Alistair Clark
Advanced Quantitative Methods
June 14th, 2024

The Effect of Exercise on Mental Health in Undergraduate Students

Introduction

Mental health issues are common among college students, with the most prevalent issues including substance abuse, anxiety, and mood disorders (Pedrelli et al, 2014). These issues can adversely affect students in many ways, including but not limited to impacting their academic performance, social interactions, and overall well-being. If left untreated, these issues can escalate, potentially leading to academic failure or dropout. College is a transitional period in a young adult's life, and the life skills learned during this time often persist into adulthood (Steineke, 2019). This makes addressing mental health issues in college students all the more integral to their success in the short term and their well-being in the long term.

One common recommendation to combat mental health issues is exercise. Physical exercise not only benefits an individual's overall physical health but mental health as well. This essay explores the causal connection between exercise and mental health among undergraduate college students. This question is important because many college-aged students in the United States are affected by mental health issues (Pedrelli et al, 2014). Understanding the positive impact that exercise has on mental health could have major implications for future health policies and student support services.

Plenty of studies have been conducted on this subject, yet very few establish a causal relationship between exercise and mental health. Instead, these studies tend to focus on establishing correlations between the two variables. While the data in this paper has its limitations, it can still be used to establish a causal relationship. The two methods that will be used in this paper to attempt to establish a causal relationship are an IV regression model and a fixed effects model.

Data collection took place over seventeen days, from May 24th, 2024 to June 10th, 2024. The two main channels through which the survey was distributed were flyers posted in the recreation center at UCSC and the social media platform Snapchat. During this period, 180 respondents attempted to finish the survey, of which 142 were successful.

The results from the IV regression show that increasing the rating of an individual's overall physical activeness had a statistically significant impact on the mental health of the individual. The fixed effects model also showed that exercising had a positive association with mental health levels.

Literature Review

Theoretical Framework. To establish motivation for studying the effect of exercise on mental health, three theories will be discussed: the Endorphin theory, the Distraction theory, and the Social Interaction theory. These will provide logical explanations for why exercise would have a causal effect on mental health. These frameworks were laid out by Lynette L. Craft, Ph.D., and Frank M. Perna, Ed.D., Ph.D., in their article titled 'The Benefits of Exercise for the Clinically Depressed'.

Endorphin Theory. The endorphin hypothesis is a theory that physical exercise will lead to the release of endorphins. Endorphins are chemicals produced in the central nervous system and the pituitary gland and act as natural painkillers/mood enhancers. Endorphins are often referred to as "feel-good" hormones because they can produce a sense of euphoria and general well-being. This idea helps explain why individuals often feel mental improvement after engaging in physical activity.

Regular physical activity increases the production of endorphins, which can help reduce symptoms of depression and anxiety. Studies have shown that individuals who engage in regular exercise have higher levels of endorphins in their bloodstream, leading to improved mood and decreased feelings of stress (Craft & Perna, 2004). This theory provides a biological explanation for the mental health benefits of exercise.

Distraction Theory. The distraction hypothesis is a theory that physical activity diverts attention from stressors and negative thoughts to focus on the physical task. By engaging in exercise, individuals temporarily shift their attention away from their worries and concerns. This can provide relief and reduce anxiety. Given that exercise is routinely performed, this can provide a significant opportunity for individuals under major mental stress to distract themselves for a short period during the day. Unlike the endorphin theory, which highlights a biomechanical explanation, the distraction theory demonstrates the cognitive aspect of how exercise can impact mental health.

Social Interaction Theory. The social interaction theory claims that a reason for physical activity being linked to improved mental health is that exercise, especially group exercise, is highly correlated with social interaction. Often, individuals suffering from mental health issues such as depression and anxiety can be socially isolated due to their condition. This engagement with other people can foster a sense of community and belonging, ultimately improving their mental health.

Final Analysis. Each of these theories provides insight into the possible mechanisms behind exercise's influence on mental health. Together, they create a cohesive framework in which exercise can benefit mental well-being.

