



Mock Exam

Social Media and Web Analytics

Spring 2023

Instructions to Candidates

The exam is graded out of **100 points**. Points for each of the three parts are clearly labelled at the beginning of each. The point allocation for each question is also clearly designated.

The **exam duration is 180 mins**.

Write your Student ID and name on every page of the exam.

Answer each question in the text box directly underneath the question. There are additional blank pages at the end of the exam if you need to cross out your work and start again. Clearly label your answers on any of the extra pages.

Answers to the questions should be based on the course material presented in lectures, Lab Sections and readings.

Remember, **your goal is to communicate**. Full credit will be given only to the correct solution which is described clearly. **Convolved and obtuse descriptions might receive low marks, even when they are correct.** Also, aim for concise solutions, as it will save you time and also help you conceptualise the key idea of the problem.

No calculators or dictionaries are permitted.

Any **communication between students during the exam**, no matter whether it is about the exam or not, **is unacceptable** and will likely result in harsh penalties. Attempting to view other student's answers will be treated in a similar manner.

To pass Social Media and Web Analytics, you must receive a grade higher than 5.5/10 for this exam.

Part A: True / False / Uncertain (20 points)

Answer *all* questions. **Each question in Part A is worth 2 points.** One point is for correctly asserting one of True/False/Uncertain and one point is for explaining your answer.

*Where relevant, your answer should relate back to findings of papers discussed in lectures. Precise citations to relevant articles are **not** required.*

Question 1.

Active lurkers have no marketing value.

Question 2.

Assigning website users to two different landing pages (A vs B) based on their previous interactions with the website and then running the regression

$$sales_i = \beta_0 + \beta_1 versionB_i + \varepsilon_i$$

Where $sales_i$ is sales of a user i on a website and $versionB_i$ takes the value 1 if a user sees version B and otherwise takes the value 0, allows the analyst to recover an unbiased estimate of the true causal effect of landing page B versus landing page A.

Question 3.

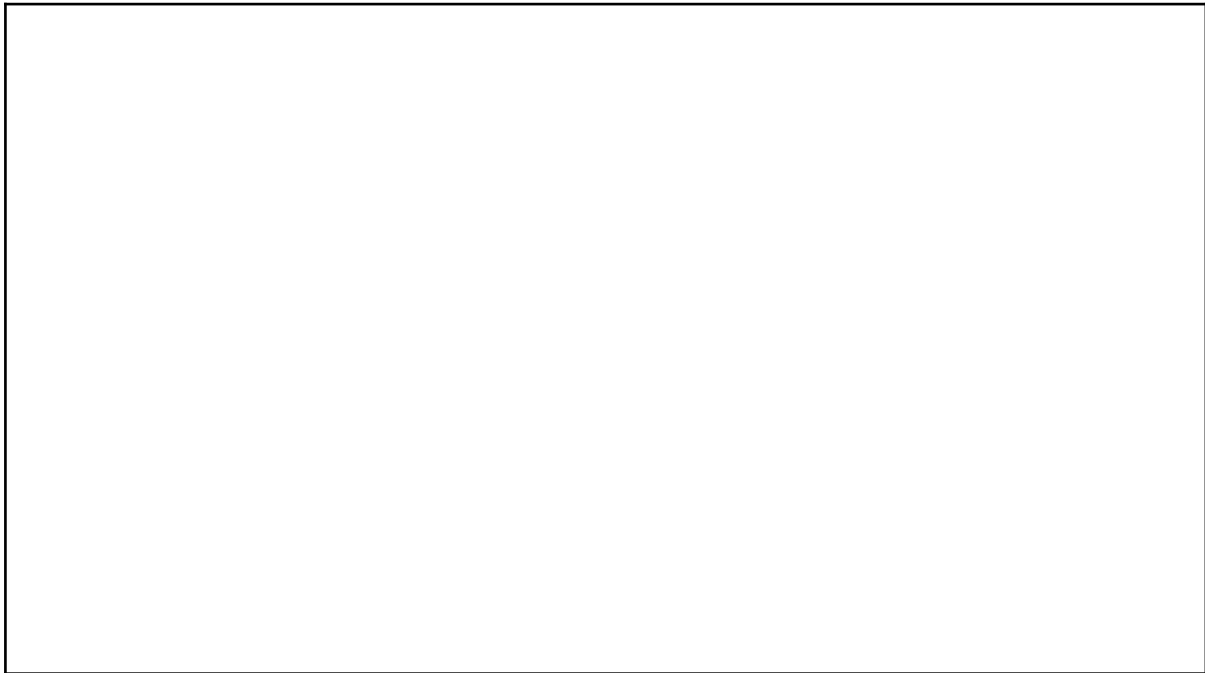
Consumer willingness to pay is the main mechanism through which a drop in online reputation influences consumer buying behaviour on EBay.

