Estimating Long-Term Treatment Effects of Combined Diet and Exercise Using Regression Adjustment and Inverse Probability Weighting

Alice Lai

Section 1: Treatment effects using observational data

i) Describe the data using summary statistics as well as appropriate graphs. Include comparisons of the variables across the two treatment groups. [15 marks]

The dataset includes 4,732 participants aged 18–55 (mean = 38.57), with a slightly left-skewed distribution (Figure 1). Gender is relatively balanced (55.85% female), and most participants are overweight (mean baseline BMI = 27.55), with a near-normal distribution (Figure 2). Only 30.33% reported chronic stress, and 87.89% lived near fast-food area. However, just 25.8% received the Diet + Exercise intervention, indicating group imbalance. Percentage distribution of participants across key baseline variables is summarized in Figure 3.

***Figure 1*: Age distribution**

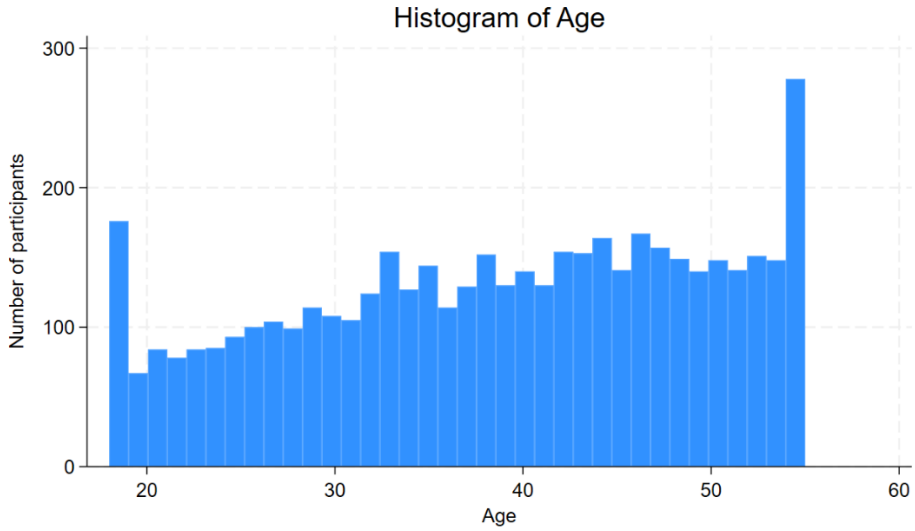


***Figure 2*: Baseline BMI distribution**

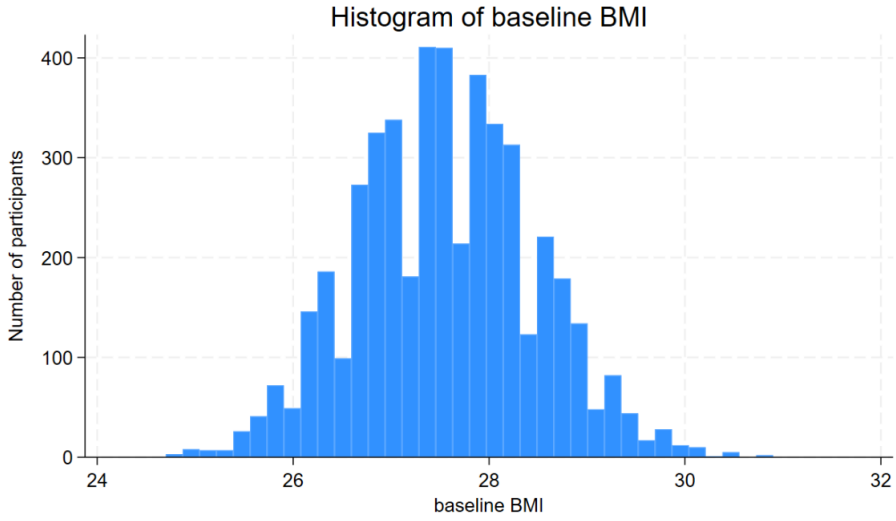
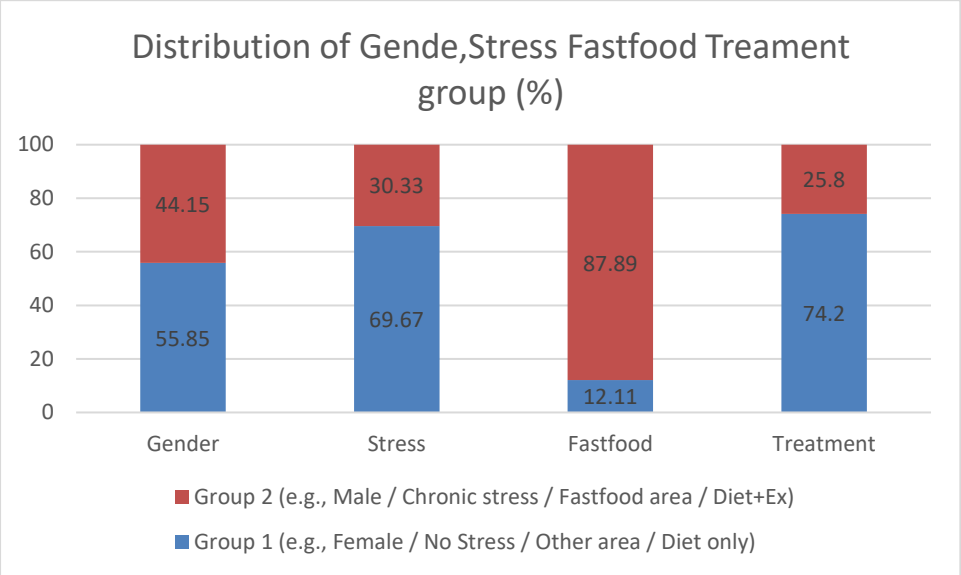


***Figure 3*: Distribution of Gender, Chronic stress, Fast-food area, Treatment group (%)**



Chronic stress was more common among females (37.0% vs. 21.9%), while fast-food exposure varied little by gender or stress. Age, stress, and fast-food access showed no strong link to baseline BMI; however, males had lower baseline BMI than females (Figure 4). The BMI change was approximately normal (Figure 5), supporting linear regression. The Diet + Exercise group saw significantly greater BMI reduction (1.81% vs. 0.64%, $p < 0.001$) (Figure 6).