The first relevant piece of research I focused on was conducted by Tanner Steineke at the University of South Dakota. The research question addressed in Steineke's study is nearly identical to the research question of this study, which is the effect of exercise on college students' overall health. Steineke's paper examines the correlations between exercise and multiple standard metrics of a person's overall health, including mental health, physical activity, overall health, diet, and sleep habits.

To collect the data necessary for this study, Steineke used SONA, a platform used by researchers at colleges and universities that takes advantage of the fact that undergraduate students regularly participate in research studies in exchange for college credit. He collected data on 91 individuals, and the survey questions were taken from SF-36 and the National College Health Assessment, which are known for their accuracy in measuring mental and physical health.

The methods used in this paper were simple OLS regression analysis to find correlations between variables. The paper only mentions relationships between variables and does not discuss any causal relationships between exercise and any of the response variables. This was the largest pitfall of this paper, as the correlations developed were significant but did not result in any meaningful conclusions.

As his research question was multifaceted, Steineke had multiple hypotheses about the findings in the data:

Hypothesis 1 stated that physical activity will be related to better self-perceived overall health (H1a) and mental (H1b) health… Hypothesis 2 states that college students will engage in lower overall levels of physical activity than they did in high school… Hypothesis 3 stated that college students' physical activity levels will be related to a healthy diet (H3a) and sleep habits (H3b) (Steineke, 2019).

All of the hypotheses were found to be true. Most relevant to the subject of this paper is Hypothesis 1b, which had a statistically significant p-value of 0.02. Additionally, the study found that physical activity accounted for six percent of the variance in mental health.

This paper found strong correlations between numerous variables that are related to this paper's research question. However, the methods used did not include finding a causal relationship, which significantly diminishes the value of the work done. Given that this research paper finds a causal relationship between physical activity levels and mental health levels, it is safe to say that my methods make a significant contribution to the work done on the topic thus far.

In another study, Rodriguez-Romo et al. (2023) investigated the relationship between physical activity and mental health in undergraduate students. More specifically, they investigated how different levels of physical activity influence mental health. The study emphasizes the connection between positive mental health and healthy behaviors that promote well-being, rather than a lack of negative symptoms.

The study collected data from 847 undergraduate students across Madrid, Spain. A survey was used and the time frame was January to May 2022. To measure mental and physical health, the Global Physical Activity Questionnaire (GPAQv2) and General Health Questionnaire (GHQ-12) were used.

The survey was cross-sectional and used stratified sampling to ensure representation from various types of universities and disciplines. The questionnaire QPAQv2 measured three types of physical activity: occupational, commuting, and leisure time. The questionnaire GHQ-12 was used to detect stress levels and any potential mental health concerns. The statistical analyses included summary statistics and logistic regression to find associations between physical activity and mental health.

The study found significant associations between high levels of physical activity and better mental health scores. Some of the results include but are not limited to: Higher commute and leisure time physical activity were positively correlated with better mental health and moderate levels of occupational physical activity were positively correlated with better mental health. The results of the logistic regression were that students who engage in high levels of leisure time physical activity and moderate levels of occupational physical activity were predicted to have around 48% lower odds of experiencing mental health issues when compared to students who have low levels of physical activity.

In conclusion, Rodriguez-Romo's study shows that physical activity, which is a broader category than exercise, is significantly associated with higher levels of mental health. While this study may only give associations between physical activity and mental health, it motivates further research into the subject.

Data

Collection Methods. The data used in this paper was collected through an online survey advertised to UCSC students via two main media forms. The primary method involved posting flyers at the UCSC Recreation Center. After obtaining permission from Brooke Yarrington, the Assistant Athletic Director of Membership and Operations, flyers were placed in high-visibility areas such as the main lobby, the basketball courts, and facility exits. Additionally, a flyer was posted at the bus stop next to the recreation center to attract the attention of people waiting for buses, encouraging them to complete the survey during their wait.

The second method of advertising the survey was through the UCSC Snapchat Group Story. This social media channel allows anyone with a UCSC email to post pictures and links to a shared feed, which remains viewable for one day. The flyer and a link to the survey were posted twice on this channel, once on May 24 and again on June 10. It is estimated that these two posts accounted for about half of the data collected from the survey.

