Ultimately, we use an Instrumental Variable (IV) approach to estimate the causal impact of water insecurity on mental health, but use other models to rationalise this approach. Given our research question, our key causal model (Model 1), is the following:

Mental Health Disorderi = 0+1Basic Water Accessi+i

However, given the significant number of physical and social factors which impact mental health ([INSERT REFERENCE ON MENTAL HEALTH BEING COMPLEX) this model clearly suffers from Omitted Variable Bias (OVB). To account for this, we propose to use four control variables when testing our causal relationship: being female, age part of the religious minority (non-Christian) and living in a rural area (Model 2). We think all of these personal attributes will independently impact the likelihood of having a mental health disorder.

Given our limitation of using just four control variables (as set out in the Capstone Project Proposal), we don’t think we can reasonably address any OVB which might be remaining in this model. It’s possible X, X or X could be remaining in the error term of Model 2, but we consider these as having a second order impact on endogeneity compared to the ones we’ve chosen.

Mental Health Disorderi = 0+1Basic Water Accessi+2Female+3Agei+4Religious Minorityi+

                   5Rurali+i

Mental healthi = 0+1WaterAccessi+2Sexi+3Agei+4Religious Minorityi+5Rurali+i

Although we consider Model 2 reasonably deals with OVB, we think using an IV approach is required to sufficiently purge endogeneity from the error term of our causal relationship. This is because it’s likely reverse causality between mental health and water access is biassing the estimate of our causal parameter, as those with chronic mental health disorders may be less capable of maintaining consistent access to water access. Given we expect water access improves mental health outcomes, it's likely our causal parameter suffers from attenuation bias as individuals with worse mental health are likely to have worse access to water.

To operationalise our IV methodology, we propose to use an individual’s distance from their drinking water source as our instrument because we consider it satisfies the relevance and exogeneity requirements. We consider it relevant as the distance to drinking water access is likely strongly correlated with whether an individual has basic water access. We also consider it satisfies the exogeneity as we think it's reasonable to assume this instrument does not directly affect mental health outcomes, except through its effect on basic water access. By using IV we can directly address this reverse causality, as it provides a means of estimating the effect of changes in water access which are unrelated to the individual’s mental health status.

This approach is given by Model 3 below, which is a just-identified two equation system:

First stage: WaterAccessi= 0+1Distance From Drinking Water+2Sex+3Sex+4Age+

             5Religious Minority+6Rural+i

econd stage: Mental healthi = 0+1WaterAccessi+2Sexi+3Agei+4Religious Minorityi+5Rurali+i

Although we consider Model 3 sufficiently accounts for the reverse causality producing endogeneity in Model 2, a complaint of this model is that our instrument could be seen as introducing new sources of OVB into the error term. If this were the case, we think variables reflecting the difficulty of travel to water, vehicle ownership and climatic conditions are the most likely sources.

On the difficulty of travel to water, we consider our rural control variable reasonably captures this as there is not sufficient variation in the topography of Ghana within the average distance individuals walk for water (15 minutes from Table 1). For the other two possible source of OVB, these issues are unresolvable for this report given time and data constraints for this report as these variables are unobserved in this dataset. Therefore, we concede vehicle ownership and climate conditions could possibly bias our results, but don’t consider that is sufficient reason to abandon our approach.

Jocelyn:

In an attempt to account for terrain and climate, we aligned topographic maps with the top 3 cities in Ghana: Accra, Kumasi, and Tamale (Ghana Cities by Population , 2024). By analysing their locations we notice that these cities are all at relatively low elevations which favour practical terrain and milder climates especially in coastal areas such as Accra. Therefore, we assume that our rural control variable reasonably captures this variation. Unfortunately, for vehicle ownership, these issues are unresolvable due to the unobserved data and time constraints. With this, we acknowledge that vehicle ownership and even other omitted variables could bias our results, but don’t consider that is sufficient reason to abandon our approach.