Question 4.

Influencers and one-on-one support are equally cost effective at driving product adoption.


Question 5.

When applied to a dataset of social media posts, a Topic Model returns whether each post is classified as positive, negative or neutral.



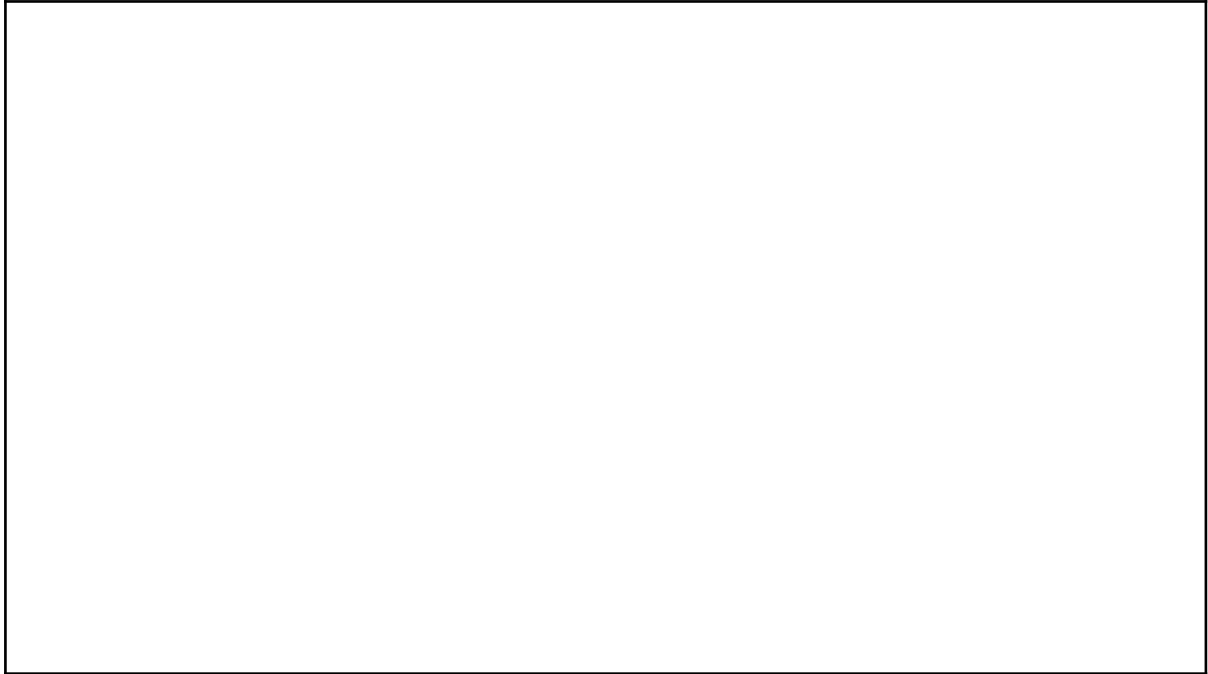
Question 6.

The sentiment within social media posts influences the demand for franchise movies such as X-men and Jurassic Park.



