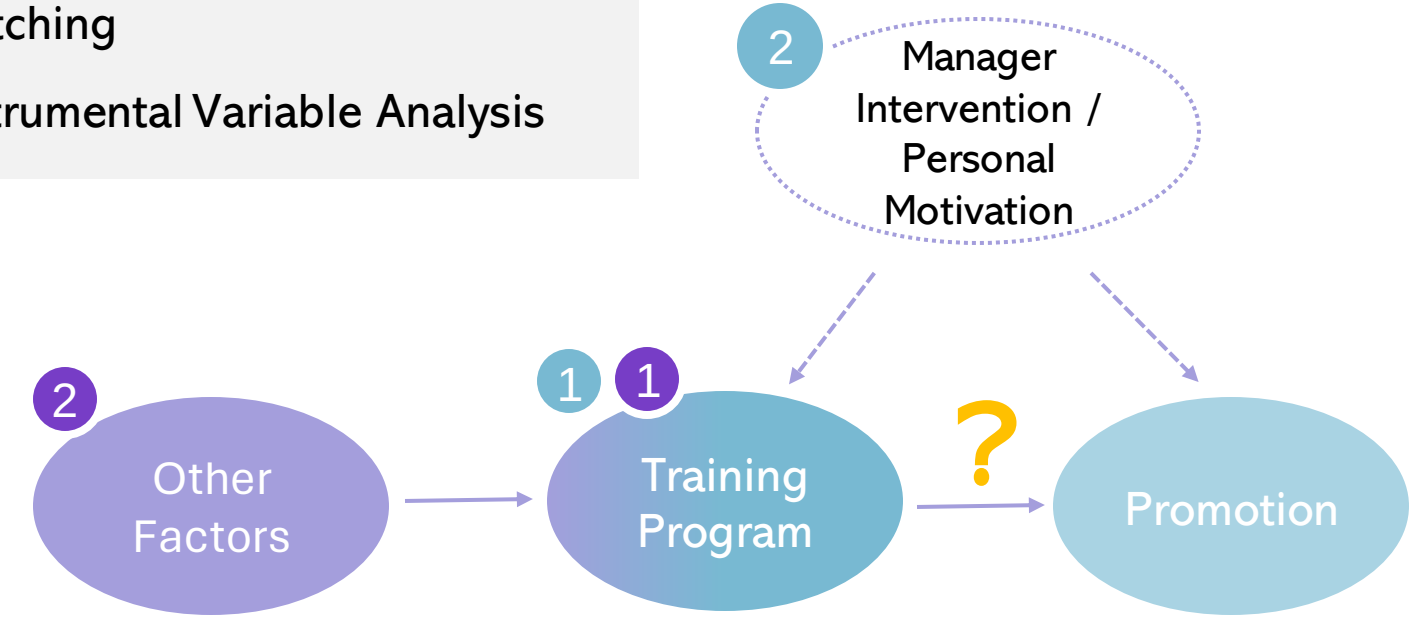
Executive Summary

Purpose	Identify effectiveness of the training program and provide recommendations
Current Issue	The experiment currently employs RCT, but faces a confounder issue
Our Solution	We used IV method to measure the true effectiveness of the training program
Result	RCT originally reflects that training program can lead to a 55.4% higher chance of promotion, which is then corrected to 87.6% following adjustment for hidden biases
Recommendations	• Conduct experiment that reimburses the transportation cost for employees participate in the training • Focus on developing soft skills, particularly networking and communication, to boost promotion prospects • Customize training to engage top performers for better career advancement opportunities

Challenges of the Data Evaluation and Our Solution

Used matching and instrumental variable analysis to address RCT compliance issues

Challenges	Way to solve the problem
1 Imbalanced Class data	1 Matching
2 Confounders (Hidden bias)	2 Instrumental Variable Analysis



