Homework - Week 8

Victor Jiao – bjiao@uchicago.edu

December 7, 2024

It was until the later parts of Problem 2 when I realized I should have done this assignment on a Jupyter notebook file. Unfortunately, it's a hassle to migrate everything there, so I am submitting my Jupyter Notebook alongside this PDF file.

A lot of the calculations are done through code, especially in the last two problems. In this file, I will provide the formulae of the calculations, as well as the results; by theory, you wouldn't have to refer to my code to understand the formulae, but the data is in the code, so you'd have to run my code to see how the values are calculated. I sincerely apologize for the inconvenience. As a redemption, my code is designed such that you can run everything at once and all results will be displayed.

1 Normal/Normal Shrinkage

1.1

Using the expected value formula provided, we get:

$$\mathrm{E}[\theta_j|\hat{\theta}_j, s_j] = \mathrm{E}[\theta_j] + (\hat{\theta}_j - \mathrm{E}[\hat{\theta}_j]) * \frac{\mathrm{Cov}(\theta_j, \hat{\theta}_j)}{\mathrm{Var}(\hat{\theta}_j)}$$

We already know that:

$$\mathrm{E}[\hat{\theta}_j|\theta_j] = \theta_j, \mathrm{Var}[\hat{\theta}_j|\theta_j] = s_j^2$$

because $\hat{\theta}_j|\theta_j, s_j \sim \mathcal{N}(\theta_j, s_j^2)$, and

$$\mathrm{E}[\theta_j] = \mu_\theta, \mathrm{Var}(\theta_j) = \sigma_\theta^2,$$

as provided.

As a result,

$$\mathrm{Cov}(\theta_j, \hat{\theta}_j) = \mathrm{Var}(\theta_j) = \sigma_\theta^2,$$

because $\mathrm{E}[\hat{\theta}_j|\theta_j] = \theta_j$.

Using the law of total variance,

$$\mathrm{Var}(\hat{\theta}_j) = \mathrm{Var}(\theta_j) + \mathrm{Var}(\hat{\theta}_j|\theta_j) = \sigma_\theta^2 + s_j^2.$$

Plugging everything in, we get:

$$\mathrm{E}[\theta_j|\hat{\theta}_j, s_j] = \mu_\theta + (\hat{\sigma}_j - \mu_\theta) * \tfrac{\sigma_\theta^2}{\sigma_\theta^2+s_j^2} = \mu_\theta + \hat{\theta}_j * \tfrac{\sigma_\theta^2}{\sigma_\theta^2+s_j^2} - \mu_\theta * \tfrac{\sigma_\theta^2}{\sigma_\theta^2+s_j^2} = \hat{\theta}_j * \tfrac{\sigma_\theta^2}{\sigma_\theta^2+s_j^2} + \mu_\theta * \tfrac{s_j^2}{\sigma_\theta^2+s_j^2}.$$

This is a normal-normal model, where $\theta^* \equiv \mathrm{E}[\theta_j|\hat{\theta}_j, s_j]$ denotes the posterior mean for θ_j given $\hat{\theta}_j$. This is what we call a "linear shrinkage" formula. Depending on the mean and standard deviation of the sample, θ^* may shrink noisy sample estimate toward prior mean based on signal-to-noise ratio. One can view this as a means of fitting metrics observed from sample into true population metrics. This idea of fitting coincides with fitted value from OLS regression of θ_j on $\hat{\theta}_j$ and inherits the usual OLS "best linear predictor" prediction.

1.2

The conditional MSE in terms of θ_j, s_j, μ_θ, and σ_θ^2 for $\hat{\theta}$ and θ^* is, respectively:

$$\mathrm{MSE}^C_{\hat{\theta}}(\theta_j, s_j) = \mathrm{E}[(\hat{\theta}_j - \theta_j)^2|\theta_j, s_j] = s_j^2, \text{ and}$$

$$\mathrm{MSE}^C_{\theta^*}(\theta_j, s_j) = \mathrm{E}[(\theta_j^* - \theta_j)^2|\theta_j, s_j] = (\tfrac{\sigma_\theta^2}{\sigma_\theta^2+s_j^2})^2 * s_j^2 + (\tfrac{s_j^2}{\sigma_\theta^2+s_j^2}) * (\theta_j - \mu_\theta)^2.$$

In the context of the in-class example, EB posterior θ_j^* shrinks the estimate for school j using hyperparameters estimated with the larger pool of students.

Mathematically, when $s_j^2 = 0$, $\mathrm{MSE}^C_{\hat{\theta}}(\theta_j, s_j) = \mathrm{MSE}^C_{\theta^*}(\theta_j, s_j)$ because the coefficient $\frac{\sigma_\theta^2}{\sigma_\theta^2+s_j^2} = 1$ and coefficient $\frac{s_j^2}{\sigma_\theta^2+s_j^2} = 0$. However, this is an edge case that would only occur when all observations in the sample are identical. (In the context of the in-class example, this edge case can be safely ignored. No schools are identical from each other.)