The motivation behind collecting data for this paper instead of using existing data was that there was much more flexibility in what variables would be available. As the literature review shows, very few of the papers that have covered this topic establish a causal relationship between exercise and mental health, and therefore their datasets would not be equipped to do so. Creating a survey allowed questions to be asked that were specifically addressing issues such as endogeneity.

Summary Statistics of Participants. The participants were 142 undergraduate students at UCSC. All participation was voluntary and all students who participated consented to their information being used in this paper. The average age of respondents, collected in Q1, was 19.5 years old. Results from Q2 indicated that roughly 60% of respondents were female, 36% were male, 3.5% were nonbinary/third gender, and 0.7% preferred not to answer. From Q3, ~24% of respondents self-identified as Asian, ~3% Black/African American, ~31.7% Hispanic/Latino, ~1.4% Native American, ~36.6% White, and ~3.5% of respondents indicated Other. In total, there were 33 questions on the survey, of which 29 were used in the final regression analysis as listed below.

Key Dependent Variables. The variables mental_health and daily_mental_health used the following Likert scale:

1: Severe mental health issues significantly impact daily life.
2: Frequent mental health problems noticeably affect daily activities.
3: Occasional stress or anxiety, generally fair mental health.
4: Generally good mental health with minor, infrequent stress.
5: Optimal mental health with no significant issues, very positive.

Key Independent Variables.
Exercise_frequency was measured on a 0-7 to indicate the number of days per week.
Exercise_duration and daily_exercise_duration used the following multiple-choice questions: 0-20, 21-40, 41-60, 61-80, 81-100 and 100+ minutes.
Exercise_intensity and daily_exercise_intensity were measured using the following scale:

1: Very Light (eg. stretching)
2: Light (e.g., walking, stretching)
3: Moderate (e.g., brisk walking, light jogging)
4: Vigorous (e.g., running, high-intensity interval training)

Instrumental Variables. The two variables used as instruments for the IV regression were friends_exercise and friends_exercise_intensity. Friends_exercise was measured with the following options: Never, Once a month, Once a week, 2-3 times a week 4-5 times a week, and daily. Friends_exercise_intensity was measured with the following scale:
1: Very light (e.g., stretching)
2: Light (e.g., walking)
3: Moderate (e.g., jogging)
4: Vigorous (e.g., running, high-intensity interval training)
5: Very vigorous (e.g., competitive sports)

Control Variables. Demographic Variables: age: Age in years (Discrete), gender: Gender (Categorical), race_ethnicity: Race/Ethnicity (Categorical)

Dietary Habit Variables: meals_per_day: Meals per Day (Discrete), fruits_veggies_consumption: Fruit and Vegetable Consumption (Likert)

Sleep Pattern Variables: sleep_hours: Hours of Sleep (Multiple Choice), sleep_consistency: Sleep Consistency (Likert)

Social and Emotional Well-being Variables: physical_health: Overall Physical Health (Likert), happiness: Happiness (Likert), stress_anxiety: Stress/Anxiety (Likert), social_satisfaction: Social Satisfaction (Likert), social_support: Frequency of feeling supported by friends and family (Likert), loneliness: Frequency of feeling lonely or isolated (Likert)

Peer Influence on Physical Activity Variables:
friends_exercise_frequency: Frequency of friends' exercise (Ordinal), friends_exercise_intensity: Intensity of friends' exercise (Ordinal), social_exercise_participation: Social Exercise Participation (Ordinal), exercise_encouragement: Exercise Encouragement (Ordinal)

Daily Variations (Collected for 10 days):
daily_meals: Daily Meals (Discrete), daily_fruits_veggies: Daily Fruit/Vegetable Consumption (Binary), daily_sleep_hours: Daily Sleep Hours (Multiple Choice), daily_mental_health: Daily Mental Health (Likert), daily_physical_health: Daily Physical Health (Likert), daily_stress_anxiety: Daily Stress/Anxiety (Likert)

For ease of use, each observation was converted to numeric. Ordinal variables such as the Likert scale or ordered categorical variables were converted to a 1-5 scale. Categorical variables were converted to sets of indicator variables, such as gender and race/ethnicity. To assess the effect of exercise thoroughly, a composite exercise variable was created, combining exercise_frequency, exercise_duration, and exercise_intensity. The standard deviations were calculated for each observation, summed, and divided by the number of variables.