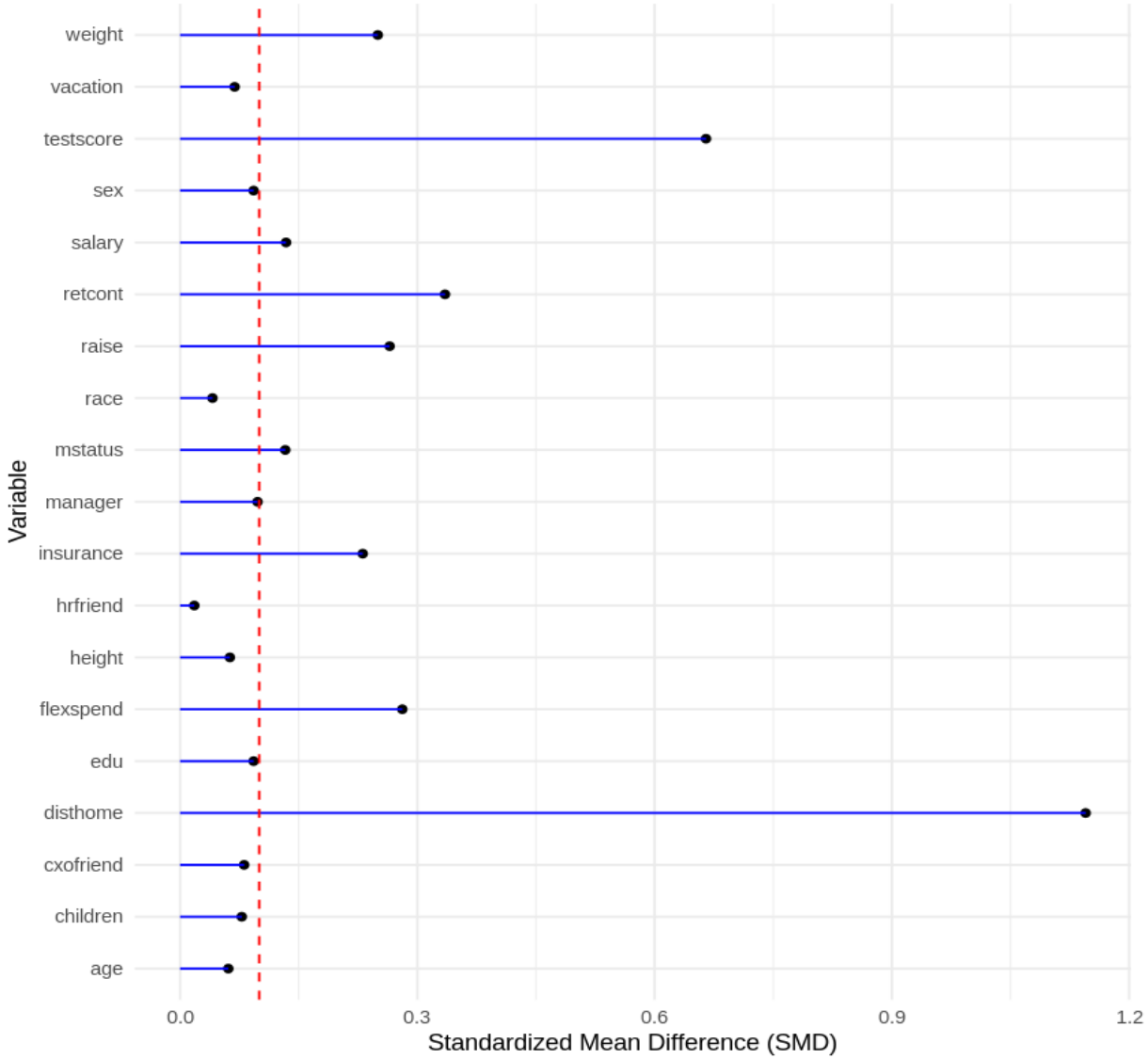
Challenge in Evaluation: Imbalanced Classes

A direct comparison between the treatment and control groups is not reliable due to group imbalance

Imbalanced Class

- Imbalanced classes can lead to misleading causal conclusions
- SMD: Serves as a metric to assess the balance between two groups

	Participated in Training	Not participated	Difference
Retirement Program (%)	5.70%	16%	**-10%**
Raise (%)	29.00%	41.50%	**-12.50%**
Distance from Home (Median)	**16**	**26**	**-10**
Test Score (Median)	58	67.5	**-9.5**