So long as $s_j^2 > 0$, linear shrinkage takes effect: θ_j^* is estimated with a fraction of the observed parameters and a fraction of the hyperparameters observed in a larger population.

We can see that $\text{MSE}^C_{\hat{\theta}}(\theta_j, s_j) < \text{MSE}^C_{\theta^*}(\theta_j, s_j)$ when $s_j^2 > (\theta_j - \mu_\theta)^2$. In other words, when the variance of the sample (think of a school) exceeds that of a higher-level sample (think of the overall distribution of the schools in a district), θ_j^* will shrink to fit the latter and be smaller in value than $\hat{\theta}_j$. Such shrinkage reduces the variance of the observed sample in order to fit the higher-level sample better. It works well when the observed sample is not a significant outlier from the average, but may introduce significant bias otherwise.

The edge case where $s_j^2 = (\theta_j - \mu_\theta)^2$, the shrinkage formula wouldn't do anything. But that would also mean that the variance observed from the sample equals that of the higher-level sample, and no fitting would be necessary.

1.3

The unconditional MSE in terms of θ_j, s_j, σ_θ^2 for $\hat{\theta}$ and θ^* is, respectively,

$$\text{MSE}^U_{\hat{\theta}}(s_j) = \int \text{E}[(\hat{\theta}_j - \theta)^2 | \theta_j = \theta, s_j] dG(\theta) = s_j^2, \text{ and}$$
$$\text{MSE}^U_{\theta^*}(s_j) = \int \text{E}[(\theta^* - \theta)^2 | \theta_j = \theta, s_j] dG(\theta) = \tfrac{\sigma_\theta^2}{\sigma_\theta^2 + s_j^2} * s_j^2.$$

The relevant notion of MSE integrates over mixing distribution G. In other words, as the question states, unconditional MSE is derived when the expectation treats θ_j as random. In the context of the in-class example, we derive unconditional MSE when we are equally interested in all schools.

We can see that so long as $s_j^2 > 0$, $\text{MSE}^U_{\theta^*}(s_j) < \text{MSE}^U_{\hat{\theta}}(s_j)$. In other words, the observed variance gets discounted by a factor that weighs in a higher-level sample variance.

1.4

School district administrators would look at all schools holistically. In other words, they demonstrate equal interest on all schools. Hence, they would prefer using the unconditional MSE.

Parents would look at an individual school. Hence, they would prefer using the MSE conditioned upon an individual school. They would also want to look at the difference between $\text{MSE}^C_{\hat{\theta}}(\theta_j, s_j)$ and $\text{MSE}^C_{\sigma^*}(\theta_j, s_j)$ to see how different the school in their neighborhood is from the district average. (It could be significantly better or significantly worse, either way.)

2 Estimates of Labor Market Discrimination

2.1

I have verified that the sample dimensions match the distribution in the question. There are 78910 applications, 97 employers, and 10453 job positions. There are 4 columns: `firm_name`, `job_id`, `callback`, and `white`.

I have also tabulated the number of applications received for each job position by race. I can confirm that most jobs have received 4 white and 4 black applications, with some exceptions. There isn't an easy way to display the data here, so please refer to my code. (I know I'm not supposed to do that, but please allow this exception. I'm also pretty sure this table is not necessary.)

2.2

We can generate a OLS regression between `callback` and `white` by running the line below:

```
sm.OLS(y, X).fit()
```

We can generate the same regression but with heteroskedasticity-robust by running:

```
sm.OLS(y, X).fit(cov_type='HC1')
```

We can generate the same regression but with job-clustered option by running:

```
sm.OLS(y, X).fit(cov_type='cluster', cov_kwds={'groups': krw_data['job_id']})
```

For brevity, you may find the three regression results at the end of the document, in section 4.1.

In all three models, the coefficient for `white` is 0.0219. This coefficient means that having a distinctly-white name is associated with a 2.19 percentage point higher likelihood of receiving a callback, ceteris paribus. This effect is statistically significant (p-value = 0.000).