Limitations. One of the largest concerns in the panel data collected is the presence of recall bias.While some variables, such as daily exercise, may be easy to recall, it may be more difficult to remember the amount of stress experienced over a week in the past. Another factor that may influence the consistency of the results is the reliance on self-reported data. Since individuals are incentivized to finish the study to enter the raffle, they may be persuaded to rush through the questions that they don't want to spend time answering. This could potentially cause the data to become inconsistent with the reality of their situations.

Another issue with how the survey was conducted is that the survey was distributed disproportionately to students who exercise frequently. This was due to the fact that the flyers were largely posted in the recreation center, which active students use. This very likely introduced sampling bias as the people who were more likely to answer were those who exercised more frequently than the average UCSC student. Similarly, while the $100 gift card incentivized more students to participate in the survey, it may have introduced a bias. Since the participation was voluntary, students who were more economically challenged may be more likely to fill out the survey. This introduces yet another potential confounding variable regarding the representativeness of the sample.

Comparing the questions from the survey used for this paper and the surveys used for other papers in the literature, it becomes clear that the questions asked in this paper were not standardized to any commonly agreed-upon metric. In Steineke's paper, two common measurement tools that were used were Sf-36 and the National College Health Assessment. These survey questions have been studied for reproducibility of results and consistency among populations, while the questions asked in the results from my study have not.

Methodology

Fixed Effects Model. To analyze the effect of exercise on mental health, a fixed effects model can be used. By allowing each individual to have their own intercept, all time-variant characteristics are controlled for, limiting potential bias. The panel data contains ten days of data for each participant, totaling to 1420 observations. The response variable, daily_mental_health was measured on the Likert scale using the same description as mentioned in the data section for mental_health. The explanatory variables were daily_exercise, daily_exercise_minutes, daily_exercise_intensity and daily_exercise_with_friends. All other time-variant variables were considered controls.

Three main assumptions are made when attempting to establish a causal effect using a fixed effects model. The most important of which is exogenous variables. Since we do not have any way to control for endogeneity in a fixed effect model, none of the explanatory variables must be correlated with the error term. The second assumption is that there are no omitted variables that are time-varying. Finally, the data must also be robust to heteroskedasticity and autocorrelation.

Since there is a high chance of autocorrelation in this data, clustered standard errors will be used in the final panel data regression. Some possible limitations of this method include omitted variables and measurement errors. Given that only 11 variables were collected, it is safe to assume that more time-varying variables are unaccounted for in this model. Compared to other models in the previous studies examined in this paper, there is more potential for bias in the results.

Instrumental Variable Regression Model.

In order to address concerns of endogeneity in the dataset, instrumental variables can be used. Using two stage least squares regression (2SLS), we can isolate exercise_composite's effect on the dependent variable mental health. The variables that will be used as instruments are friends_exercise and friends_exercise_intensity. These variables were chosen because they can be assumed to be uncorrelated with individuals mental health but correlated with their exercise habits.

The first stage of 2SLS is regressing the instruments on the endogenous variable. This tests the inclusion restriction requirement. As long as the instruments are significantly correlated with the endogenous variable, they satisfy the requirement. The second stage is conducting the iv regression using the instrumental variables. In cases where there are two instruments, overidentification tests can be used to test the exclusion restriction.

Some potential shortcomings of this method are not being able to identify a strong instrument. In the case that the instrument(s) do not satisfy either the inclusion or exclusion restrictions, the results they produce usually become biased. As long as the inclusion and exclusion criteria are met, the instruments are considered strong instruments and can be used to infer causality of the explanatory variable on the dependent variable.