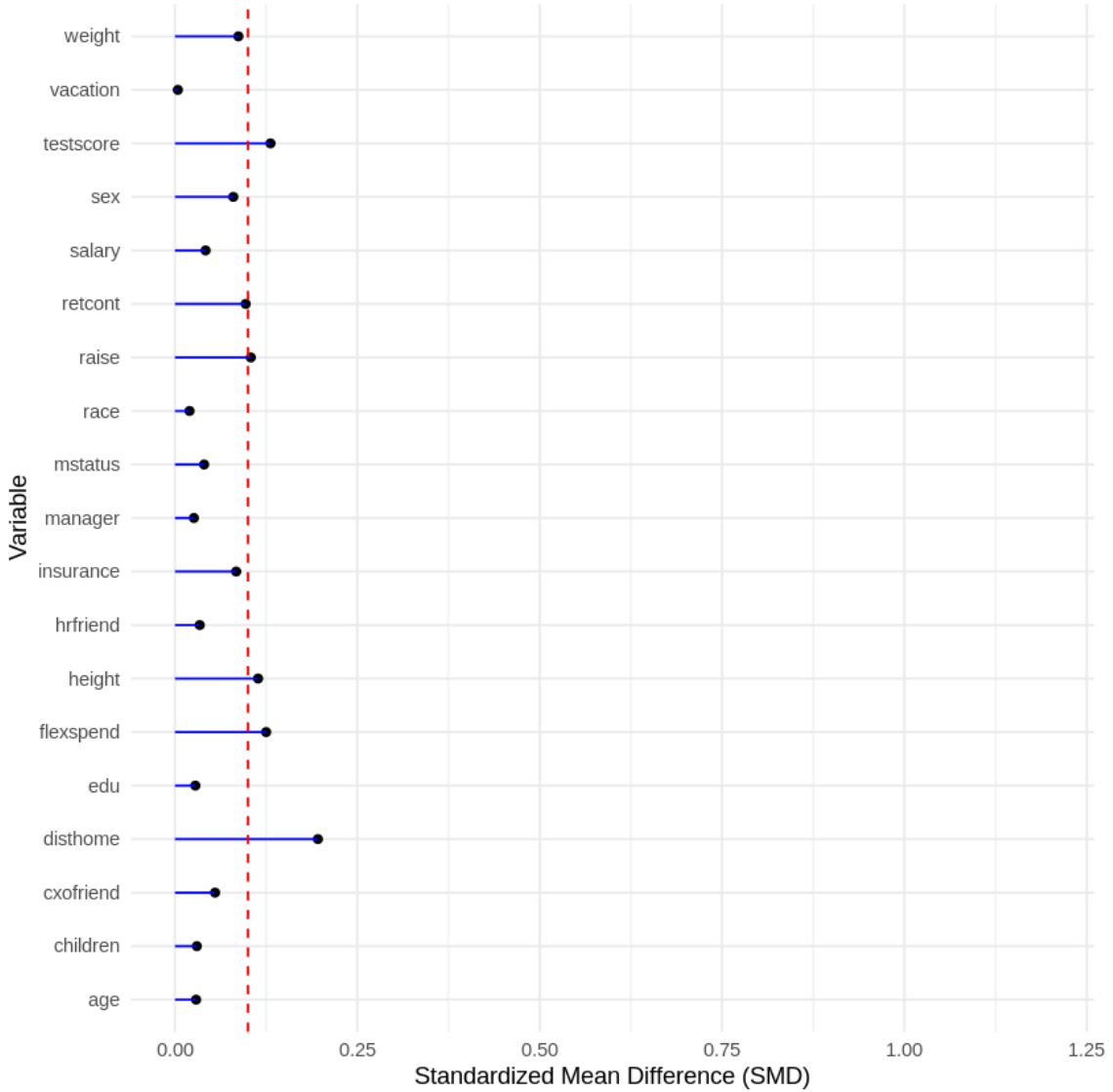
Challenge in Evaluation: Balanced Classes through Matching

Treatment and control groups are comparable after apply propensity score matching techniques

Balance through Matching

- Matching: Pairing individuals who are similar in certain characteristics
- SMD < 0.1: Shows that groups are balanced