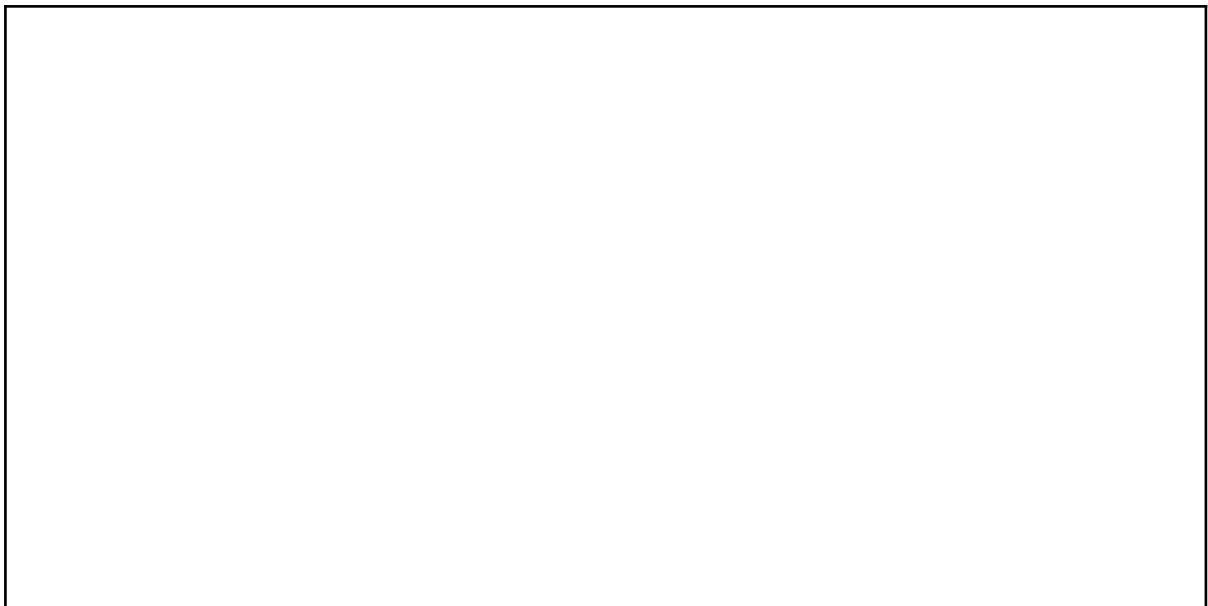
Question 7.

The VADER sentiment lexicon was specifically designed for use with social media text.




Question 8.

Improvised marketing interventions that feature humour are more likely to go viral.



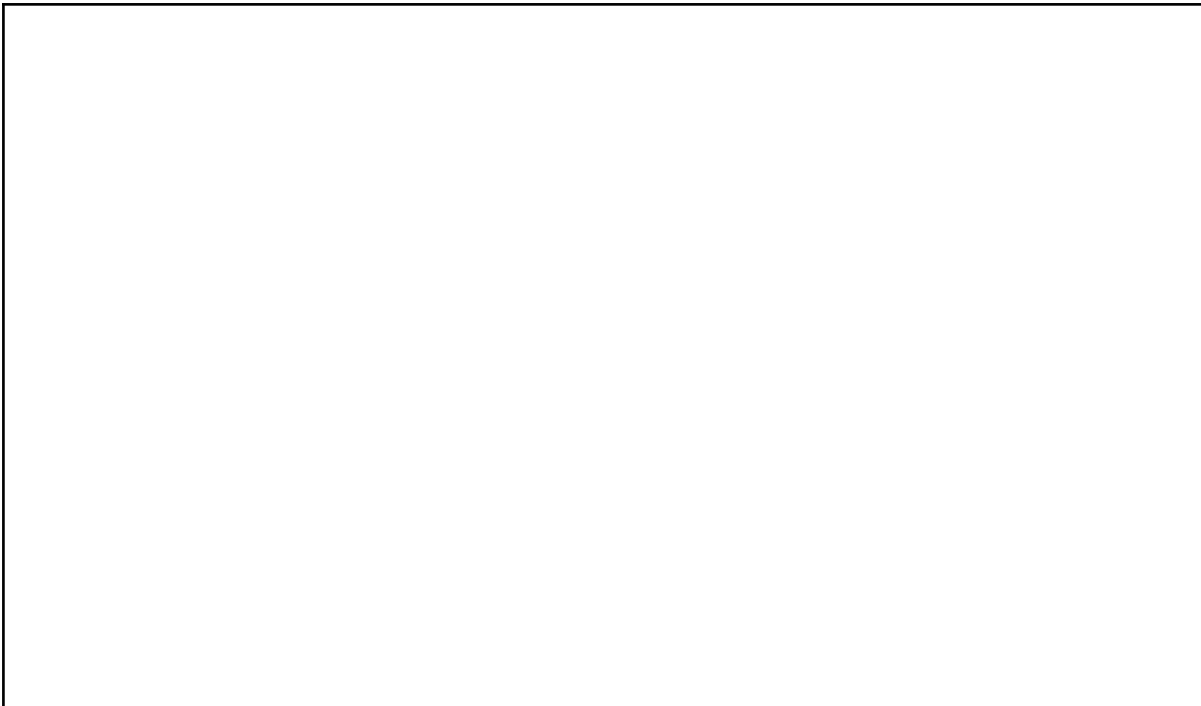
Question 9.