***Figure 4*: Baseline BMI by gender group**

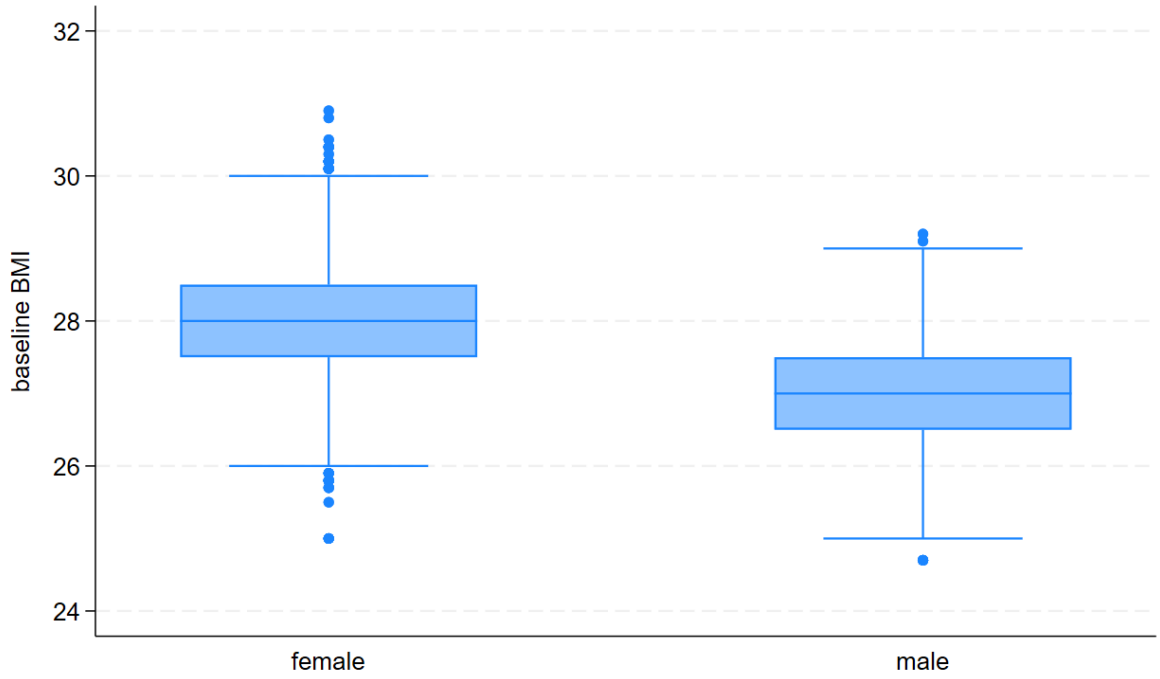


***Figure 5*: Histogram of average BMI percentage change after 10 years**

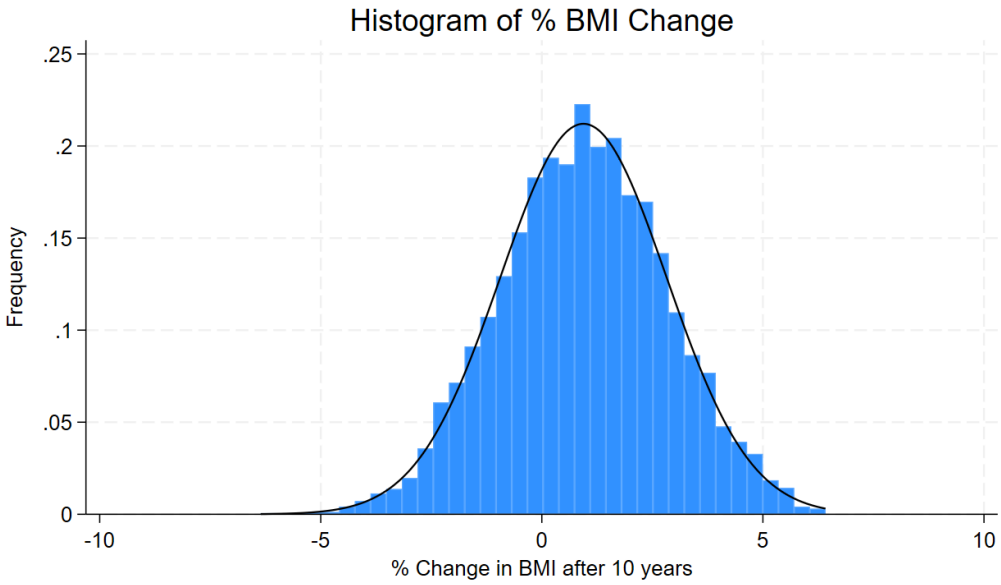
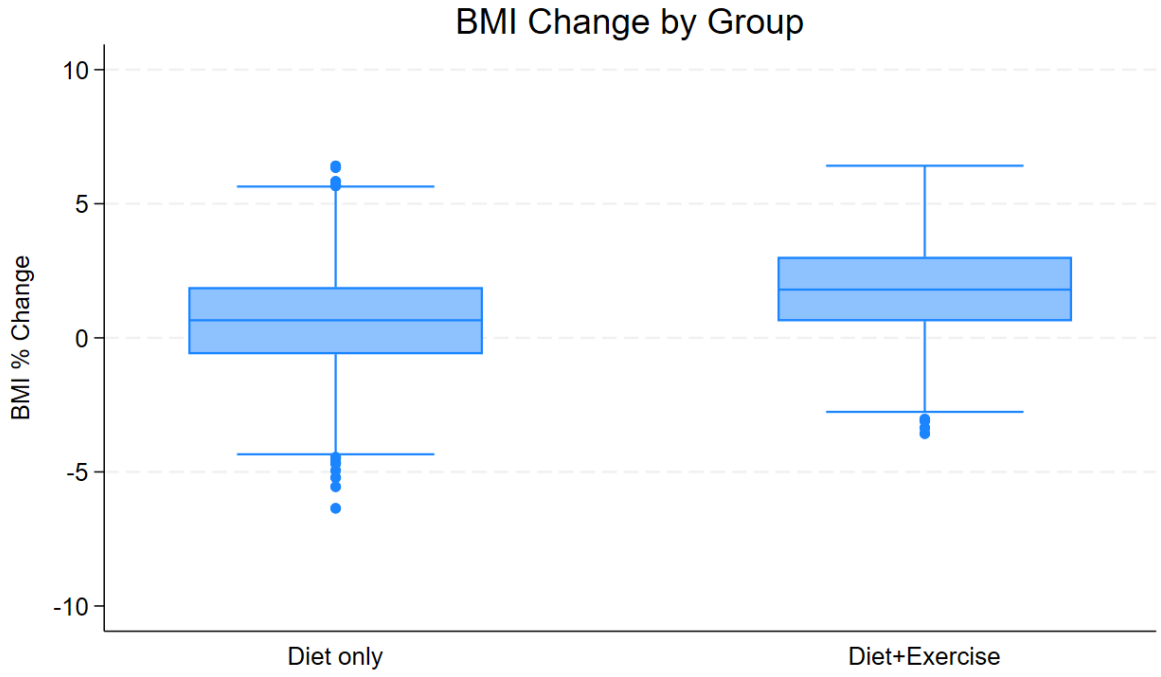


***Figure 6*: BMI percentage change by Treatment group**



BMI change was moderately negatively correlated (Figure 7) with age (r = -0.497), slightly stronger with a quadratic age term (r = -0.501). A weak but significant negative correlation (Figure 8) was also found between baseline BMI and BMI change (r = -0.203, $p < 0.001$). Males and participants without chronic stress showed greater improvements; no significant link was found between fast-food exposure and BMI change.