	Participated in Training	Not participated	Difference
Retirement Program (%)	7.30%	8.40%	**-1%**
Raise (%)	32.30%	34.30%	**-2.00%**
Distance from Home (Median)	**19.29**	**20.96**	**-1.67**
Test Score (Median)	58.01	59.91	**-1.9**



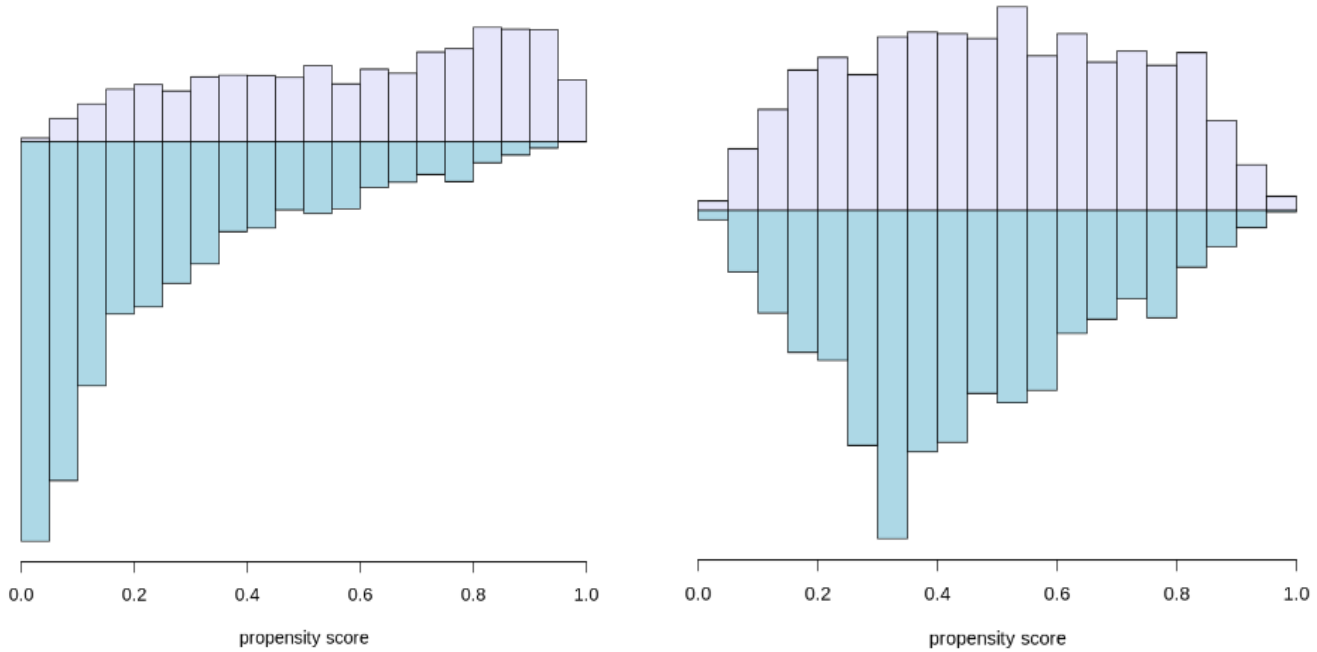
Appendix B - PSM Matching was selected based on 3 criteria

Challenge in Evaluation: Matching

Treatment and control groups are comparable after propensity score matching

Propensity score before and after matching

Propensity Score:
Probability to be assigned to the treatment group



- Training and control group show similar distribution **after matching.**
- This indicates that for each propensity score in the treatment group, a similar control can be found.

Challenge in Evaluation: Hidden Bias

Instrumental Variable Analysis was conducted to address potential hidden biases

Endogeneity Problem:

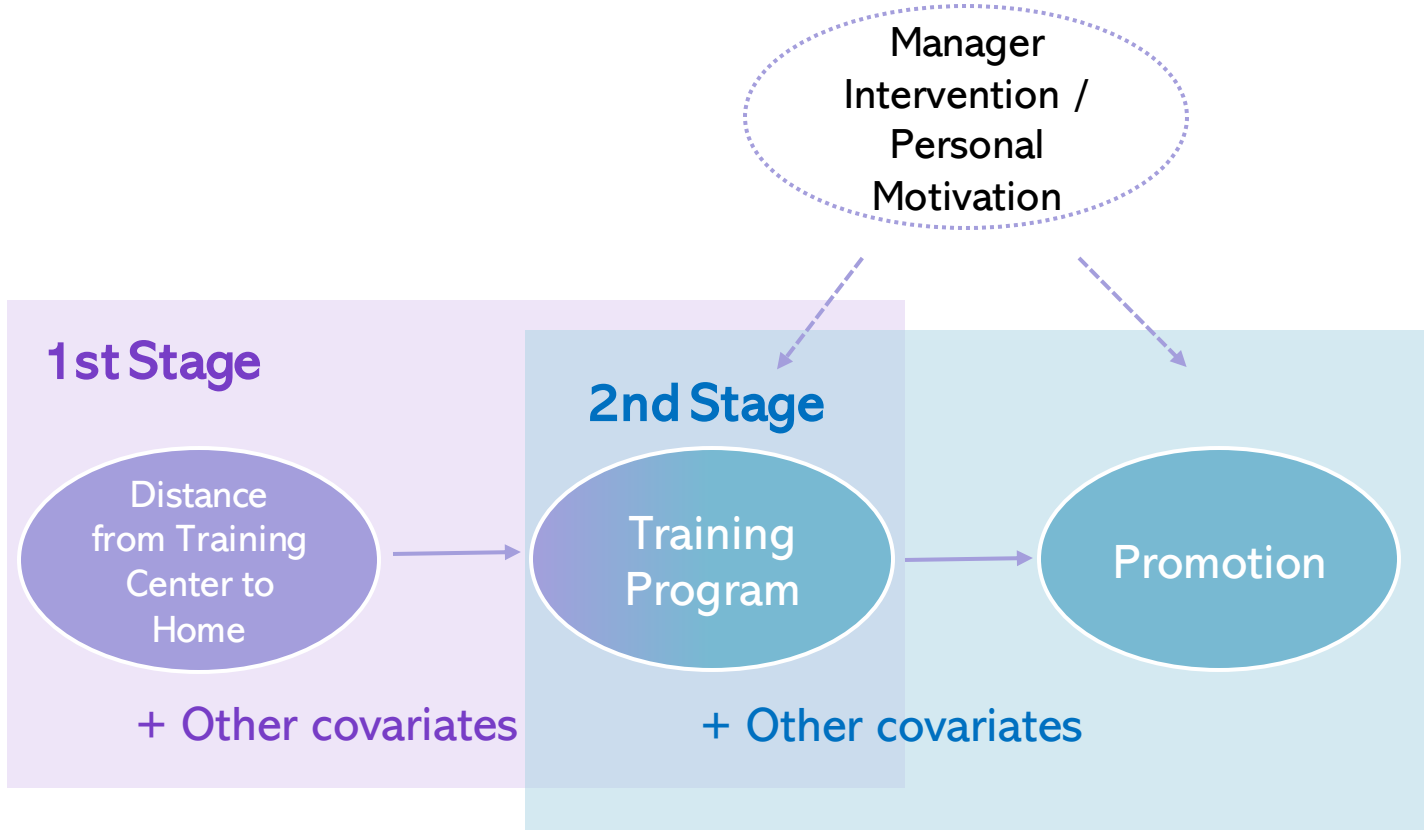
Some unobserved variables, including manager intervention and personal motivation, could affect if an employee attends the training program and gets promotion.

Solution:

- Identify instrumental variable (Distance from training facility to home)
- Perform Instrumental Variable Analysis to isolate the influence of attending the training program from promotion

Approach:

Perform Two-stage Regression Analysis



Training Impact Evaluation: 1st Stage Results

Four variables are negatively correlated to the odds of attending the training program

Stage 1: Use Training Program as Target Variable

Variables	Estimates	Change in Odds
Distance from training facility to home	-0.033	-3.3%
Got a raise last year	-0.154	-14.3%
Enrolled in the retirement program	-0.363	-30.5%
Test score	-0.013	-1.3%

Example interpretation:
An additional mile in the distance from training facility to home is associated with a 3.3% decrease in the odds of an employee attending the training program.