In a network of Twitter users, the edges are defined as connections between users and the nodes are the Twitter users.



Question 10.

Influential retweets of a firm's own tweets increase viewing of TV shows if the firm's tweet is informative.



Part B: Short Answer Questions (40 points)

Answer **all** questions in this section. Point allocations per question are clearly indicated by each question and/or sub-question.

Question 1 (10 points). Difference in Differences.

- (a) [4 points] Provide an intuitive explanation of the parallel trends assumption. In your answer include a graph that illustrates the assumption



- (b) [6 points] Show that if the parallel trends assumption does not hold, an analyst cannot recover an unbiased estimate of the causal effect of a treatment in a difference in differences setting.

HINT: Start by constructing a table similar to that below with mean values of an outcome variable in each cell.

	Before	After
Treatment		
Control		

Question 2 (10 points). Online Word of Mouth and New Product Demand.

- (a) [3 points] List three mechanisms through which online word of mouth can influence new product demand.

(b) [7 points] Explain which mechanisms that connect word of mouth to consumer demand are important for movies with either small or medium size expected box office.

HINT: Highlight why any differences you identify might occur.

Question 3 (10 points). Why Users Contribute to Social Media.

(a) [4 points] Define the terms Intrinsic Utility and Image Related Utility

(b) [6 points] Which of the motivations you defined in (a) influence the decisions of social media users? Are these findings the same for all users?