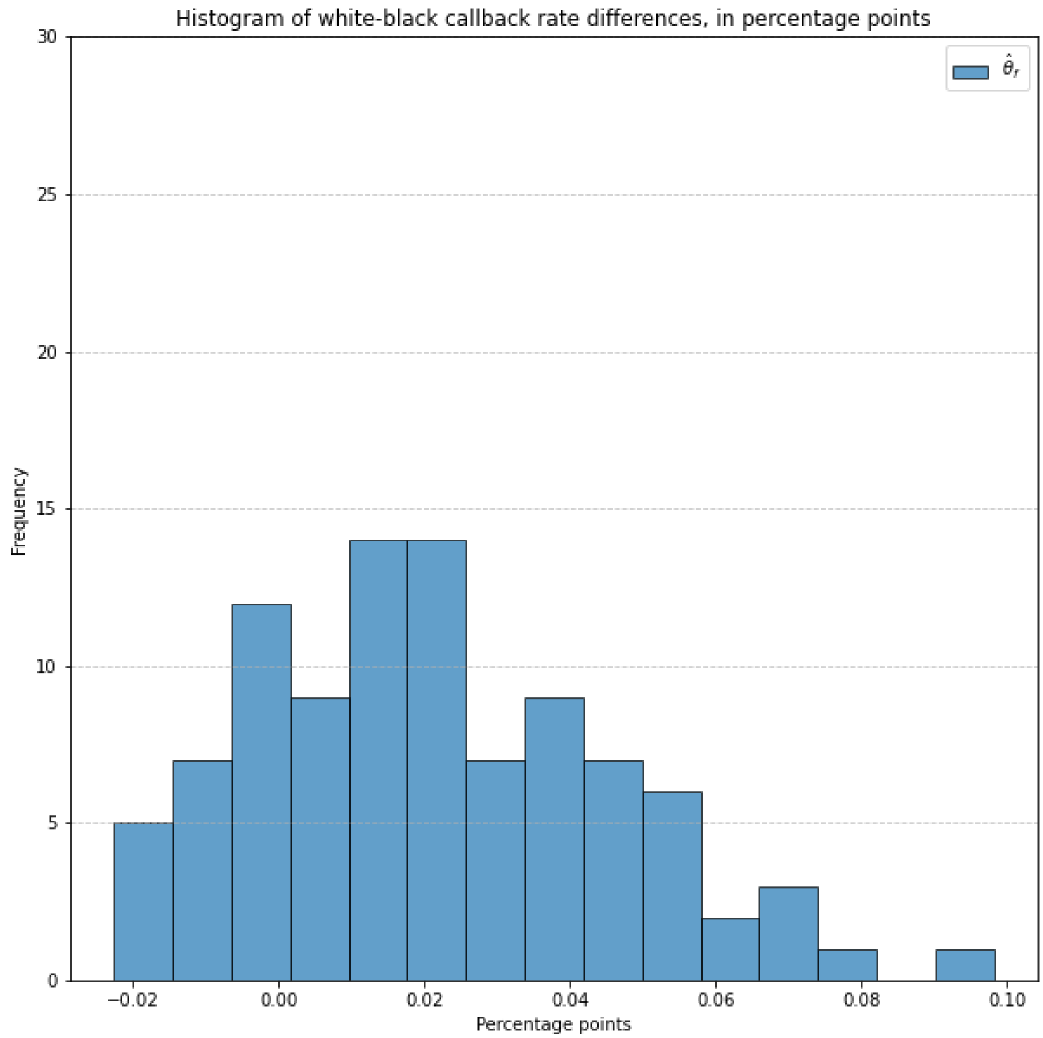
The slope standard error is 0.003 for generic model, 0.003 for heteroskedasticity-robust model, and 0.002 for job-clustered model.

Heteroskedasticity-robust model relaxes the assumption of equal variances by allowing for the possibility that some observations may have larger or smaller error variances. The similar standard error for heteroskedasticity-robust model compared to the generic model suggests that heteroskedasticity correction by itself did not drastically alter the estimated precision.

Clustering by job effectively means grouping the applications for each job, assuming that observations within the same job are related. The standard errors are then calculated by considering the variation across these job-level clusters rather than across individual observations alone. Usually, introducing clustering adds a realistic complexity and may contribute to higher slope standard error, but that's not the case here. A likely reason is that every job in this experiment consistently shows a similar difference in callback rates between distinctly white and distinctly black names, and clustering by job could reveal a really stable pattern.

2.3

Below is the histogram of the firm-level white-black callback rate differences:



I have set the y-axis to 30 to ensure scale uniformity with subsequent histograms.

The variance of the estimated white-black callback gaps across firms is expected to be larger than that of the true underlying gaps due to the addition of estimation error variance. This occurs likely because each estimate incorporates not only the true firm-specific effect but also random sampling fluctuations in the estimation process.

2.4

We use $\hat{\mu}_\theta$ as the estimator for μ_θ, and s^2_θ as the estimator for σ^2_θ. I have included four significant figures for each value.

We can calculate $\hat{\mu}_\theta$ by averaging all estimated contact gaps $(\hat{\theta}_f)$ across all firms:

$$\hat{\mu}_\theta = \frac{\sum_{f=1}^{N} \hat{\theta}_f}{N} = 0.02111.$$

We first compute the observed variance of $\hat{\theta}_f$:

$$s^2_{\hat{\theta}} = \frac{\sum_{f=1}^{N} (\hat{\theta}_f - \hat{\mu}_\theta)^2}{N-1}.$$

We then take the mean of the square standard errors to estimate the variance due to noise:

$$\bar{s}^2_f = \frac{\sum_{f=1}^{N} s^2_f}{N}.$$

The observed variance of the estimates includes both the true variance of the underlying gaps and the variance due to estimation error:

$$s^2_{\hat{\theta}} = \hat{\sigma}^2_\theta + \bar{s}^2_f \Rightarrow \hat{\sigma}^2_\theta = s^2_{\hat{\theta}} - \bar{s}^2_f.$$

