Inequality in motion

Comparing socioeconomic and demographic characteristics relating to physical activity

Sarah Collier

Abstract:

This project examines the sensitivity of socioeconomic (SES) status and demographic factors in regards to physical activity levels using Fitbit data from the All of Us database from the National Institute of Health. I am proposing an idea of a simplified economic model that treats physical activity as a function of socioeconomic and demographic characteristics. This model aims to take these intertwined variables and represent them in a measurable way. This work aims to lay the foundation for a future regression model and contribute to the understanding of how disparities in SES status and demographic factors contribute to physical activity levels and ultimately unequal health outcomes.

Introduction:

Physical activity is a well documented determinant of one's health. (Dhuli et al., 2022) Despite its necessity, access to the time and opportunity to be active is unequal. Individuals with a lower SES and in marginalized demographic groups face resource constraints that limit their ability to participate in exercise. A 2002 study by Dr. Vern Seefeldt and colleagues found that "African American and Hispanic American women, workers from lower socioeconomic strata and undernourished individuals are less likely to maintain an active lifestyle." (Seefeldt, Malina, & Clark, 2002)

The Fitbit is a wearable health measuring tool that tracks the metrics of an individual like heart rate, steps, and exercise. Although the Fitbit uses advanced technology, its accuracy may vary from one context to another, but this underestimation is acceptable, on average, for steps and heart rate. (Chevance et al., 2022). Specifically, I will be looking at the heart rate training zones which are accurate in evaluating a person's physiological effort, which is a strong proxy for exercise. (Eckard et al., 2019). In this study this inaccuracy is directly incorporated into my model as an error term which will help distinguish true from observed activity.

The goal of this project is to test a simplified calculus based economic model that can evaluate inequality in exercise specifically using SES status and demographic factors. This is a preliminary step towards a future econometrics model that could more accurately determine the sensitivity between SES status and demographic factors and which has a stronger correlation with low levels of physical activity. This work aims to help understand sources of uneven physical activity that could contribute to larger health disparities.

Conceptual Model:

To separate SES factors and demographic characteristics, I used a simplified calculus based economic model.

$i = index\ per\ person\ in\ dataset\ (person(i))$

A_i^{obs} = *Observed activity per person in dataset (ex daily steps for person (i))*

S_i = *SES per person in dataset*

D_i = *Demographic group of person in dataset*

f(i) = By how much SES and demographic factors combine to determine activity (This variable will be unknown until there is a regression model)

$\varepsilon_i = Error\ term\ for\ unobserved\ influences\ and\ device\ error.$

$$A_i^{obs} = f(S_i, D_i) + \varepsilon_i$$

This model demonstrates the idea that observed physical activity is determined by socioeconomic and demographic factors while also acknowledging both other activity determinants as well as Fitbit inaccuracy. To explore the influence of each input, I will consider how activity would change when SES changes while demographics remain fixed, and vice versa. The equation for the sensitivities are shown below:

$$\frac{\partial A_i^{obs}}{\partial S_i} = \frac{\partial A_i^{obs}}{\partial D_i}$$

This equation takes the derivative of both SES status and demographic groups over activity level. This serves as a step toward a more complete regression model, which could estimate the relative contributions of SES and demographic characteristics while also incorporating the error structure above. The regression model mentioned above, however, is future work and will not be stated in this proposal.

Literature Review:

Existing literature illustrates that physical activity is unequal across populations. Both SES status and demographic factors influence physical activity levels, but few studies look at these factors independently to determine which has a greater impact on activity.

Fitbit data from the All of Us database demonstrates that there are clear demographic differences in physical activity. In the study done by Jeong et al. (2025) differences in ex, age, BMI, and geography affect exercise levels. Their study uses a Gini model to measure the inequality. It also highlights the lack of SES-focused analyses. This serves as an example of how to use All of Us Fitbit data to measure physical activity.

SES factors are important when analyzing exercise rates. According to Guthold et al. (2020), economic inequality is strongly associated with physical inactivity, especially in high- and middle-income countries. This model shows that through eliminating SES disparities, youth physical activity could be increased. This paper supports treating SES as a structural determinant of activity separate from demographics.

Demographic disparities remain influential even after controlling for SES. A national study of racial groups shows that Black, Asian, and Hispanic adults are less likely to be physically active

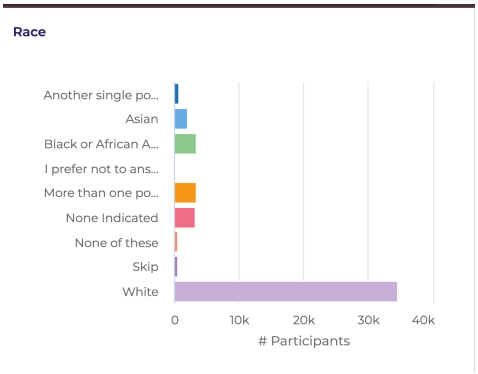
than White adults, and even after adjusting for income, education, and gender. This indicates that demographic factors should be viewed independently beyond SES.

This proposed research aims to use the information from the literature above looking at both SES and demographic variables to determine physical activity levels. Specifically focusing on filling the gap of direct comparison of the two variables.