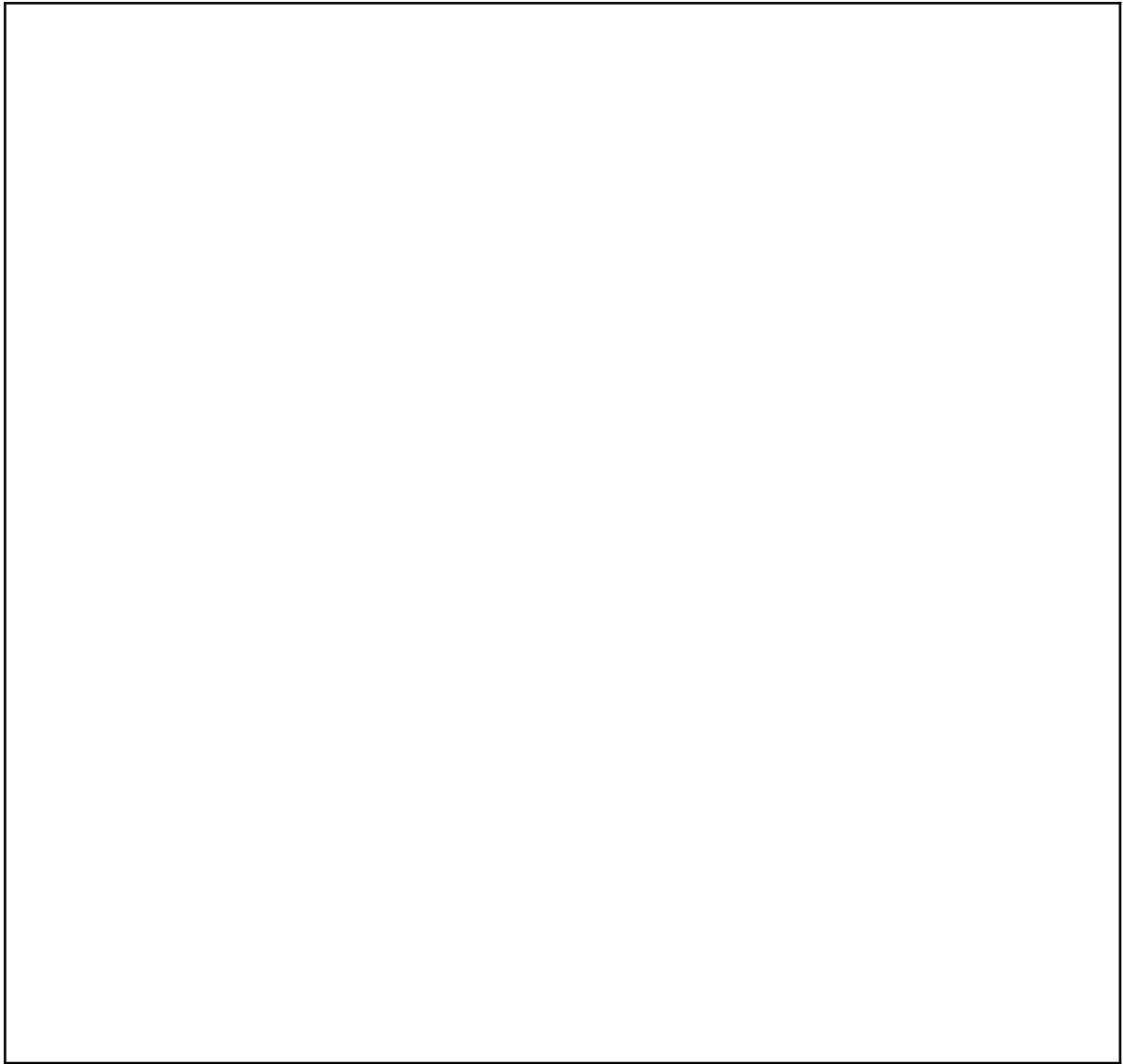
Question 4 (10 points). Topic Modelling.

- (a) [4 points]. Assume that you have a data set of Trip Advisor reviews about hotels that has already been cleaned and converted to a Document-Term-Matrix called `reviews_dtm`.

Write R code that estimates a structural topic model with 7 topics that would output the same output if run more than once

HINT: Assume all required R libraries are loaded.

- (b) [6 points] Explain how a marketing analytics team that works for Hilton Hotels could use sentiment analysis together with topic modelling to propose new managerial strategies that improve the consumer experience.



(Exam continues on the next page.)

Part C: Structured Essay Question (40 points)

Answer **all** questions.

Scenario:

In your position as marketing scientist at Meta, you are tasked to work with the World Bank on a project in India that examines whether Facebook Ads that discuss malaria prevention techniques change rural household behaviour. The World Bank is interested in whether the ad campaigns are effective on average, and whether the effects differ between relatively poorer households compared to relatively richer ones.

The Ads will be targeted to reach individuals with a Facebook account in a random subset of districts, and you will have access to surveys administered before and after the ad campaign from households across all rural districts in India with questions regarding their use of malaria prevention aids, and characteristics about their household demographics and income.

Question 1 (5 points).

Identify the marketing problem, and why it is relevant.

Question 2 (5 points).

State the main research question(s) that your team needs to answer.

Question 3 (5 points).