***Figure 7*: Association between Age and Percentage Change in BMI**

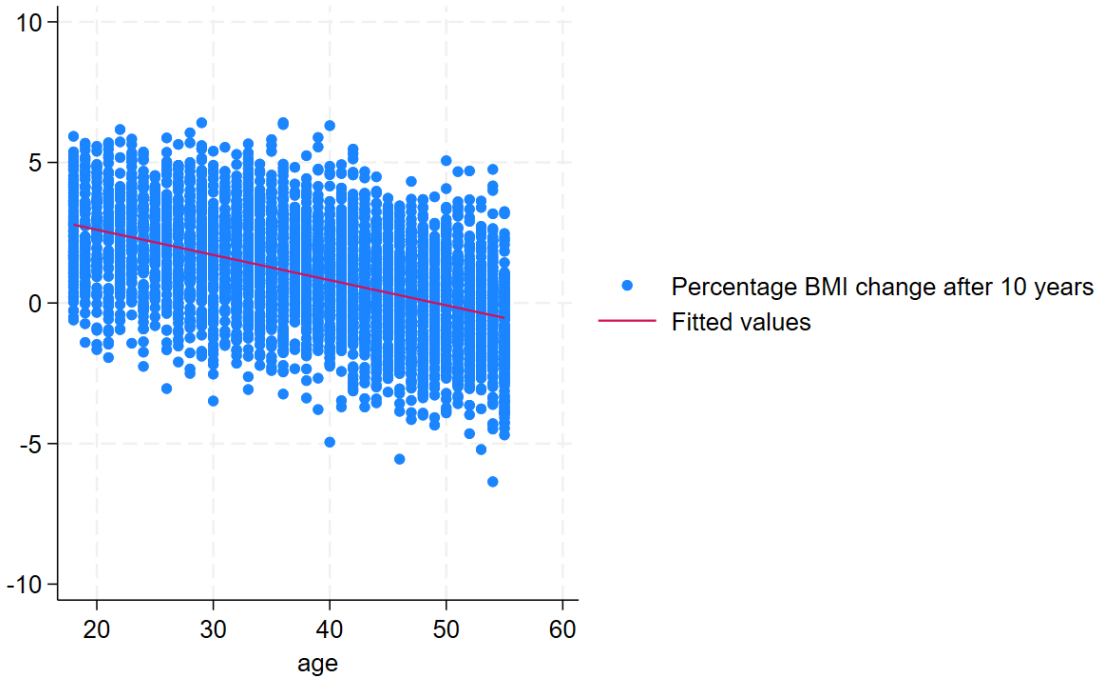
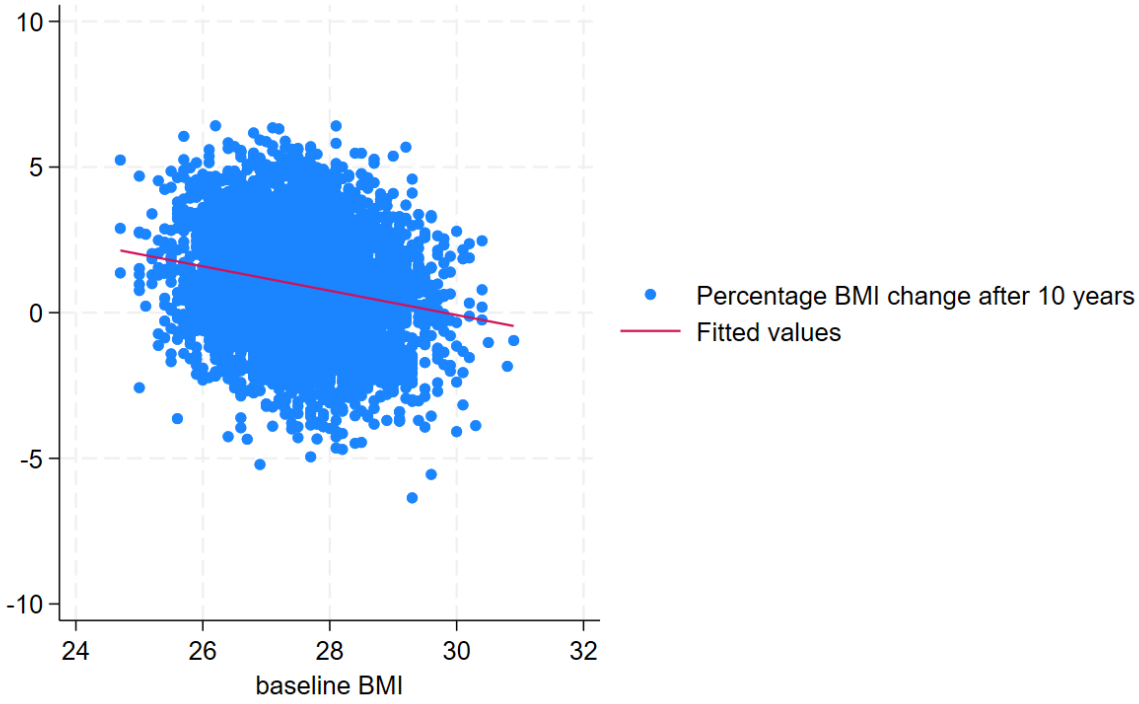


***Figure 8*: Association between baseline BMI and Percentage Change in BMI**



Substantial group imbalances were observed (Table 1). The Diet + Exercise group was significantly younger (33.4 vs. 40.4 years), had lower baseline BMI (27.27 vs. 27.65), a higher proportion of males (53.5% vs. 40.9%), and fewer reports of chronic stress (21.1% vs. 33.5%) than the Diet-only group (all $p < 0.001$). While fast-food proximity differed slightly (87.1% vs. 88.2%), the difference was not significant.

***Table 1*: Baseline Characteristics by Treatment Group**

Variable	Diet only (n=3,511)	Diet + Exercise (n=1,221)	p-value
Age	40.4	33.4	<0.001
Baseline BMI	27.65	27.27	<0.001
% Male	40.9%	53.5%	<0.001
% Chronic stress	33.5%	21.1%	<0.001
% Fast-food area	88.2%	87.1%	0.23

ii) Use regression adjustment to calculate the average treatment effect (ATE) justifying any choices made. Interpret and comment on its statistical significance. Explain why this estimator might be problematic in this case. [25 marks]

Linear regression was applied due to the approximate normality of the BMI percentage change outcome. Different covariate combinations were tested to address confounding.

A naïve model (Model 1) yielded an unadjusted ATE of 1.164 percentage points ($p < 0.001$), reflecting the crude group difference. A fully adjusted model (Model 2), controlling for age, age², baseline BMI, gender, stress, and fast-food exposure, produced a reduced ATE of 0.423 ($p < 0.001$), indicating much of the initial effect was due to covariate imbalance. Model 3 excluded stress and fast-food exposure, which were insignificant and minimally explanatory. Likelihood-ratio testing showed no significant difference in fit between Model 2 and Model 3 ($p = 0.0864$). Diagnostics (RESET, linktest, hettest, residuals) showed no misspecification or heteroskedasticity; AIC/BIC supported improved model fit.

Potential effect modification by age and baseline BMI was explored. Subgroup analyses (Tables 2 & 3) and scatterplots (Figures 9 & 10) showed that younger individuals and those with lower baseline BMI had greater BMI changes. Still, the treatment effect remained consistent across covariates. In age-based plots, both groups followed parallel declining slopes, with similar patterns for baseline BMI, suggesting no effect modification. Models 4 and 5 formally tested treat × age and treat × BMI interactions. Neither was statistically significant, and model fit did not improve. Thus, age and baseline BMI were retained as confounders rather than effect modifiers.

Table 2*: *Average % BMI change by age subgroup

Age	Average % BMI change
<30	+2.17
30-39	+1.42
40-49	+0.42
50-55	-0.41

Table 3*: *Average % BMI change by baseline BMI subgroup

Baseline BMI	Average % BMI change
Normal(18.5-24.9)	+3.17
Overweight(25-29.9)	+0.95
Obese(>=30)	-0.40

Figure 9*: *Unadjusted BMI Change by Age and Treatment Group with Group-Specific Averages