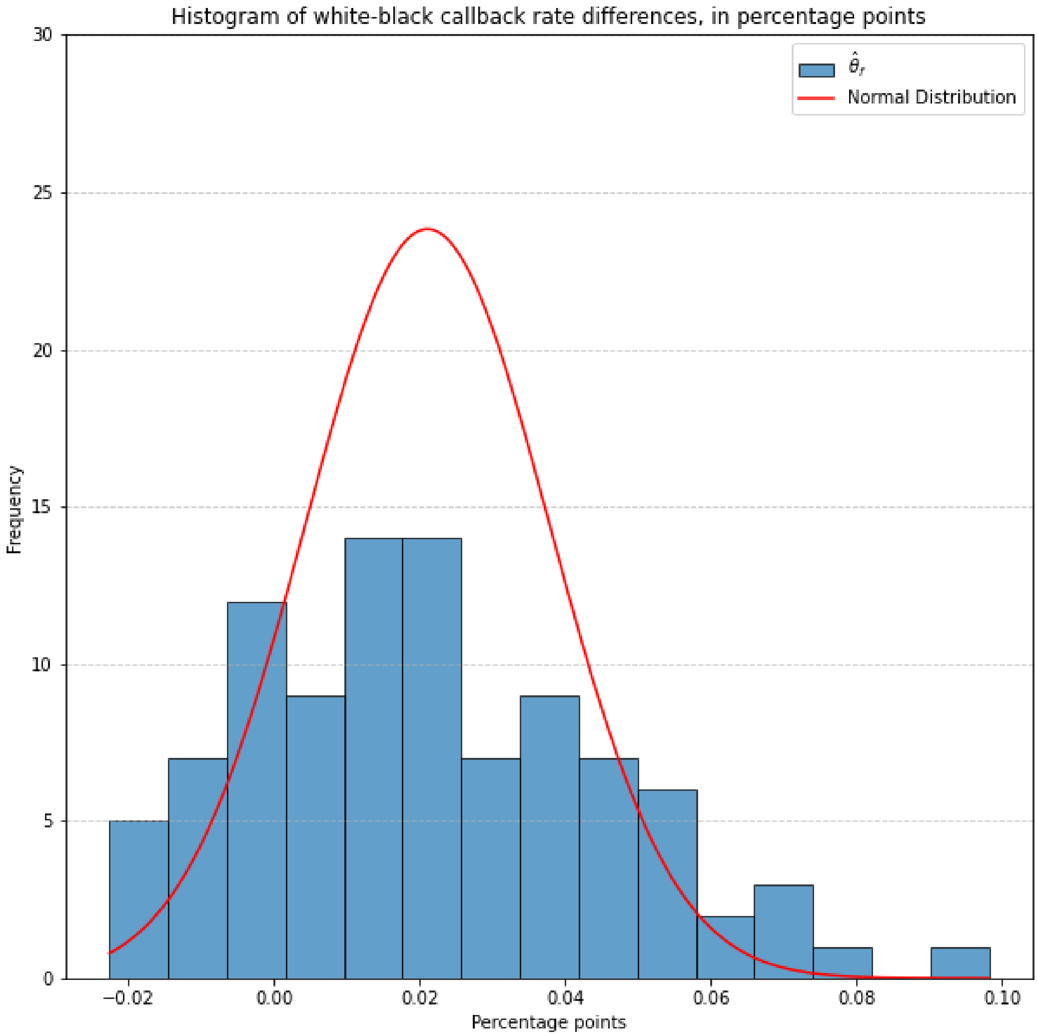
We find, from calculations in code, that:

$$\bar{s}^2_f = 0.0003101,\ s^2_{\hat{\theta}} = 0.0005905,\ \hat{\sigma}^2_\theta = 0.0002804.$$

As a result, the proportion of estimation due to noise is:

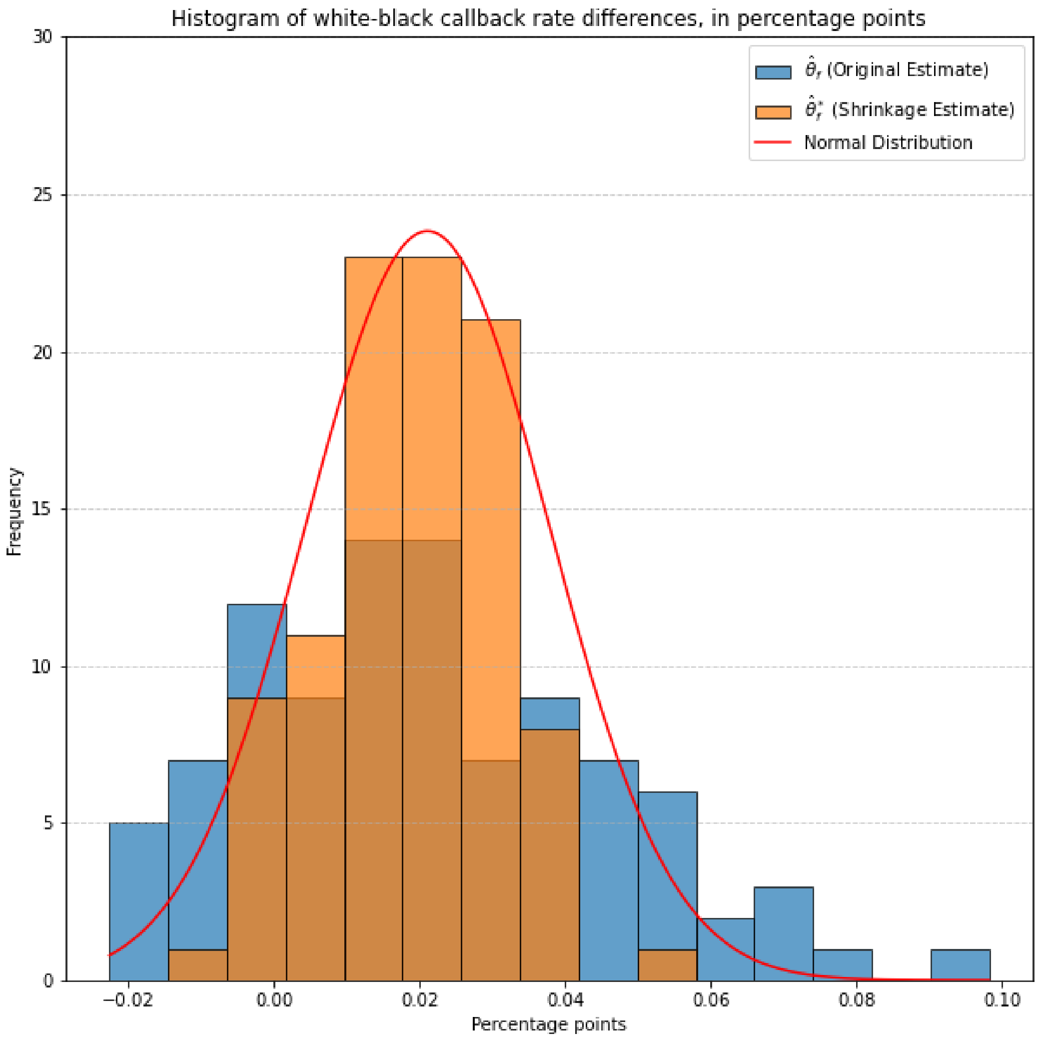
$$\frac{\bar{s}^2_f}{s^2_{\hat{\theta}}} * 100\% = 52.51\%.$$

Below is the updated histogram overlaid with a normal distribution with $\mu_\theta = \hat{\mu}_\theta$, and $\sigma_\theta = \hat{\sigma}_\theta$:



2.5

Below is the updated histogram overlaid with the distribution of $\hat{\theta}_f^*$:



The variance of $\hat{\theta}_f^*$ across firms is smaller than $\hat{\sigma}_\theta^2$. This is because because the shrinkage mechanism reduces the influence of noisy, extreme values by pulling them toward the firm-clustered standard deviation. The linear shrinkage formula reflects a trade-off between variability and reliability inherent in shrinkage estimators.

2.6

We first consider the MSE of $\hat{\theta}_f$. We know that conditional upon θ_f, we have:

$$\hat{\theta}_f|\theta_f \sim \mathcal{N}(\theta_f, s_f^2) \Rightarrow \mathrm{E}[(\hat{\theta}_f - \theta_f)^2|\theta_f] = s_f^2.$$

However, since this conditional expectation does not need to depend on θ_f, taking unconditional expectation yields:

$$\mathrm{MSE}(\hat{\theta}_f) = \mathrm{E}[(\hat{\theta}_f - \theta_f)^2] = s_f^2.$$

We then consider the MSE of $\hat{\theta}_f^*$. For the brevity of formulae, we define shrinkage weight $w = \frac{\hat{\sigma}_\theta^2}{\hat{\sigma}_\theta^2 + s_f^2}$. Hence,

$$\hat{\theta}_f^* = w_f\hat{\theta}_f + (1 - w_f)\hat{\mu}_\theta.$$

We first subtract θ_f from the linear shrinkage estimates:

$$\hat{\theta}_f^* - \hat{\theta}_f = w_f(\hat{\theta}_f - \theta_f) + (1 - w_f)(\hat{\mu}_\theta - \theta_f).$$

We see that

- $(\hat{\theta}_f - \theta_f)|\theta_f \sim \mathcal{N}(0, s_f^2)$, and
- $\theta_f \sim \mathcal{N}(\mu_\theta, \sigma_\theta^2)$, so $(\mu_\theta - \theta_f) \sim \mathcal{N}(0, \sigma_\theta^2)$.