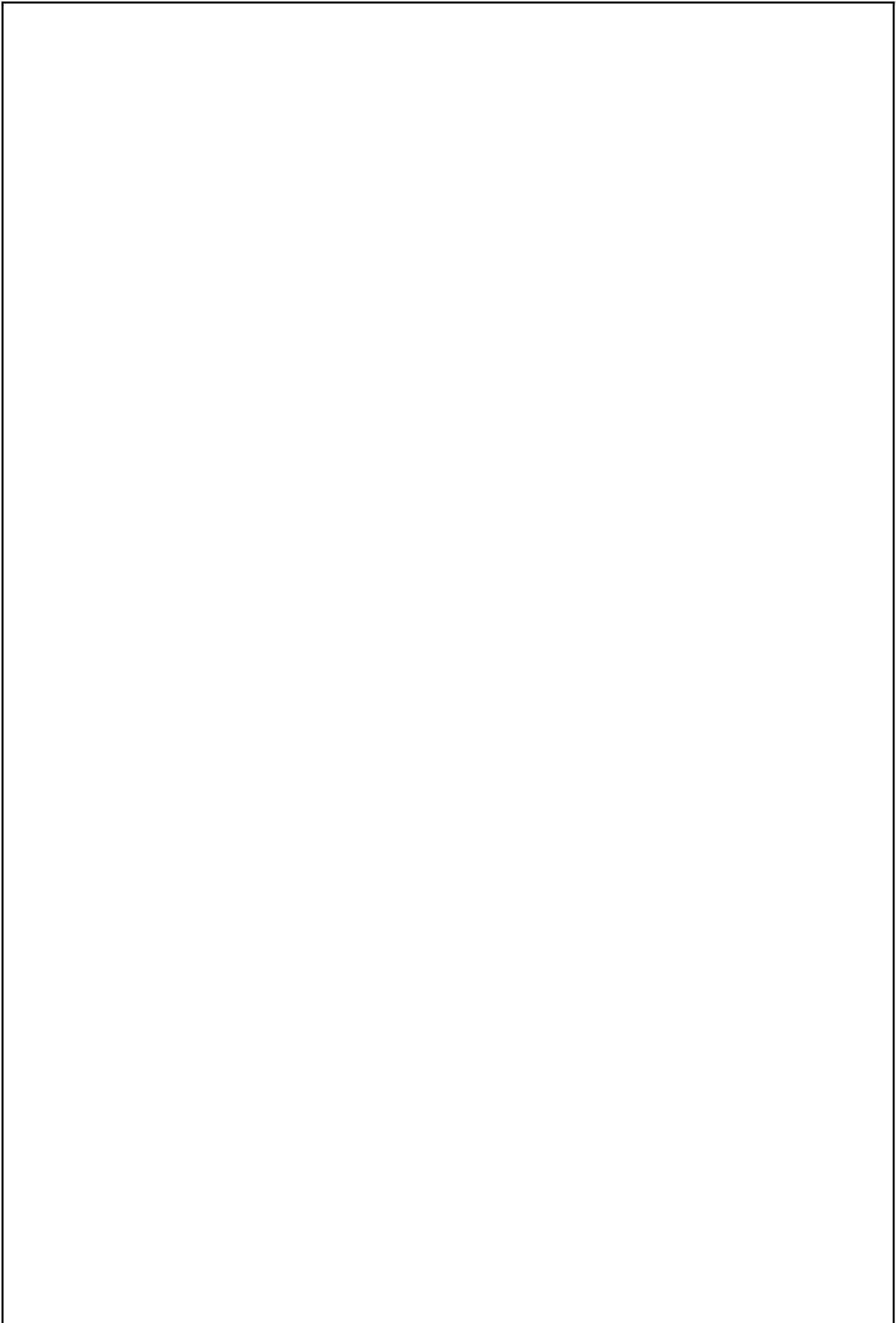
What are the hypotheses being tested? Write them using simple, yet precise terminology.

Question 4 (5 points).

What data would you need to complete the task? Specify the variables you would use and whether you are using cross section, time series or panel data to answer your question. Justify your answer.

Question 5 (10 points).

Outline the empirical framework you will use, specifying any regression equations or other statistical techniques you will need to answer the question. Also explicitly state any assumptions that need to hold for your analysis to deliver causal estimates.



Question 6 (3 points).

How does the empirical modelling you propose deliver answers to the hypotheses outlined in (3) and the research question in (2)?



Question 7 (7 points).

What results would you expect to find? Link these back to explicit managerial strategies that the World Bank could pursue in the future to promote public health awareness campaigns.