Results

Fixed Effects Model Results. To test whether the random effects model or fixed effects model would be more appropriate for the data, the Hausman test was used. Comparing these two models returned the following table.

Table 1: Hausman Test: Fixed Effects vs. Random Effects

	Coefficients		Difference	
Variable	(b) fe	(B) re	(b-B)	Std. Err.
daily_exercise	-0.0472689	-0.1034573	0.0561884	0.0346362
daily_exercise_minutes	0.0113285	0.0400353	-0.0287068	0.0097809
daily_exercise_intensity	-0.0114907	-0.0053147	-0.006176	0.0106538
daily_exercise_encouragement	-0.0390649	-0.070692	0.0316271	0.0214941
daily_meals	0.0051501	0.0259155	-0.0207654	0.0217512
daily_fruits_veggies	-0.1008164	-0.0554651	-0.0453514	0.0160304
daily_sleep_hours	-0.025056	0.0051316	-0.0301876	0.0115743
daily_physical_health	0.0946789	0.2042852	-0.1096063	0.0159687
daily_stress	-0.2198543	-0.2476792	0.0278249	0.0103186
daily_exercise_with_friends	0.2026393	0.2003613	0.002278	0.0107392
daily_exercise_motivation	0.0252365	0.019953	0.0052835	0.0088304

Note: b = Consistent under H0 and Ha; obtained from xtreg.
B = Inconsistent under Ha, efficient under H0; obtained from xtreg.
Test of H0: Difference in coefficients not systematic
chi2(11) = (b-B)'[(V_b-V_B)$\hat{}$(-1)](b-B) = 62.82
Prob >chi2 = 0.0000

Based on the results, it can be concluded that the fixed effects model is the more appropriate model to use. The Hausman test indicates that the fixed effects model is appropriate due to the presence of unobserved variables correlated with the regressors. One potential source of bias is reverse causality, which would happen if mental health influences the regressors. This is especially plausible in the fixed effects model given that responses were recorded daily. Another source of bias could be autocorrelation, which is also caused by daily responses being recorded. To mitigate this bias, we can include cluster standard errors in the final regression model.

Table 2: Fixed-Effects Regression: Daily Mental Health

	daily_mental_health					
Variable	Coefficient	Robust Std. Err.	t	P>|t|	[95% Conf. Interval]	
daily_exercise	-0.0472689	0.1458571	-0.32	0.746	-0.3356184	0.2410805
daily_exercise_minutes	0.0113285	0.0569766	0.20	0.843	-0.1013103	0.1239673
daily_exercise_intensity	-0.0114907	0.0618039	-0.19	0.853	-0.1336727	0.1106913
daily_exercise_encouragement	-0.0390649	0.0856601	-0.46	0.649	-0.2084092	0.1302793
daily_meals	0.0051501	0.0499054	0.10	0.918	-0.0935094	0.1038096
daily_fruits_veggies	-0.1008164	0.056973	-1.77	0.079	-0.2134481	0.0118153
daily_sleep_hours	-0.025056	0.0456302	-0.55	0.584	-0.1152638	0.0651517
daily_physical_health	0.0946789	0.043227	2.19	0.030	0.009222	0.1801358
daily_stress	-0.2198543	0.0416037	-5.28	0.000	-0.3021019	-0.1376067
daily_exercise_with_friends	0.2026393	0.0688355	2.94	0.004	0.0665562	0.3387223
daily_exercise_motivation	0.0252365	0.0374014	0.67	0.501	-0.0487035	0.0991765
_cons	3.905392	0.2808569	13.91	0.000	3.350157	4.460627

Number of obs = 1,420
Group variable: ID
Number of groups = 142
R-squared: Within = 0.0861, Between = 0.3800, Overall = 0.2512
F(11,141) = 5.06
Prob >F = 0.0000
sigma_u = 0.64656707, sigma_e = 0.62133815, rho = 0.51989023
(Std. err. adjusted for 142 clusters in ID)

Based on these results, we find no significant association between daily_exercise, daily_exercise_minutes, or daily_exercise_encouragement. However, one of the four key independent variables ended up being statistically significant in the regression, which was daily_exercise_with_friends. This variable was associated with a .2 higher daily mental health score and had a t-score of 2.94. Relating these results to the theoretical framework discussed earlier in the paper, we can see evidence in support of the social interaction hypothesis. This hypothesis theorizes that exercise can improve mental health by allowing individuals to feel a sense of community or belonging, which would be the case when exercising with friends.