We can see that $\hat{\theta}_f^* - \theta_f$ is a linear combination of two independent normal random variables with zero means. Its variance is simply the sum of variances weighted by the squares of coefficients. Knowing that the variance of $\hat{\theta}_f - \theta_f$ is s_f^2, and that of $\mu_\theta - \theta_f$ is σ_θ^2,

$$\mathrm{Var}(\hat{\theta}_f^* - \theta_f) = w_f^2 * s_f^2 + (1 - w_f)^2 * \sigma_\theta^2.$$

Since the expectation of $\hat{\theta}_f^* - \theta_f$ is zero, the MSE is just the variance. We replace w with the original shrinkage weight representation and get:

$$\mathrm{MSE}(\hat{\theta}_f^*) = (\tfrac{\hat{\sigma}_\theta^2}{\hat{\sigma}_\theta^2 + s_f^2})^2 * s_f^2 + (\tfrac{s_f^2}{\hat{\sigma}_\theta^2 + s_f^2}) * \hat{\sigma}_\theta^2.$$

The average of the original MSEs across 97 firms, from computation in code, is $\frac{\sum_{f=1}^{97} \mathrm{MSE}(\hat{\theta}_f)}{97} = 0.0003101$.

Likewise, the average of the shrinkage MSEs across 97 firms is $\frac{\sum_{f=1}^{97} \mathrm{MSE}(\hat{\theta}_f^*)}{97} = 0.0001325$.

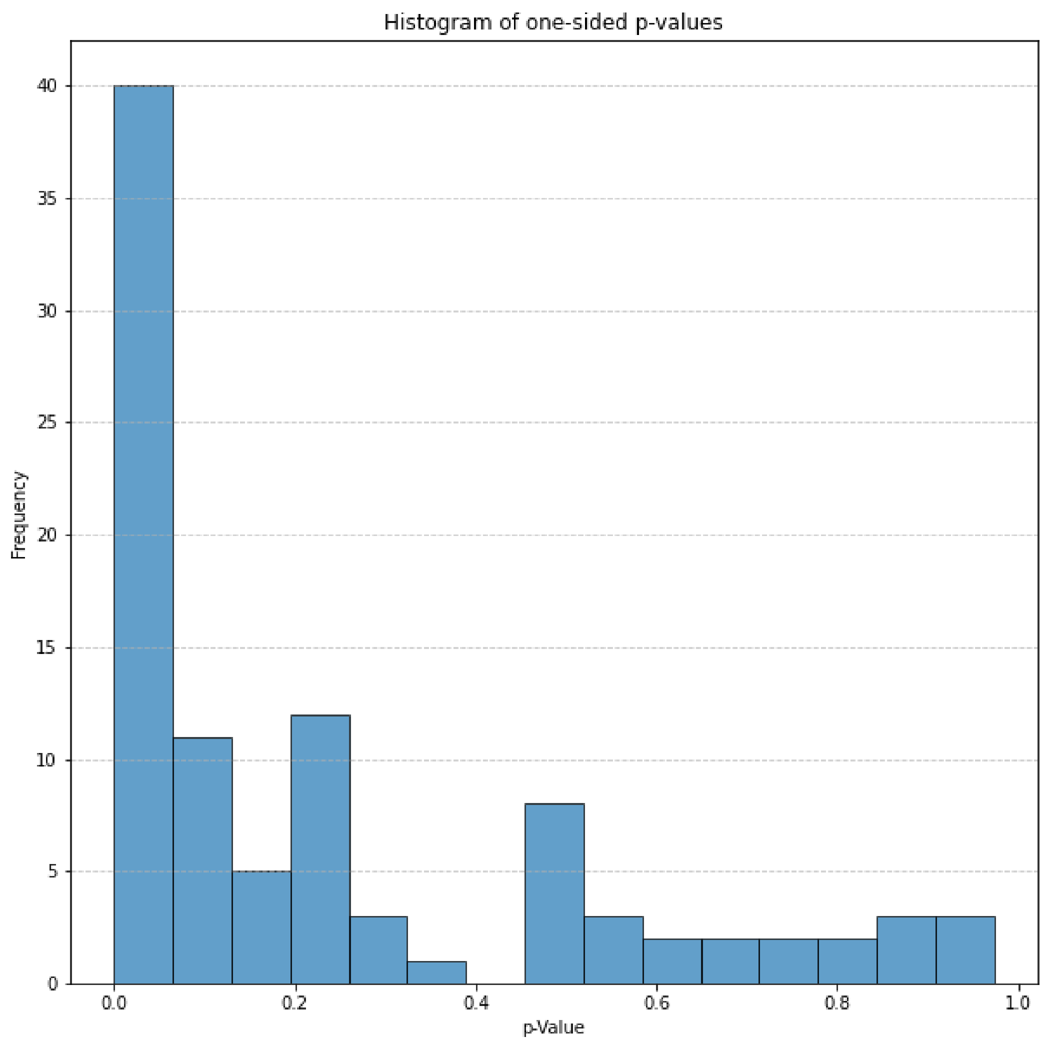
The percentage improvement rate is, therefore:

$$\sum_{f=1}^{97} \frac{\mathrm{MSE}(\hat{\theta}_f) - \mathrm{MSE}(\hat{\theta}_f^*)}{\mathrm{MSE}(\hat{\theta}_f)} * 100\% = \frac{0.0003101 - 0.0001325}{0.0003101} * 100\% = 57.28\%.$$