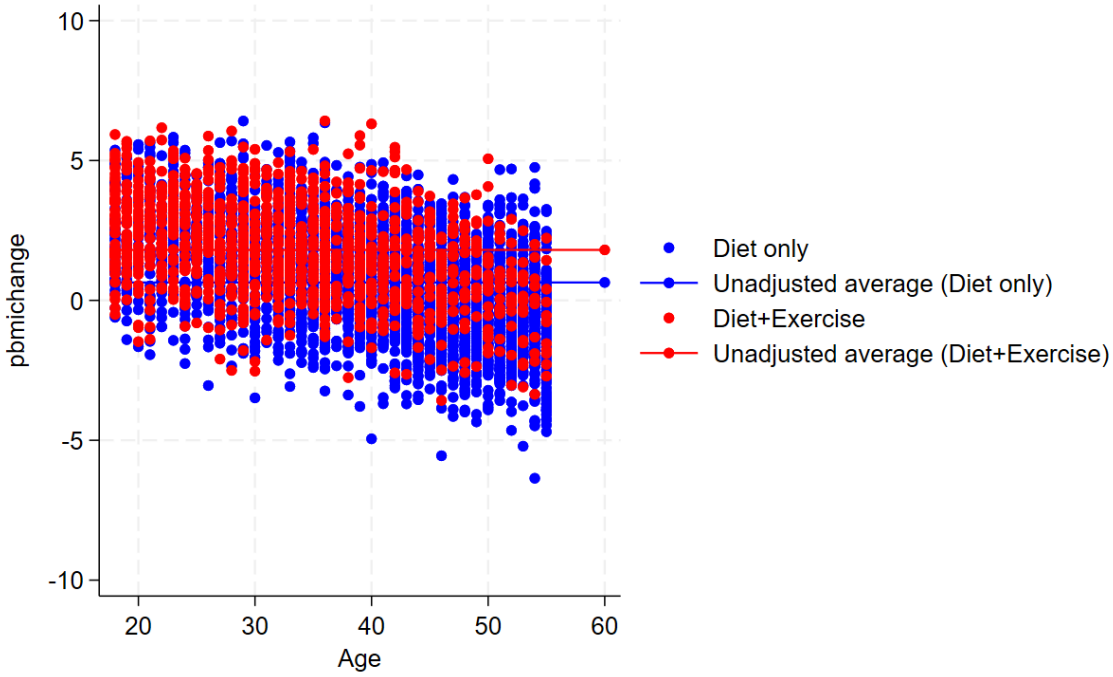
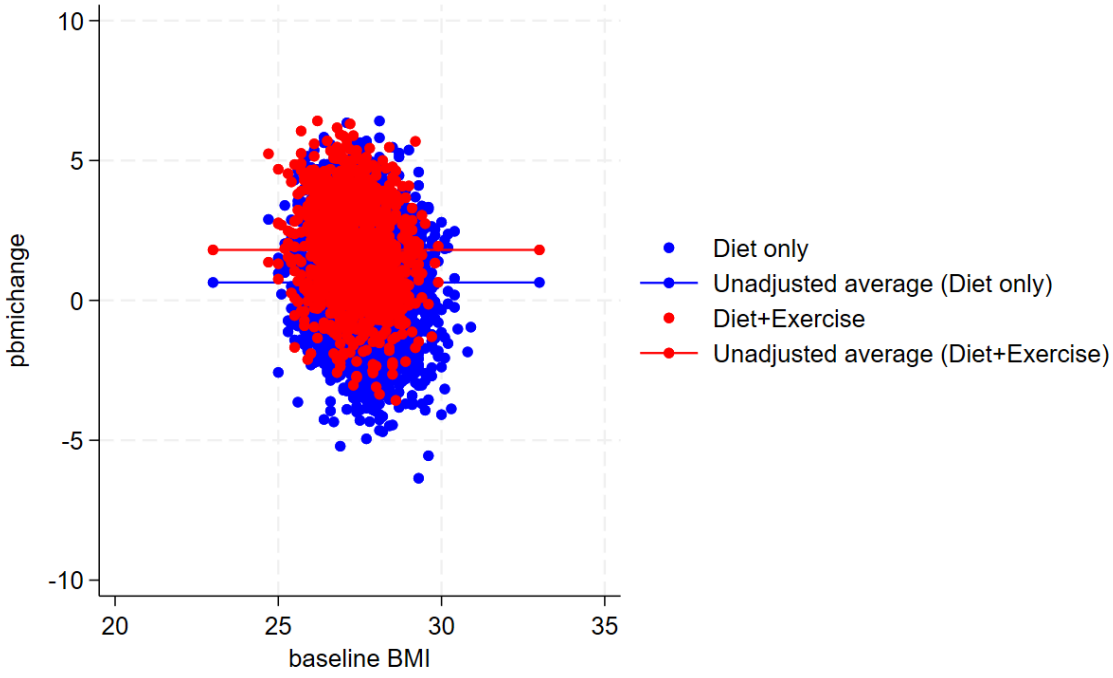


Figure 10*: *Unadjusted BMI Change by baseline BMI and Treatment Group with Group-Specific Averages



Model 3 emerged as the best-performing model (Table 4), balancing explanatory power, simplicity, and statistical validity. The coefficients of Model 3 are as follows:

$$\%\ \boldsymbol{BMI\ change} = 0.434 * treat + 0.007 * age - 0.001 * age^2 - 0.286 * baselineBMI + 0.286 * male$$

***Table 4*: Linear regression Model with different covariant**

Model	Variables Included	Interaction(s)	Significant Variables (p < 0.05)	Treat Coef.	Pseudo R^2	AIC	BIC
1(Naïve)	treat	None	treat	1.164***	0.073	19051.71	19064.63
2(Full)	treat age age2 bmi0 male stress fastfood	None	treat age2 bmi0 male	0.423***	0.305	17703.98	17755.68
3(Parsimonious)	treat age age2 bmi0 male	None	treat age2 bmi0 male	0.434***	0.304	17704.88	17743.65
4(Age interaction)	Treat x age age age2 bmi0 male	Treat x age(ns)	age, bmi0, stress	0.227(ns)	0.304	17705.73	17750.97
5 (baseline BMI interaction)	Treat x bmi0 age age2 bmi0 male	Treat x BMI (ns)	age, bmi0, stress	1.84(ns)	0.304	17706.12	17751.36

The final treatment effect from Model 3 was 0.434 (95% CI: 0.325–0.543, $p < 0.001$), indicating a significant and clinically meaningful effect of the Diet + Exercise intervention. This estimate closely aligned with that from the regression-adjustment command (ATE = 0.441, 95% CI: 0.325–0.558), reinforcing the result.

Despite this, regression adjustment has limitations. It assumes correct model specification, including linearity and inclusion of all relevant covariates and interactions. Misspecification may introduce bias. Unlike IPW, RA does not ensure covariate balance. Given treatment group imbalances and differing covariate distributions, there is a risk of extrapolation—especially when predicting outcomes for older individuals in the underrepresented treatment group. While diagnostics supported the model, results should be interpreted cautiously due to model dependence and possible residual confounding.

iii) **Use inverse probability weighting to estimate the ATE justifying any choices made. Check balance and overlap and comment on your results. [25 marks]**

Given substantial baseline imbalances in age, gender, stress, and BMI across treatment groups, inverse probability weighting (IPW) was applied. A probit model estimated propensity scores for receiving the combined intervention and calculating the average treatment effect (ATE).

Model 4 (age, BMI, stress) was selected after comparing six alternatives (Table 5). Although more complex models included gender, fast-food exposure, quadratic age terms, or gender interactions, none significantly improved model fit. Model 4 had the lowest AIC (4747.77) and performed comparably to the full model in likelihood-ratio testing ($\chi^2 = 0.20$, $p = 0.9039$), with strong goodness-of-fit (Pearson $\chi^2 = 2066.95$, $p = 0.7031$). Covariates were selected based on prior associations with treatment assignment.