This daily_exercise_with_friends had the second-largest magnitude, right behind daily_stress, which used a Likert scale to measure how much stress each individual experienced on a given day. Despite not finding a significant correlation between daily_exercise, daily_exercise_minutes, or daily_exercise_encouragement and daily_mental_health, daily_exercise_with_friends would be an interesting variable to explore in further research.

Instrumental Variable Regression Model Results. The explanatory variable for this regression model is exercise_composite, which was created by combining three different measures of physical activity (details are provided in the data section). This variable was found to be endogenous by the endogeneity test (Table 6). To mitigate this bias, two instruments were used. These were the variables friends_exercise and friends_exercise_intensity. To test for inclusion, first stage IV was used. Since two instruments are used in the regression, the overidentification test can be used to test if the exclusion restriction is satisfied.

Table 3: Instrumental Variables 2SLS Regression

	mental_health			
Variable	Coefficient	Std. Err.	z	P>|z|
exercise_composite	0.5049436	0.1253108	4.03	0.000
age	0.0514623	0.0152654	3.37	0.001
gender_female	2.967084	0.2676837	11.08	0.000
gender_male	2.886565	0.2410205	11.98	0.000
gender_nonbinary	2.386279	0.2639885	9.04	0.000
gender_prefer_not_say	0 (omitted)			
race_asian	-0.0977538	0.045047	-2.17	0.030
race_black	-0.1120128	0.1015745	-1.10	0.270
race_hispanic	-0.0542156	0.0414991	-1.31	0.191
race_native_am	-0.6255213	0.1464042	-4.27	0.000
race_other	-0.0630585	0.1010547	-0.62	0.533
race_white	0 (omitted)			
meals_per_day	-0.0367269	0.0332437	-1.10	0.269
fruits_veggies_per_day	0.1216177	0.0431779	2.82	0.005
sleep_hours	0.1417732	0.0282575	5.02	0.000
sleep_consistency	-0.1161299	0.0257765	-4.51	0.000
happiness	0.37331	0.0352207	10.60	0.000
stress	0.3333585	0.0241355	13.81	0.000
social_satisfaction	0.002699	0.0237385	0.11	0.909
social_support	-0.0394924	0.0269131	-1.47	0.142
loneliness	0.1568717	0.0263975	5.94	0.000
_cons	-3.456088	0.4249415	-8.13	0.000

Instrumented: exercise_composite
Instruments: age, gender_female, gender_male, gender_nonbinary, race_asian, race_black, race_hispanic, race_native_am, race_other, meals_per_day, fruits_veggies_per_day, sleep_hours, sleep_consistency, happiness, stress, social_satisfaction, social_support, loneliness, friends_exercise, friends_exercise_intensity
Number of obs = 1,420, Wald chi2(19) = 1799.25, Prob>chi2 = 0.0000, R-squared = 0.5007, Root MSE = 0.59471

The results of the regression analysis were that exercise_composite had a statistically significant impact on the dependent variable, mental_health. The coefficient of exercise_composite was .504, meaning that a one-unit increase in exercise_composite caused an increase of .504 in an individual's mental health levels.

Other significant coefficients include gender, happiness, stress, and loneliness. Female-identifying individuals had the highest expected mental health levels, followed by males, followed by nonbinary, and finally, those who preferred not to answer.

Happiness, stress, and loneliness were measured on a Likert scale with positive behaviors rated at 5 and negative behaviors rated at 1.

When dealing with instrumental variables, it is important to check whether the inclusion and exclusion restrictions have been satisfied. In the table below, the regression is tested for the inclusion restriction.