The employees with the following properties are **less likely** to attend training program:

1. Live far away from the training facility
2. Received a raise in salary last year
3. Enrolled in the retirement program
4. Achieved a higher test score

Training Impact Evaluation: Results

Attending training program increases the chance of getting promoted by 87.6%

Stage 2: Use Promotion as Target Variable

Variables	Estimates	Change in Odds
Training Program	1.945	600.0%
Married	-0.606	-45.4%
Single	-1.035	-64.5%
Age	0.039	4.0%
Number of vacation days taken	-1.173	-69.1%
Has HR connection	0.394	48.3%
Has C-level connection	0.800	122.5%
Test Score	0.118	12.5%

→

- Attending training program is **positively & causally** related to getting promotion
- Increases the chance of promotion by **87.6%**

→

Employees with **higher chance** to get promoted may have the following characteristic:

1. Took fewer vacation days last year
2. Have HR & C-executive connections
3. Achieved a higher test score

Recommendations

Recommendations to scale the ACME's Career Training Program and IV analysis

Training Program:

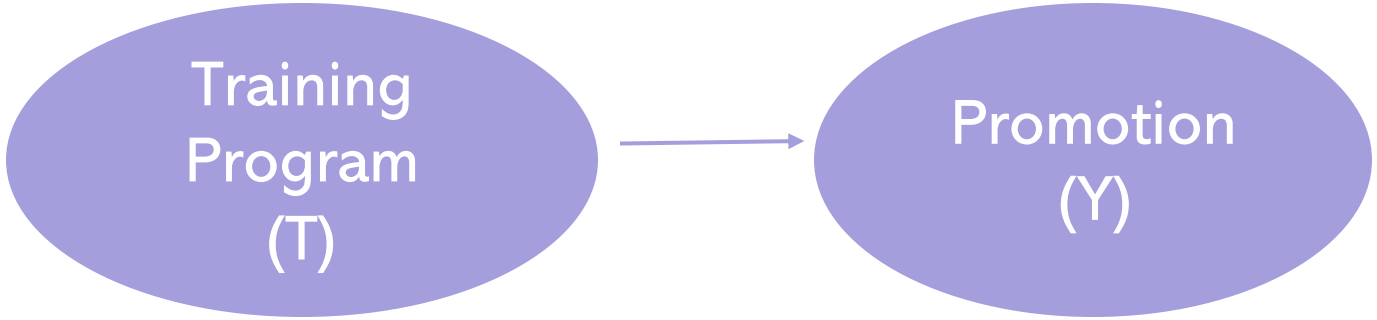
- Conduct experiments on reimbursing transportation costs for employees who participate in the training program
- Focus on developing soft skills, particularly networking and communication, to boost promotion prospects
- Customize training to engage top performers for better career advancement opportunities

Instrumental Variable Analysis:

- Improve the randomization in selecting treatment and control group to avoid imbalance in covariates that could cause confounding issues

Thank You

Appendix A – DAG representation



1. Determine the causal impact of the Career 2030 training on employee promotions.
2. Identify strategies to scale the training program, focusing on content enhancement and structural improvement

Appendix B - Matching Method Decision Criteria

	1:1 Matching	PSM Matching	IPTW
Positivity Assumption Violation	No Violation	No Violation	High Weights
Num of Covariates SMDs over 0.1	3 covariates (2 over 0.2)	2 covariates close to 0.1	
Number of employees/Pairs	636 pairs	1646 pairs	

Appendix C - Fine-Tuning Propensity Score Matching for Covariate Balance via Caliper Trials

- Propensity Score Matching (PSM) utilized to find optimal pairs with lower Standardized Mean Differences (SMDs)
- Caliper set to 0.80 for optimal results, ensuring no violation of positivity assumption