Table 5: **Probit Model with different covariant**

Model	Variables Included	Interaction(s)	Significant Variables (p < 0.05)	Pseudo R^2	AIC	BIC
1	age, bmi0, male, stress, fastfood	None	age, bmi0, stress	0.1229	4751.56	4790.33
2	age, age², bmi0, male, stress, fastfood	age × age	bmi0, stress	0.1231	4752.66	4797.90
3	age group dummies, bmi0, male, stress, fastfood	categorical age	agegrp, bmi0, stress	0.1155	4795.45	4847.14
4	age, bmi0, stress	None (parsimonious)	age, bmi0, stress	0.1229	4747.76	4773.61
5	age, bmi0, male, stress, age × male	Treat × age (ns)	age, bmi0, stress	0.1231	4750.563	4789.33
6	age, bmi0, male, stress, bmi0 × male	Treat ×baseline BMI (ns)	age, bmi0, stress	0.1232	4750.13	4788.90

The ATE estimated via IPW was 0.432 (95% CI: 0.297–0.567, $p < 0.001$), indicating that the Diet + Exercise group has average 0.43 percentage point greater BMI reduction over 10 years versus the Diet-only group. Weight diagnostics showed an average of 2.01 and a maximum of 45.49, suggesting some participants with extreme scores contributed disproportionately. A sensitivity analysis using trimmed weights (capped at 10) yielded a slightly larger ATE of 0.551, confirming that the main result was not driven by outliers and reinforcing estimate credibility.

Covariate balance improved substantially post-weighting. Before adjustment, standardized mean differences (SMDs) for age, BMI, and stress were -0.70, -0.43, and -0.28. After IPW, all SMDs dropped to near zero (≤ 0.02), and variance ratios approached 1, indicating successful balance. This was supported by the overidentification test ($\chi^2 = 1.40$, $p = 0.8443$). Density plots (Figures 11 & 12) for age and BMI showed strong alignment between groups after weighting.

***Figure 11*: Balance Plot for age Before and After Weighting**

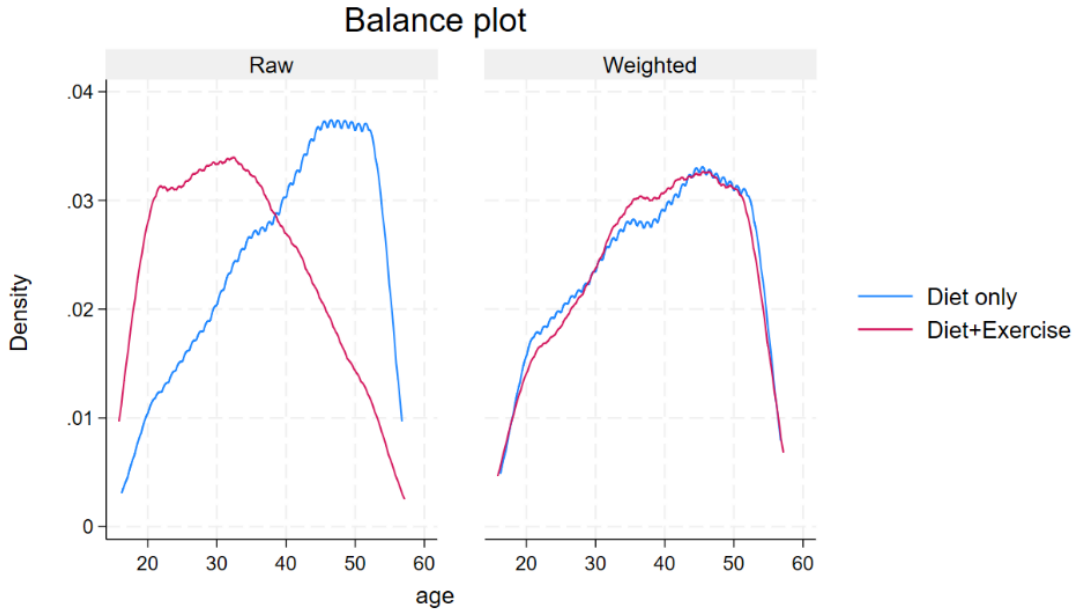
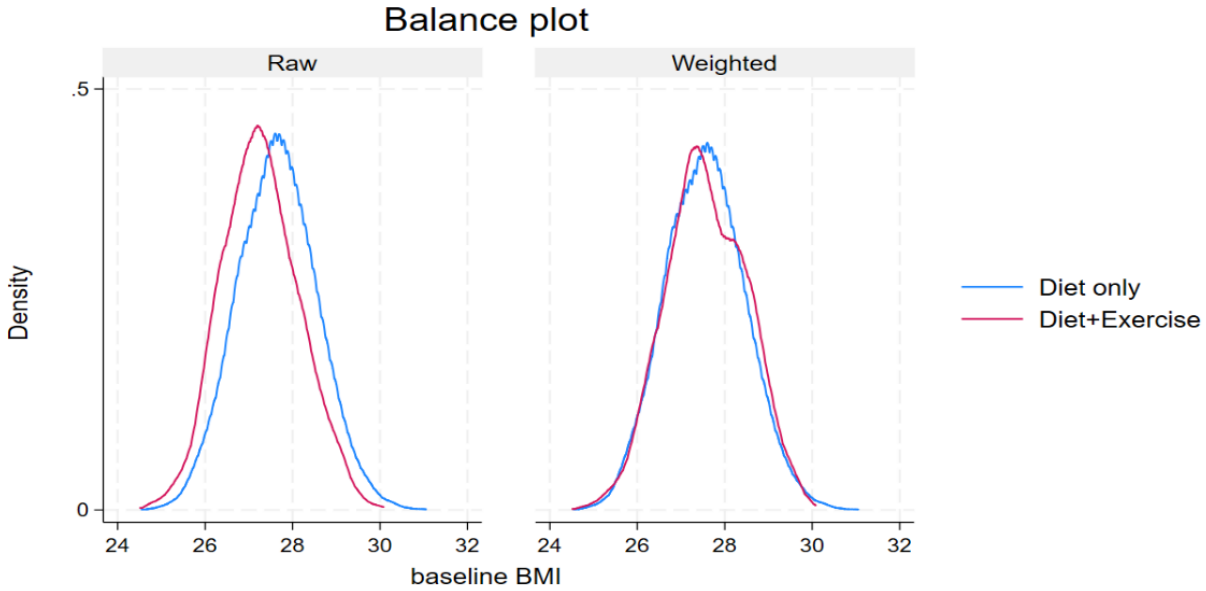


***Figure 12*: Balance Plot for baseline BMI Before and After Weighting**



Propensity score overlap (Figure 13) was adequate. While Diet-only participants clustered at lower scores, both groups shared common support between 0.2 and 0.6, minimizing extrapolation risk and supporting internal validity. Weighted scatterplots of BMI change by age and baseline BMI (Figures 14 & 15) showed good covariate coverage and similar trends across groups. Together, these diagnostics indicate IPW assumptions were met and the treatment effect can be interpreted causally.