Data

I will use data from the All of Us Database https://allofus.nih.gov/ . The Fitbit data is outlined in the activity section. I will use the personal survey questions (ex: Can you afford food each month? Is your neighborhood safe?) as a proxy for SES. I will create groups based on responses to each question then evaluate each group. The demographic data is outlined with each Fitbit response.

Although I have not yet finished analyzing all the data, using the image below it is evident that there is a clear difference between the participants with access to Fitbits across race groups. This is important because it supports the importance of demographic factors in physical activity inequalities. I will make sure other variables like age of participants and education level are constant to make sure to only measure demographic and SES factors.



Additionally in my future work I will create visualizations in R to demonstrate my summary statistics. An example of my code is below.

```
library(dplyr)
library(ggplot2)

# 1. Mean Steps by Income
df %>%
  group_by(income) %>%
  summarise(mean_steps = mean(steps, na.rm = TRUE)) %>%
  ggplot(aes(x = income, y = mean_steps, fill = income)) +
  geom_col() +
```

```
  labs(title = "Mean Daily Steps by Income Group",
       x = "Income Group",
       y = "Mean Daily Steps") +
```

Methods

This research idea is a quantitative secondary analysis of data. I will start by computing average steps per participant then grouping by SES variables. I will look at the answers to the survey questions (ex: Can you afford food each month? Is your neighborhood safe?) to create SES groups that I will compare with the demographic variables race and gender. I will then compare the means of activity and plug the information into my model. Next, I will use R to visualize the differences. Future work will then quantify these differences.

Research Design

I will be treating activity as an outcome influenced by SES and demographic factors. Using a calculus based economic model, I aim to simplify the two factors. I will be using exploratory comparisons in this stage and plan to expand my model to eventually build toward an econometric analysis.

Limitations

Although there is a lot of data and research on the issue of physical activity inequality, it is challenging to independently evaluate SES and demographic factors due to their overlap. It is difficult to determine if demographic factors may have influenced SES. Additionally, this work includes an error term to help offset other variables that can influence activity. (ex: psychosocial and environmental factors). In continuation as mentioned above, the Fitbits do not extract step and heart rate count with complete accuracy. This model has attempted to offset these factors through the error term, but in future work they will be more explicitly outlined.

Conclusion

This project proposes a simplified calculus based economic model to analyze how socioeconomic (SES) and demographic factors shape exercise rates. Through modeling activity as a function of these characteristics, plus error measurement, this model allows for clearer separation of influences. Despite existing literature demonstrating SES and demographic related activity inequality, their individual importance remains understudied. This proposal is the preliminary step in creating a future regression model in order to better understand how muli-layered inequality affects physical activity and ultimately an individual's overall health.

Works Cited

Bennie, J. A., De Cocker, K., Teychenne, M. J., Trost, S. G., Biddle, S. J. H., & Salmon, J. (2020). The epidemiology and social context of the physical activity and sedentary behaviour of adults in Australia: a review of the literature. *International Journal of Behavioral Nutrition and Physical Activity*, *17*(1), 1–13. https://doi.org/10.1186/s12966-020-01039-x

Chevance, G., Golaszewski, N. M., Tipton, E., Hekler, E. B., Buman, M., Welk, G. J., Patrick, K., & Godino, J. G. (2022). Accuracy and precision of energy expenditure, heart rate, and steps measured by combined-sensing Fitbits against reference measures: Systematic Review and Meta-analysis. *JMIR mHealth and uHealth*, *10*(4), e35626. https://doi.org/10.2196/35626

Dhuli, K., Naureen, Z., Medori, M. C., Fioretti, F., Caruso, P., Perrone, M. A., Nodari, S., Manganotti, P., Xhufi, S., Bushati, M., Bozo, D., Connelly, S. T., Herbst, K. L., & Bertelli, M. (2022). Physical activity for health. *Journal of Preventive Medicine and Hygiene*, *63*(2 Suppl 3), E150–E159. https://doi.org/10.15167/2421-4248/jpmh2022.63.2S3.2756

Eckard, M. L., Kuwabara, H. C., & Van Camp, C. M. (2019). Using heart rate as a physical activity metric. *Journal of Applied Behavior Analysis*, *52*(3), 718–732. https://doi.org/10.1002/jaba.581

Ma, K., Wang, M., Yin, J., Li, X., Wang, Y., & Wei, S. (2022). Effects of 12-week moderate-to-vigorous physical activity intervention on physical fitness and cognition among college students in different provinces of China. *Frontiers in Psychology*, *13*. https://doi.org/10.3389/fpsyg.2022.9409869

Seefeldt, V., Malina, R. M., & Clark, M. A. (2002). Factors affecting levels of physical activity in adults. *Sports Medicine*, *32*(3), 143–168. https://doi.org/10.2165/00007256-200232030-00001

Tandon, S., Bales, R., Sanner, H., Rastegari, B., Kordy, M., Ghasemian-Dabiri, A., Zakeri, A., Zakeri, H., Kordy, K., & Jentsch, V. (2024). Associations between physical activity and mental health in emerging adulthood: findings from the Student Wellbeing Study. *npj Digital Medicine*, *7*(1), 1–11. https://doi.org/10.1038/s41746-024-01358-4

Undergraduate Research Center. (n.d.). *How to write an abstract*. University of California, Davis. Retrieved from https://urc.ucdavis.edu/how-write-abstract