There are five stratas in the design, indexed with $j = 1, \ldots, 5$. The number of treated and control units in stratum j equals N_j^T and N_j^C for the two groups. In the design, we have

$$N_1^T = 568, N_2^T = 248, N_3^T = 464, N_4^T = 377, N_5^T = 343, \tag{1}$$
$$N_1^C = 143, N_2^C = 10, N_3^C = 3, N_4^C = 857, N_5^C = 634. \tag{2}$$

The total number of units in the experiment is $N = \sum_{j=1}^{5} N_j^T + N_j^C = 3,647$. Within each stratum, treatment assignment is completely randomized. The primary outcomes are binary variables (APPLY, GRANTED). For simplicity, we therefore focus here on a single binary outcome Y.

Let p_j^T (p_j^C) be the share of treated (control) in stratum j for which $Y = 1$, the natural estimator of the average treatment effect in that stratum is the difference-in-means estimator $\hat{\tau}_j = p_j^T - p_j^C$. The estimator of the overall treatment effect is a weighted average of each treatment effect

$$\hat{\tau} = \sum_{i=1}^{N} \lambda_j \hat{\tau}_j, \tag{3}$$

where $0 \leq \lambda_j \leq 1$ and $\sum_{j=1}^{5} \lambda_j = 1$. A natural choice for λ_j is proportional to the inverse of the variances:

$$\lambda_j = \frac{\frac{N_j^T N_j^C}{N_j^T + N_j^C}}{\sum_{j=1}^{5} \frac{N_j^T N_j^C}{N_j^T + N_j^C}}. \tag{4}$$

This estimator is equivalent to the coefficient in front of a treatment indicator from a regression of Y on a treatment indicator and dummy variables for the different strata. The variance can be estimated as

$$\widehat{\mathrm{Var}(\hat{\tau})} = \lambda_j^2 \left(\frac{p_j^T(1 - p_j^T)}{N_j^T - 1} + \frac{p_j^C(1 - p_j^C)}{N_j^C - 1} \right). \tag{5}$$

To test whether a treatment effect exists, we will test the sharp null with Fisher's exact test using the estimator in equation (3) with the weights in equation (4). To construct 95 percent confidence interval, we will also use the variance estimator in equation (5).

For the outcome variable APPLY, previous experiments have suggested that the application rate after six months is around ten percent. Using this figure, we can simulate the power of Fisher's exact test for different application rates in the treatment group. For simplicity, we assume that the ex-ante probability of applying is the same for the different strata.

Figure 1 shows the power for different values of the treatment effect (i.e., $\tau = E(p^T) - E(p^C)$). As can be seen in the figure, a power of 80 percent can be reached with a treatment effect of $\tau \approx .135 - .1 = .035$, i.e., 3.5 percentage points, which is substantially less than what was found in the previous experiments (where the estimate was around 10 percentage points). While the intervention is somewhat more restricted this time, it still seems as if the study is well-powered to discover differences in the application rates.

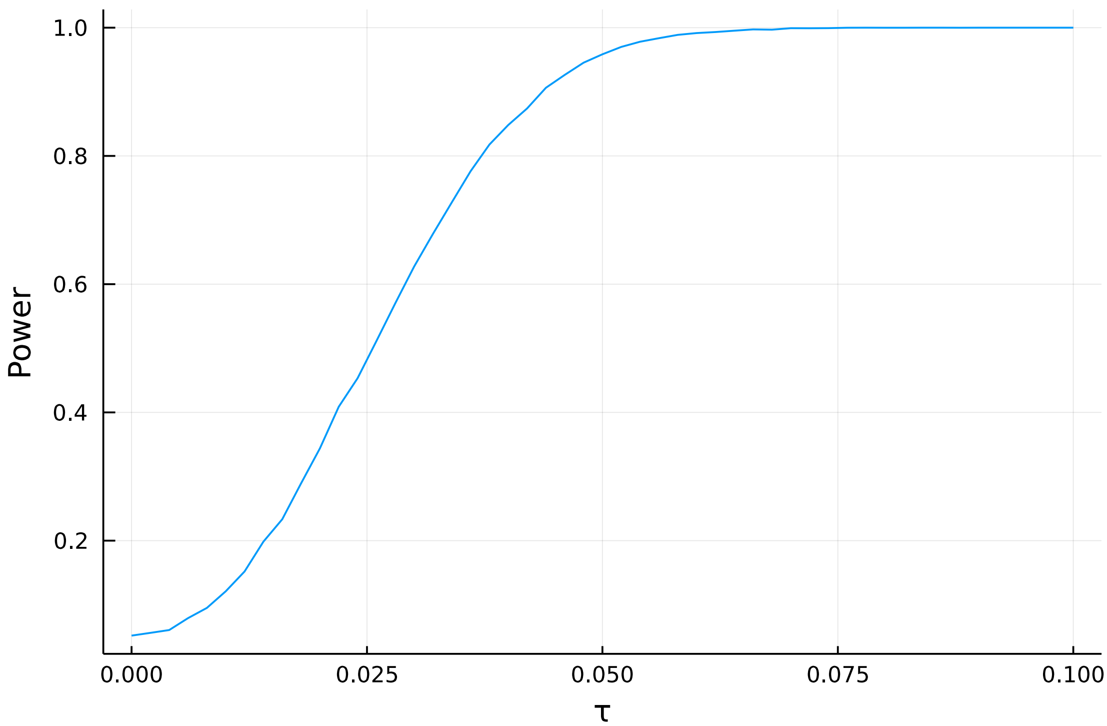

Figure 1: Power as a function of treatment effect