***Figure 13*: Balance Plot for baseline BMI Before and After Weighting**

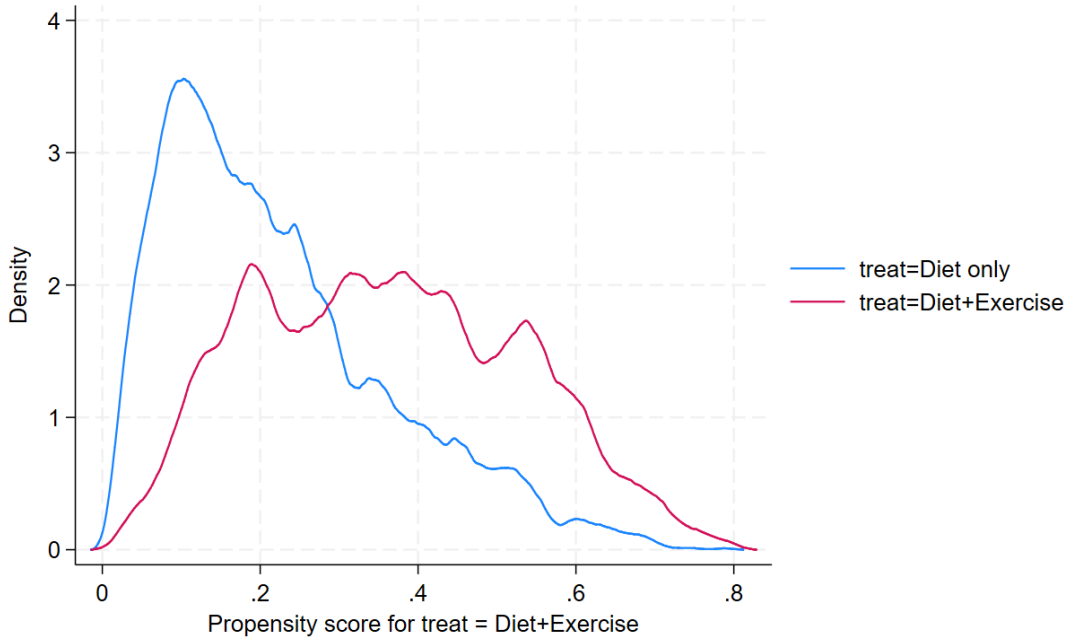


***Figure 14*: Weighted BMI Change by Age and Treatment Group (IPTW-adjusted)**

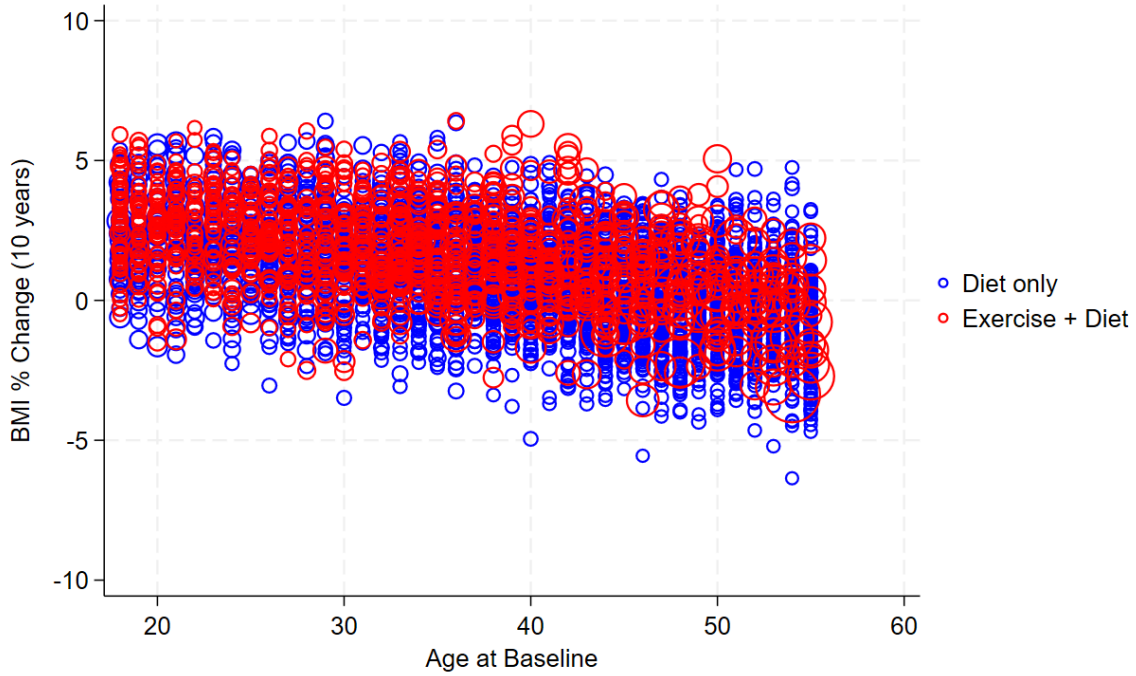
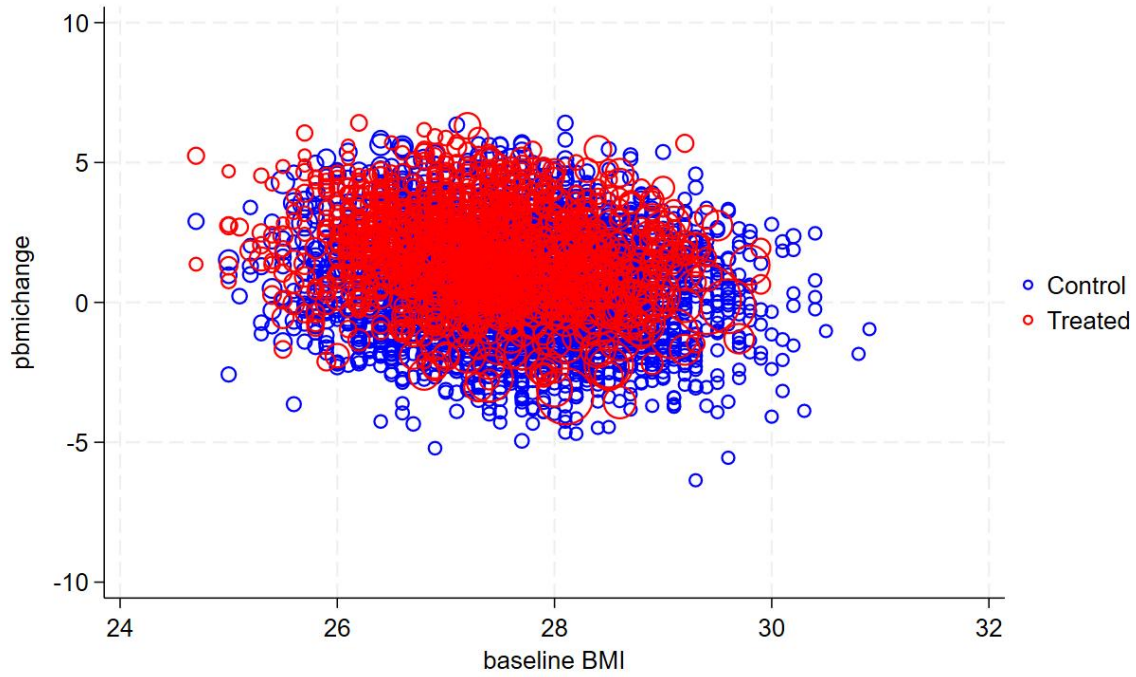


***Figure 15*: Weighted BMI Change by baseline BMI and Treatment Group (IPTW-adjusted)**



iv) What can you conclude about the long-term effects of the exercise and a dietary intervention vs the dietary intervention alone? Discuss whether the collection of additional variables and the use of other statistical models/methodology could change your conclusions. [10 marks]

The combined intervention led to significantly greater BMI reduction, supported by consistent estimates from regression adjustment (ATE = 0.441) and IPW (ATE = 0.529).

However, interpretation remains uncertain, as both methods assume no unobserved confounding(1). Unmeasured factors—like motivation, adherence, social support, psychological health, and environment—could bias results if they influence both treatment and BMI change.

RA is sensitive to model misspecification(2); omitting nonlinear terms (e.g., age²) or interactions (e.g., treatment × baseline BMI) may bias estimates. Its modest R^2 (~0.30) suggests substantial unexplained variation. IPW requires well-specified propensity scores. Large weights (up to 45) and limited overlap are concerning, though trimming improved the ATE (0.551).

