Winter 2025: Public Policy 775

Problem Set 1: Understanding Relationships in Data

Due: February 3rd at 11:59 PM

Introduction

Having heard about your excellent training at the Ford School, the Department of Health and Human Services (HHS) has called on you to evaluate the effects of social media on mental health. Below are some questions that they may be interested in to help them understand whether TikTok should be banned in the United States for mental health reasons.

Question 0: Correlation vs. Factual Statements (which may be right or wrong)

In Lecture 2, we discussed the differences between statements about correlations and other factual statements that do not convey information about a correlation. Now that you have a deeper understanding that correlation requires variation, consider the following statements. Which ones describe a correlation, and which ones do not?

(a) Most social media influencers with large followings post regularly and consistently.
(b) People who use social media heavily are typically more stressed than those who use it less.
(c) Among content creators, those who focus on educational topics tend to have higher engagement rates than those who focus on personal blogs.
(d) Locations with the highest levels of social media use tend to have greater exposure to misinformation.
(e) Young people who spend more time on social media are more likely to report experiencing cyberbullying compared to those who spend less time.

Question 1: The Cause of Increased Anxiety

A public health committee has hired three experts to figure out the primary cause of increasing anxiety levels in a certain population. The first expert says that social media is the main driver, as it fosters constant comparison, fear of missing out (FOMO), and information overload. The second expert argues that economic uncertainty, such as job insecurity and rising debt levels, is the primary contributor. The third expert claims that societal pressures, including academic and workplace performance expectations, have significantly amplified anxiety levels.

Confused by the seemingly conflicting explanations, the committee hires you, a quantitative analyst, to adjudicate between these three possibilities. What would you tell them? How would you proceed? Is just one of them correct, or could they all be correct?

Question 2: Evaluating the Effectiveness of the TikTok Ban

On January 18, 2025, TikTok became unavailable to U.S. users following the enactment of a law banning the app due to its impact on mental health and national security concerns. The service was restored approximately 14 hours later after President Donald Trump signed an executive order delaying the ban's enforcement by 75 days to allow for negotiations.

You've been tasked with evaluating the effectiveness of this ban in reducing TikTok usage and understanding its broader implications. To do so, you will rely on a self-report questionnaire to collect data. For example, you measure anxiety using the Generalized Anxiety Disorder Scale (GAD-7). Use your knowledge of measurement concepts, as discussed in Lecture 5, to answer the following questions:

(a) Suppose you want to measure the effect of the TikTok ban on anxiety levels using a **before-after analysis**. You plan to compare self-reported anxiety levels among TikTok users before the ban and after the ban. What issue might prevent you from measuring the causal effect of the TikTok ban on the GAD-7 scores?

(b) Suppose that you obtained a causal effect of TikTok on anxiety from the ban. What issues might arise from interpreting changes in the GAD-7 due to the TikTok ban?

(c) Suppose that the negotiations between TikTok and Donald Trump fail, and the app is permanently banned 75 days after being restored. Do you expect the causal effect of this ban to be the same as the one from the 1-day ban? Why or why not?

Question 3: Evaluating a Field Experiment Paying People to Stop Using TikTok

Suppose you conducted a randomized controlled trial (RCT) to study the impact of TikTok use on mental health. You recruited participants by hanging flyers around the University of Michigan's campus, which described the study and invited individuals to participate. After consenting, participants were randomly assigned to one of two groups:

- **Treatment group:** Paid $20 per week to completely stop using TikTok for 6 weeks. They are only paid if they didn't use TikTok at all that week.
- **Control group:** Received no payment and continued using TikTok as usual.

The primary outcome of interest was mental health, measured using the **GAD-7**. The study also added 10 secondary mental health outcomes (e.g., stress, loneliness, social connectedness).

(a) In observational studies comparing people who use TikTok with those who don't, there is often a risk of selection bias. Explain how selection bias could confound your comparison of those who do and do not use TikTok without an RCT. Explain how this RCT design helps eliminate selection bias and why randomization is critical to isolating the causal impact of TikTok use on mental health.

(b) Does this study estimate the effect of a TikTok ban? If not, what does it estimate? Discuss how the design of the study might influence the interpretation of its findings in relation to a TikTok ban.

(c) The researchers use the GAD-7 to measure anxiety levels as their primary outcome. Discuss whether this choice addresses the issue of construct validity in linking TikTok use to mental health. Does randomization solve the issue of construct validity? What challenges remain in ensuring that the GAD-7 fully captures the impact of TikTok use?

(d) The researchers analyze 10 additional secondary outcomes (e.g., stress, loneliness, social connectedness). Explain the risks of **multiple hypothesis testing** in this context. What is the Type I error associated with testing all 11 hypotheses? How might this Type I error rate affect the validity of the study's conclusions?

(e) Participants were recruited via flyers at the University of Michigan's campus. How might differences in observable characteristics (e.g., age, socioeconomic status, pre-existing TikTok usage) between participants and the broader population of TikTok users limit the external validity of the study's findings?

(f) Participants in this study volunteered to take part and agreed to stop using TikTok for $20 per week. How might this selection process limit the generalizability of the

study's findings? Specifically, discuss how people who agree to participate in such a study might differ from the broader population of TikTok users.

(g) Suppose the research team increases the weekly payment to stop using TikTok from $20 to $300. How might the higher incentive attract a different group of participants? Would this improve or limit the generalizability of the study's findings?

Question 4: Assessing Feasibility

Before running the experiment, the HHS approached you and asked you to assess the feasibility of the study. The primary outcome is anxiety, as measured on the GAD-7 scale. Suppose the mean GAD-7 score is a 4.9 with a standard deviation of 4.1. You have enough resources to allocate 300 subjects to treatment and 300 subjects to control.

(a) At conventional levels, what is the minimum detectable effect size for the experiment at conventional levels of significance and power? That is, $\alpha = 0.05$ and $\beta = 0.2$

(b) Suppose that the HHS is worried about the cost of the experiment. To save money, they want to allocate 90% of subjects to the control group and 10% of the subjects to the treatment group. How does that affect your minimum detectable effect size? Why?

(c) The study aims to test 11 mental health outcomes. To account for multiple hypothesis testing, you decide to lower the significance level to $\alpha = \frac{0.05}{11} = 0.0045$. How does this affect your minimum detectable effect size?

(d) Given the MDEs that you calculated, how might you go about deciding whether the study had a large enough sample size to reliably reject substantially significant effects?