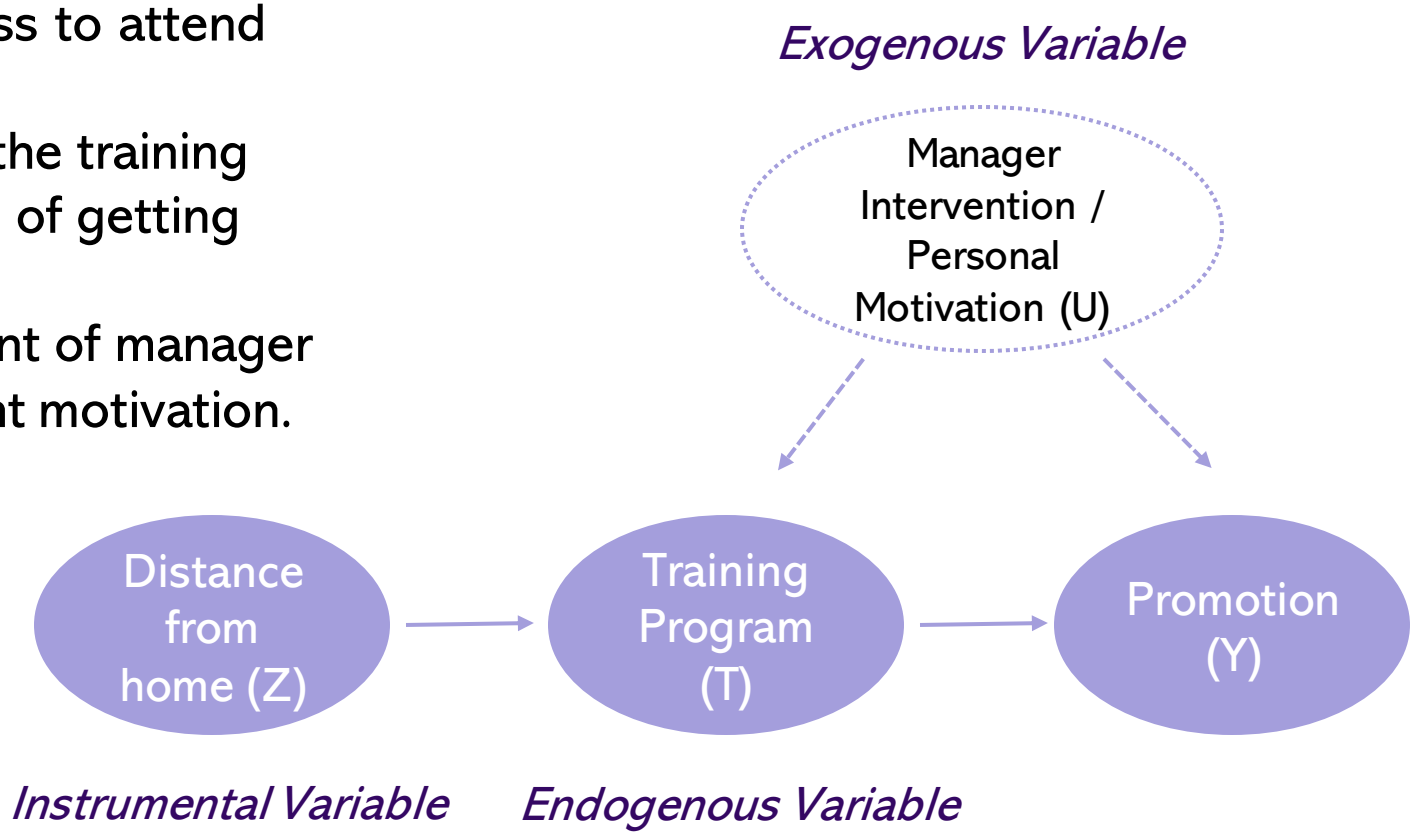
Caliper Trials and Selection for 1:1 Matching

Caliper	Num of Pairs	Positivity Assumption Violation	Covariates SMDs over 0.1 Count
0.6	1566	Weak	1
0.8	1646	No	2
1.0	1760	Weak	2 (1 over 0.2)
1.2	1856	Weak	2 (1 over 0.2)
1.5	1991	No	4 (2 over 0.2)

Appendix D - Instrumental Variable Analysis

IV Assumptions:

- *Relevance:* Traveling distance may be adversely correlated to employees' willingness to attend training.
- *Exclusion:* Distance from home to the training facility is not related to the chance of getting promoted.
- *Exogeneity:* Distance is independent of manager intervention or career advancement motivation.



Appendix E - IV Analysis Stage 1 Results

Stage 1:

- Correlation coefficient of -0.16 confirms that a longer distance from home negatively impacts training participation.