Given these limitations, future analyses could use doubly robust (DR) estimators or instrumental variable (IV) approaches. DR offers protection if either model is correctly specified(1), while IVs address unobserved confounding(3)—though no valid IVs were available here.

Section 2: Modelling cost data

i) Describe the statistical models you could fit to the data, providing justification for your choice of models and your selection of variables for inclusion in the models [15 marks]

Several statistical approaches can be considered to model the cost data, each tailoring specific characteristics commonly observed in healthcare costs—namely, non-negativity, positive skewness, and heteroscedasticity.

Ordinary Least Squares (OLS) regression serves as a useful baseline due to its interpretability(4); however, it relies on strong assumptions of normally distributed and homoscedastic residuals, which typically violated in cost data. Therefore, OLS may yield biased or inefficient estimates and even predict negative costs.

A more flexible alternative is the Generalised Linear Model (GLM), which accommodates skewed distributions and non-linear covariate effects. GLMs with a Gamma distribution and either an identity or log link are commonly used. The identity link yields additive effects interpretable as differences in mean cost, while the log link models multiplicative effects, useful for understanding cost ratios.

Two-part model is suitable when many individuals incur zero costs: the first part models the probability of incurring any cost (e.g., via logistic regression), and the second part applies a GLM to positive costs only.

Variable selection is guided by theoretical relevance and empirical performance(4,5). Key covariates include treatment group, age, gender, baseline BMI, chronic stress, and fast-food proximity, with potential nonlinear terms (e.g., age²) or interactions (e.g., treatment × age). Covariates are retained if they confound the treatment effect or improve model fit, based on diagnostics such as AIC, residual plots, and link tests.

ii) Given you have fitted a number of models, describe the process you would use to select or recommend the most appropriate model. [10 marks]

A structured selection process is essential to identify the most appropriate model (4,5). The model should align with the objective—typically estimating population mean costs—and reflect data features such as positive skewness, non-negativity, and heteroscedasticity. Hence, GLMs (with Gamma distribution and identity or log link) or two-part models are often more appropriate than OLS.

After fitting candidate models (e.g., OLS, GLM, two-part), selection is based on statistical performance, diagnostics, and interpretability. Fit statistics like AIC and BIC guide comparisons, with lower values indicating better fit. Residual and Q-Q plots, along with heteroscedasticity tests (e.g., Breusch-Pagan), assess assumption validity. Predicted vs. observed plots evaluate calibration, and out-of-sample validation may be used where feasible.

Interpretability and policy relevance are also important. Identity-link GLMs may be preferred when absolute cost differences matter (5). When models perform similarly, the parsimonious option with sufficient explanatory power is preferred. Ultimately, the selected model should balance fit, assumption validity, and clarity to support sound economic decision-making.

Reference List

1. Funk MJ, Westreich D, Wiesen C, Stürmer T, Brookhart MA, Davidian M. Doubly robust estimation of causal effects. *Am J Epidemiol*. 2011;173(7):761–767.

2. Chin BH. Regression Adjustments for Estimating the Global Treatment Effect in Experiments with Interference [Internet]. 2022 [cited 2025 Mar 31]. Available from: https://arxiv.org/abs/2202.01169

3. Basu A, Heckman JJ, Navarro-Lozano S, Urzua S. Use of instrumental variables in the presence of heterogeneity and self-selection: an application to treatments of breast cancer patients. *Health Econ*. 2007;16(11):1133–1157.

4. Barber J, Thompson S. Multiple regression of cost data: use of generalised linear models. *J Health Serv Res Policy*. 2004;9(4):197–204.

5. Kearns B, Ara R, Wailoo A, et al. Good practice guidelines for the use of statistical regression models in economic evaluations. *Pharmacoeconomics*. 2013;31(8):643–652.