3 Multiple Testing to Detect Discrimination

3.1

Below is the histogram of the p-values after computing a one-tailed z-test for all 97 firms:



The histogram is very right-skewed. Most firms have very low p-values on the z-test. (Need something else here)

3.2

The FDR is the probability that a firm's true $\theta_f = 0$ given that the null hypothesis is rejected. Using Bayes' Rule:

$$\text{FDR}(\bar{p}) = \frac{\Pr[R_f(\bar{p})=1|\theta_f=0] * \Pr[\theta_f=0]}{\Pr[R_f(\bar{p})=1]}.$$

We break down the terms of this fraction below:

- $\Pr[R_f(\bar{p}) = 1|\theta_f = 0]$ denotes the probability of rejecting the null ($R_f(\bar{p}) = 1$) when $\theta_f = 0$, which is $\Pr[p_f \leq \bar{p}] = \bar{p}$.
- $\Pr[\theta_f = 0]$ denotes the proportion of firms for which the null hypothesis is true. This is exactly π_0 provided in the question.
- $\Pr[R_f(\bar{p}) = 1]$ is the total probability of rejecting the null hypothesis across all firms. In other words, it is the probability that a firm's p-value falls at or below the threshold value $\bar{p}$. Hence, $\Pr[R_f(\bar{p}) = 1] = \Pr[p_f \leq \bar{p}]$.

Plugging everything in, we get:

$$\text{FDR}(\bar{p}) = \frac{\bar{p}\pi_0}{\Pr[p_f \leq \bar{p}]}.$$

FDR measures the proportion of rejected null hypotheses that are actually false positives, providing a way to assess the reliability of the discoveries. The numerator of the formula, $\bar{p}\pi_0$, represents the expected number of false positives, where $\bar{p}$ is the probability of rejection under the null, and π_0 is the proportion of firms for which the null hypothesis is true. The denominator, $\Pr[p_f \leq \bar{p}]$, reflects the total number of rejections, including both true and false positives. By expressing FDR as the ratio of these two components, the formula captures the balance between detecting true effects and the risk of mistakenly rejecting true null hypotheses. It demonstrates a trade-off between setting a threshold with more rejections and higher FDR versus one with fewer rejections and lower FDR.

3.3

The numerator $\sum_{f=1}^{F} 1\{p_f > b\}$ counts the number of firms with p-values greater than the threshold b. These are firms for which there is little evidence to reject the null hypothesis. From the code, we get that $\sum_{f=1}^{F} 1\{p_f > b\} = 19$.

For the denominator, under the null hypothesis, p-values are uniformly distributed between 0 and 1. The probability of observing a p-value greater than b is $(1 - b)$. By multiplying this probability by the total number of firms F, we get the expected number of firms with p-values greater than b if all firms were under the null hypothesis.

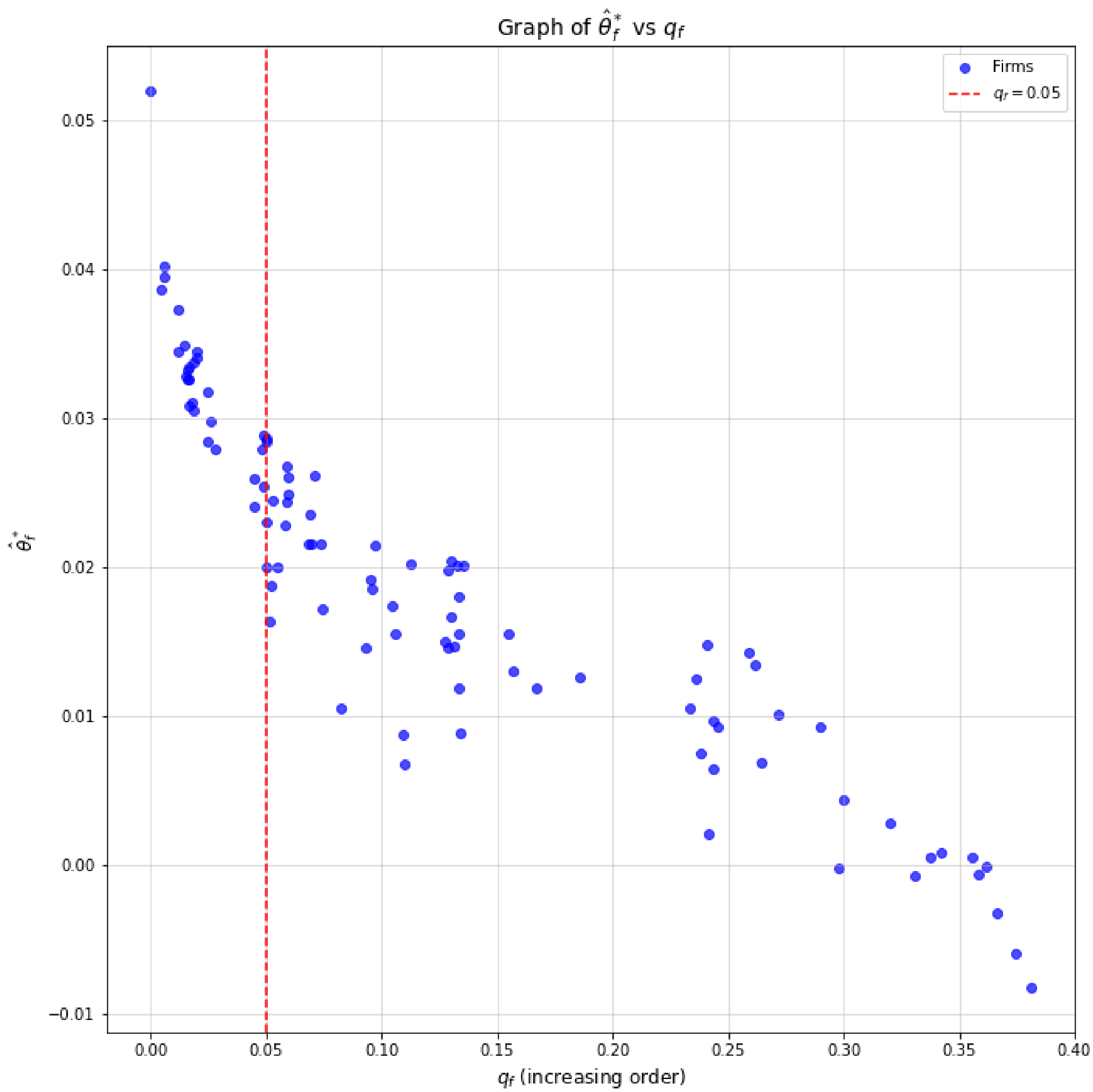
$\hat{\pi}_0$ is the ratio between the aforementioned terms, which, as the problem suggests, provides an upper bound on the proportion of firms that do not discriminate against distinctively black names:

