Pre-Analysis Plan

This document provides a preliminary outline of econometric models that will be used to estimate the average treatment effect of the provision of social comparison information on groundwater use decisions.

1 Blocking Variables

We utilized several blocking variables to guide the assignment of treatment and control among wells and their owners in the study area. These blocks were determined by using groundwater use data from previous years (2017 and 2018) to determine which factors best predict water use. This process resulted in block along:

1. $GWMD$: A vector of groundwater management district (GWMD) dummy variables which constitute the comparison group in the experiment.

2. $Q2_{i,2018}^{owner}$: A dummy variable which = 1 if the i^{th} well's owner's mean 2018 water use, across all their wells, is greater than their GWMD's median 2018 water use at the well-level.

2 Primary Econometric Model

Let w_i represent the water use by the i^{th} well in 2019. The primary econometric model takes the following form

$$w_i = \alpha * GWMD + \beta_1 * Q2_{i,2018}^{owner} + \beta_2 * Treated_i + \varepsilon_i \tag{1}$$

where α is a vector estimated coefficients related to GWMD dummy variables, β_1 is a coefficient capturing the effect of the dummy variable $Q2_{i,2018}^{owner}$, $Treated_i$ is a dummy variable that indicates whether the owner of the i^{th} well received comparison information, β_2 is the treatment effect of interest, and ε_i is an idiosyncratic error term. Standard errors are clustered at the owner level.

3 Secondary Econometric Models

3.1 Model 1

As in the primary model, let w_i represent the water use by the i^{th} well in 2019. The model takes the following form

$$\begin{aligned} w_i =& \alpha * GWMD + \beta_1 * Q2_{i,2018}^{owner} + \beta_2 * Treated_i + \beta_3 * (Treated_i \ X \ Q2_{i,2018}^{owner}) \\ &+ \beta_4 * Well\ Capacity_i + \beta_5 * \#\ of\ Wells + \varepsilon_i \end{aligned}$$

where β_3 measures the treatment effect conditional on the i^{th} well's owner's mean 2018 water use, across all their wells, being greater than their GWMD's median 2018 water use at the well-level. β_4 measures the impact of well capacity, which constitutes a flow constraint imposed by local aquifer conditions which limits the volume of water that can be pumped within a given unit of time (e.g. $\frac{gal.}{min}$). β_5 measures the impact of the total number of wells owned by the owner of the i^{th} well.

3.2 Model 2

As in the primary model, let w_i represent the water use by the i^{th} well in 2019. The model takes the following form

$$\begin{aligned} w_i =& \alpha * GWMD + \beta_1 * Q2_{i,2018}^{owner} + \beta_2 * Treated_i + \beta_3 * (Treated_i \ X \ Q2_{i,2018}^{owner}) \\ &+ \beta_4 * Well\ Capacity_i + \beta_5 * \#\ of\ Wells + \beta_6 * \gamma_i + \beta_7 * (\gamma_i \ X \ Treated_i) + \varepsilon_i \end{aligned}$$

where γ_i is a dummy variable which = 1 when the i^{th} well's 2018 water uses was greater than their comparison group's median well-level water use in 2018 and β_6 is a coefficient measuring the dummy variable's effect. β_7 measures the effect of higher than comparison group median 2018 water use conditional on inclusion in the treatment group.

4 Data Inclusion Rules

Outliers: Our research focuses on groundwater use among irrigators of row crops. As such we will drop any groundwater pumping observations that use less than 5 acre feet of water in 2019 as these wells were likely used for livestock watering rather than row crop irrigation.

Attrition: Any wells whose owner was involved in the social comparison mailer but does not report pumping in 2019 while reporting in 2018 will be dropped from the analysis.

Noncompliance: All wells associated with a mailer that was returned due to incorrect mailing address provided by the State of Colorado will be dropped from the analysis.