Table 4: First-Stage Regression Summary Statistics

	exercise_composite			
Variable Prob >F	R-sq.	Adjusted R-sq.	Partial R-sq.	F(2,1399)
exercise_composite 0.0000	0.2100	0.1987	0.0310	22.3704

Minimum eigenvalue statistic = 22.3704
Critical Values:
2SLS size of nominal 5% Wald test: 19.93, 11.59, 8.75, 7.25
LIML size of nominal 5% Wald test: 8.68, 5.33, 4.42, 3.92

The results are that the two instrumental variables are significantly correlated with the explanatory variable. This means that the inclusion restriction is satisfied. Next, the exclusion restriction is tested. This can only be done when two variables are used, otherwise, the exclusion restriction must be argued by logic.

Table 5: Tests of Overidentifying Restrictions

	Sargan (score) chi2(1)	Basmann chi2(1)
Statistic	2.05889	2.03139
p-value	0.1513	0.1541

These results show that the two instrumental variables are not significantly correlated with the error term, meaning that they satisfy the exclusion restriction. Finally, we test to make sure that the explanatory variable is indeed endogenous. This test returned a statistically significant p-value, so it is concluded that the explanatory variable is indeed endogenous.

Table 6: Tests of Endogeneity

	Durbin (score) chi2(1)	Wu-Hausman F(1,1399)
Statistic	16.111	16.0549
p-value	0.0001	0.0001

H0: Variables are exogenous

Conclusion

The findings from this research provide significant evidence in support of exercise having a positive impact on college students' mental health levels. By using both fixed effects and instrumental variable regression models, concerns about potential biases and sources of endogeneity were addressed, strengthening the validity of these results.

The results of the fixed effects model revealed a strong positive correlation between exercising with friends on a given day and having higher mental health levels. This is corroborated by the theoretical framework mentioned in the beginning of this paper. While these were the only results that were statistically significant from the fixed effect model, they still demonstrate an interesting relationship.

The instrumental variables regression model made a strong contribution to the claim that the composite variable for exercise frequency, duration, and intensity has a positive effect on mental health levels. Based on the tests done in the results section, the instruments used were considered strong, and allowed the claim of causality to be made.

And based on these results, there are some policy implementations I would recommend to the University. An increase in socially interactive physical education classes could be a good way to increase mental health levels. These types of classes would be very beneficial according to the results, as they would increase likelihood of exercising with friends or roommates as well as increase students exercise frequency.

While this study has provided multiple valuable insights, there are some issues with the data that must be addressed. Relying solely on self-reported data comes with multiple sources of bias such as sampling and recall bias. Collecting non-standard measures of mental health and physical health also came with its own set of issues, as the data was not comparable to other studies.

In future research on this topic, some changes that would need to be made to improve this study would include using standardized questionnaires such as those used in previous studies, collecting data over a longer period, collecting more lifestyle variables and collecting a representative sample of the population. These improvements would strengthen the claims of causality.

References

Pedrelli, P., Nyer, M., Yeung, A., Zulauf, C., & Wilens, T. (2014, August 21). College students: Mental health problems and treatment considerations. PubMed Central (PMC). https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4527955/

Steineke, T. (2019, November 21). The Effect of Exercise on College Students' Overall Health The Effect of Exercise on College Students' Overall Health. USD RED | University of South Dakota Research. https://red.library.usd.edu/cgi/viewcontent.cgi?article=1072&context=honors-thesis#:~:text=Performing%20exercise%20not%20only%20helps,%2C%20and%20brain%2C%20at%20ease

Craft, L. L., & Perna, F. M. (2004). The benefits of exercise for the clinically depressed. PubMed Central (PMC). https://www.ncbi.nlm.nih.gov/pmc/articles/PMC474733/

Rodríguez-Romo, G., Acebes-Sánchez, J., García-Merino, S., Garrido-Muñoz, M., Blanco-García, C., & Diez-Vega, I. (2023). Physical Activity and Mental Health in Undergraduate Students. International Journal of Environmental Research and Public Health, 20, 195. https://doi.org/10.3390/ijerph20010195