Coefficients:

	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	1.125252	0.645326	1.744	0.08121	.
disthome	-0.033405	0.005063	-6.598	4.16e-11	***
raiseYes	-0.154306	0.078104	-1.976	0.04819	*
salary$20-$40k	-0.014253	0.210071	-0.068	0.94591	
salary$40-$80k	-0.018480	0.163848	-0.113	0.91020	
salaryUnder $20k	-0.043093	0.208621	-0.207	0.83635	
children	0.005583	0.035646	0.157	0.87554	
mstatusmarried	0.126274	0.116330	1.085	0.27771	
mstatussingle	0.021599	0.093271	0.232	0.81687	
age	0.001659	0.003717	0.446	0.65540	
sexMale	0.057902	0.072328	0.801	0.42339	
edu	0.080361	0.144215	0.557	0.57737	
vacation	0.006321	0.148395	0.043	0.96602	
hrfriendYes	0.002580	0.070970	0.036	0.97100	
cxofriendYes	0.076103	0.072471	1.050	0.29367	
managerYes	-0.018965	0.158716	-0.119	0.90489	
insuranceCovered & Medicaid	-0.030495	0.126188	-0.242	0.80904	
insuranceCovered & Medicare	-0.463231	0.512862	-0.903	0.36641	
insuranceMedicaid	-0.096395	0.105090	-0.917	0.35900	
insuranceMedicare	-0.559180	0.363350	-1.539	0.12381	
insuranceMedicare & Medicaid	-0.013304	1.425972	-0.009	0.99256	
insuranceOther	0.006910	0.172018	0.040	0.96796	
flexspendYes	0.095575	0.077925	1.226	0.22001	
retcontYes	-0.363465	0.134002	-2.712	0.00668	**
raceother	0.125958	0.170011	0.741	0.45877	
racewhite	0.055165	0.100658	0.548	0.58366	
testscore	-0.013631	0.002711	-5.028	4.95e-07	***

Appendix F - IV Analysis Stage 2 Results

Stage 2:

- Seven covariates strongly associated with promotion:
  - Older employees who took fewer vacation days, got higher test scores, and have HR/C-level executive connections are more strongly associated with getting promoted

Coefficients:

	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	-6.680555	0.979352	-6.821	9.02e-12	***
managerYes	-0.064522	0.213427	-0.302	0.762414	
trainingest	1.944700	0.811102	2.398	0.016503	*
raiseYes	-0.174337	0.105295	-1.656	0.097784	.
salary$20-$40k	-0.251214	0.278041	-0.904	0.366254	
salary$40-$80k	-0.068262	0.215214	-0.317	0.751106	
salaryUnder $20k	0.161177	0.276506	0.583	0.559955	
children	-0.040388	0.047357	-0.853	0.393751	
mstatusmarried	-0.605551	0.161585	-3.748	0.000179	***
mstatussingle	-1.035295	0.131393	-7.879	3.29e-15	***
age	0.038788	0.004903	7.911	2.56e-15	***
sexMale	0.175744	0.096723	1.817	0.069219	.
edu	0.315525	0.193500	1.631	0.102970	
vacation	-1.174464	0.197169	-5.957	2.57e-09	***
hrfriendYes	0.393998	0.093799	4.200	2.66e-05	***
cxofriendYes	0.800568	0.094496	8.472	< 2e-16	***
insuranceCovered & Medicaid	-0.063139	0.170453	-0.370	0.711070	
insuranceCovered & Medicare	0.122994	0.679901	0.181	0.856446	
insuranceMedicaid	-0.038417	0.138695	-0.277	0.781790	
insuranceMedicare	0.224185	0.493354	0.454	0.649534	
insuranceMedicare & Medicaid	10.015467	313.065948	0.032	0.974479	
insuranceOther	-0.105036	0.221407	-0.474	0.635212	
flexspendYes	-0.188710	0.104085	-1.813	0.069827	.
retcontYes	0.288687	0.178464	1.618	0.105744	
raceother	-0.104336	0.223029	-0.468	0.639919	
racewhite	-0.063759	0.133500	-0.478	0.632934	
testscore	0.117680	0.005012	23.481	< 2e-16	***