Appendix : Stata Code

```
/***************
 Step 0
****************/

clear all

//set working directory
cd "C:\Users\user\Desktop\0403 Further Stat"

capture log close
log using "0403", replace


/***************
 Step 1
****************/
use BMIData_0.dta, clear

label define gender_lbl 1 "male" 0 "female"
label values male gender_lbl

label define treat_lbl 1 "Diet+Exercise" 0 "Diet only"
label values treat treat_lbl

label define stress_lbl 1 "Chronic stress" 0 "/"
label values stress stress_lbl

label define fastfood_lbl 1 "Fastfood neighbourhood" 0 "/"
label values fastfood fastfood_lbl

label variable pid "Patient id"
label variable bmi0 "baseline BMI"
label variable pbmichange  "Percentage BMI change after 10 years"


// what's in the data?
describe

//summary stats
summarize

//summarize detail
summarize,detail
```

```
/***************
 Step 2 Variable analysis
****************/

/* ------------------------------------------------ */
/* Understand the distribution of variables in population  */
/* ------------------------------------------------ */

// Oridanal variables
tabulate male // counts of male/female
tabulate stress //% people under chronic stress
tabulate fastfood // % living in fast-food dense areas
tabulate treat // distribution of treatment groups

//Continuous variables
histogram bmi0, frequency title("Histogram of baseline BMI") ///
    xtitle("baseline BMI") ytitle("Number of participants")
histogram age, frequency title("Histogram of Age") ///
    xtitle("Age") ytitle("Number of participants")

/* ------------------------------------------------ */
/* Distribution across variables in population  */
/* ------------------------------------------------ */
by male, sort:tabulate stress
by male, sort:tabulate fastfood
by stress, sort:tabulate fastfood

graph box age, over(male)
graph box age, over(stress)
graph box age, over(fastfood)
graph box bmi0, over(male)
graph box bmi0, over(stress)
graph box bmi0, over(fastfood)

/***************
 Step 3 Outcomes analysis
****************/
histogram pbmichange, normal title("Histogram of % BMI Change") ///
    xtitle("% Change in BMI after 10 years") ytitle("Frequency")
ttest pbmichange, by(treat)
graph box pbmichange, over(treat) title("BMI Change by Group") ytitle("BMI % Change")

/***************
 Step 4 Variable and outcomes
****************/
// Continuous 1: Age
twoway ///
```

```
    (scatter pbmichange age, mcolor(%30) msymbol(o)) ///
    (lowess pbmichange age, bwidth(0.8) lcolor(red) lwidth(medium)) ///
, ///
    title("BMI % Change vs Age", size(medium)) ///
    xtitle("Age at Baseline", size(small)) ///
    ytitle("BMI % Change", size(small)) ///
    legend(order(1 "Individual Data" 2 "Lowess Trend") ring(0) pos(1)) ///
    graphregion(color(white)) ///
    ylabel(, angle(horizontal)) ///
    xlabel(, angle(horizontal))

// Continuous 2: Baseline BMI
twoway (scatter pbmichange bmi0), ///
title("BMI % Change vs baseline BMI") xtitle("baseline BMI") ytitle("BMI % Change")


// Supplement with Pearson correlation for Continuouscovariates
pwcorr pbmichange age, sig
gen age2 = age^2
pwcorr pbmichange age2, sig
pwcorr pbmichange bmi0, sig

// Binary 1: Gender
graph box pbmichange, over(male)title("BMI % Change by Gender") ytitle("BMI % Change")

// Binary 2: Stress
graph box pbmichange, over(stress)title("BMI % Change by Stress Status") ytitle("BMI %
Change")

// Binary 3: Fast food exposure
graph box pbmichange, over(fastfood)title("BMI % Change by Fast Food Environment")
ytitle("BMI % Change")

// Supplement with t-tests for binary covariates
ttest pbmichange, by(male)
ttest pbmichange, by(stress)
ttest pbmichange, by(fastfood)


/*****************************************************************/
/* SECTION II: Selection on observables       */
/*****************************************************************/

/***************
 Step 1 Naive treatment effect
***************/
```

```
//unadjusted mean bmi change
mean pbmichange, over(treat)

//test defference in mean cost
ttest pbmichange, by(treat)

//keep mean pbmichange by treatment groups
global a = r(mu_1) // stores the control group mean in the global macro a
global b = r(mu_2) // stores the treatment group mean in the global macro b
global diff= $a -$b // stores the difference in mean pbmichange in the global macro diff
display "Difference in unadjusted means: $diff" // displays what is stored in the
global macro diff

/***************
 Step 2 Descriptive stats by treatment group
***************/
/* ------------------------------------------------ */
/* Check if there is imbalance between treatment group   */
/* ------------------------------------------------ */
bysort treat: summarize

/* ------------------------------------------------ */
/* t-tests for continuous variables  */
/* ------------------------------------------------ */
// Continuous 1: Age
ttest age, by(treat)

//by age group(18-55)
gen agegrp = .
replace agegrp = 1 if age < 30
replace agegrp = 2 if age >=30 & age < 40
replace agegrp = 3 if age >=40 & age < 50
replace agegrp = 4 if age >=50
label define agegrp 1 "<30" 2 "30–39" 3 "40–49" 4 "50–55"
label values agegrp agegrp

mean pbmichange, over(agegrp)

twoway (scatter pbmichange age, sort) (lfit pbmichange age), saving("overall", replace)

twoway (scatter pbmichange age if treat==0, sort mcolor(blue) legend(label(1 "Diet only"))
xtitle("Age") ytitle("pbmichange")) ///
(scatteri $a 18 $a 60, recast(connected) lcolor(blue) mcolor(blue) legend(label(2
"Unadjusted average (Diet only)"))) ///
(scatter pbmichange age if treat==1, mcolor(red) legend(label(3 "Diet+Exercise"))) ///
 (scatteri $b 18 $b 60, recast(connected) lcolor(red) mcolor(red) ///
 legend(label(4 "Unadjusted average (Diet+Exercise)"))), saving("pic1", replace)
```

```
// Continuous 2: Baseline BMI
ttest bmi0, by(treat)

//by baseline bmi(24.7-30.9)
gen bmi0grp = .
replace bmi0grp = 1 if bmi0 < 18.5
replace bmi0grp = 2 if bmi0 >= 18.5 & bmi0 < 25
replace bmi0grp = 3 if bmi0 >= 25 & bmi0 < 30
replace bmi0grp = 4 if bmi0 >= 30

label define bmi0grp_lbl ///
    1 "Underweight (<18.5)" ///
    2 "Normal (18.5–24.9)" ///
    3 "Overweight (25–29.9)" ///
    4 "Obese (≥30)"

label values bmi0grp bmi0grp_lbl

mean pbmichange, over(bmi0grp)

twoway (lfit pbmichange bmi0 if treat==0, lcolor(blue)) ///
       (lfit pbmichange bmi0 if treat==1, lcolor(red)) ///
       (scatter pbmichange bmi0 if treat==0, mcolor(blue)) ///
       (scatter pbmichange bmi0 if treat==1, mcolor(red))

twoway (scatter pbmichange bmi0 if treat==0, sort mcolor(blue) legend(label(1 "Diet only"))
xtitle("baseline BMI") ytitle("pbmichange")) ///
(scatteri $a 23 $a 33, recast(connected) lcolor(blue) mcolor(blue) legend(label(2
"Unadjusted average (Diet only)"))) ///
(scatter pbmichange bmi0 if treat==1, mcolor(red) legend(label(3 "Diet+Exercise"))) ///
 (scatteri $b 23 $b 33, recast(connected) lcolor(red) mcolor(red) ///
 legend(label(4 "Unadjusted average (Diet+Exercise)"))), saving("pic1", replace)

/* ---------------------------------------------- */
/* Chi-squared tests for categorical variables  */
/* ---------------------------------------------- */
tabulate male treat, chi2
tabulate stress treat, chi2
tabulate fastfood treat, chi2


/***************
 Step 3 Treatment effect using a linear regression
****************/
//Treatment effect using a linear regression
```

```
//3.1 Naive model(unadjusted)
regress pbmichange treat
estimates store model1

linktest
hettest
rvfplot

//3.2 Full model
regress pbmichange treat age age2 bmi0 male stress fastfood
estimates store model2
ovtest
linktest
hettest
rvfplot

//3.3 Parsimonious Model
regress pbmichange treat age age2 bmi0 male stress
estimates store model3
ovtest
linktest
hettest
rvfplot

//3.4 Age × Treatment
regress pbmichange treat##c.age age age2 bmi0 male stress
estimates store model4
ovtest
linktest
hettest
rvfplot

//3.5 Sensitivity Model：Baseline BMI interaction
regress pbmichange treat##c.bmi0 age age2 bmi0 male stress
estimates store model5
ovtest
linktest
hettest
rvfplot

//test restrictions implied by model2
lrtest model2 model3

estimates table model1 model2 model3 model4 model5, star stats(aic bic r2 N)
```

```
/***************
 Step 4 Regression adjustment (RA)
****************/
teffects ra (pbmichange age age2 bmi0 male stress) (treat)

/***************
 Step 5 Inverse probability weighting (IPW)
****************/

//modelling the propensity score

//5.1 IPW1
probit treat age bmi0 male stress fastfood
estimates store model1
estat gof

//5.2 IPW2
probit treat c.age c.age#c.age bmi0 male stress fastfood
estimates store model2
estat gof

lrtest model1 model2

//5.3 IPW3
probit treat i.agegrp bmi0 male stress fastfood
estimates store model3
estat gof

//5.4 IPW4
probit treat age bmi0 stress
estimates store model4
estat gof

lrtest model1 model4

//5.4 IPW5
probit treat age bmi0 male stress c.age#i.male
estimates store model5

//5.4 IPW6
probit treat age bmi0 male stress c.bmi0#i.male
estimates store model6


// Choose the model with better AIC/BIC and good fit
estimates table model1 model2 model3 model4 model5 model6, star stats(aic bic r2 N)
```

```
//Estimate ATE using IPW (based on chosen PS model)
teffects ipw (pbmichange) (treat age bmi0 stress,probit),aeq


//Check Covariate Balance and Overlap
tebalance overid

tebalance summarize age bmi0 stress,baseline
tebalance summarize age bmi0 stress


// Check overlap (propensity distribution by treatment)
tebalance density age
tebalance density bmi0

teffects overlap, ptlevel(1)


// Visualize weights
predict double pscore,ps tlevel(1)
gen double wei=1/pscore if treat==1
replace wei=1/(1-pscore) if treat==0

twoway (scatter pbmichange age if treat==0 [w=wei], sort mcolor(blue)
msymbol(circle_hollow)legend(label(1 "Diet only")) xtitle("Age at Baseline") ytitle("BMI %
Change (10 years)")) (scatter pbmichange age if treat==1 [w=wei], mcolor(red)
msymbol(circle_hollow) legend(label(2 "Exercise + Diet")) )

summ wei

twoway (scatter pbmichange bmi0 if treat==0 [w=wei], sort mcolor(blue)
msymbol(circle_hollow)legend(label(1 "Control")) xtitle("baseline BMI") ytitle("pbmichange"))
(scatter pbmichange bmi0 if treat==1 [w=wei], mcolor(red) msymbol(circle_hollow)
legend(label(2 "Treated")) )

// Sensitivity test
gen wei_trim = wei
replace wei_trim = 10 if wei > 10

reg pbmichange treat [pw=wei_trim]
```