$$\hat{\pi}_0 = \frac{\sum_{f=1}^{F} 1\{p_f > b\}}{(1-b)F} = \frac{19}{0.5*97} = 0.3918.$$

3.4

For brevity, please refer to section 4.2 for the comprehensive list of firms and their respective $\hat{\theta}_f$, s_f, p_f, q_f, and $\hat{\theta}^*_f$. There are 28 firms in total.

Please refer to the scatterplot below for the relationship between $\hat{\theta}^*_f$ and q_f, for all 97 firms, as well as the threshold q-value.



There is a clear negative correlation between q-values and shrinkage estimates. As q-value increases, the shrinkage estimates generally go down. However, note that when $0.05 \leq q_f \leq 0.15$, the negative trend isn't strong, which may testify that the threshold 0.05 can miss some firms that also show decent amounts of discrimination.

4 Supplemental Materials

4.1 OLS Regression Results

Below is the result for the generic OLS regression:

Dep. Variable:	callback	**R-squared:**	0.001
Model:	OLS	**Adj. R-squared:**	0.001
Method:	Least Squares	**F-statistic:**	50.08
Date:	Fri, 06 Dec 2024	**Prob (F-statistic):**	1.49e-12
Time:	18:25:39	**Log-Likelihood:**	-46075.
No. Observations:	78910	**AIC:**	9.215e+04
Df Residuals:	78908	**BIC:**	9.217e+04
Df Model:	1		
Covariance Type:	nonrobust		

	coef	**std err**	**t**	**P> |t|**	**[0.025**	**0.975]**
const	0.2408	0.002	110.210	0.000	0.236	0.245
white	0.0219	0.003	7.076	0.000	0.016	0.028

Omnibus:	15330.467	**Durbin-Watson:**	1.743
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	18751.134
Skew:	1.143	**Prob(JB):**	0.00
Kurtosis:	2.310	**Cond. No.**	2.62

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

Note that the generic OLS regression assumes homoskedasticity and independence among all variables. In real life, these assumptions are often violated. For instance, if the variance of `callback` depends on `white` (or other unobserved factors), the OLS standard errors may be inaccurately estimated, leading to incorrect inferences about the significance of coefficients.

Below is the result for a heteroskedasticity-robust OLS regression:

Dep. Variable:	callback	**R-squared:**	0.001
Model:	OLS	**Adj. R-squared:**	0.001
Method:	Least Squares	**F-statistic:**	50.08
Date:	Fri, 06 Dec 2024	**Prob (F-statistic):**	1.49e-12
Time:	18:25:39	**Log-Likelihood:**	-46075.
No. Observations:	78910	**AIC:**	9.215e+04
Df Residuals:	78908	**BIC:**	9.217e+04
Df Model:	1		
Covariance Type:	HC1		

	coef	**std err**	**z**	**P> |z|**	**[0.025**	**0.975]**
const	0.2408	0.002	111.835	0.000	0.237	0.245
white	0.0219	0.003	7.077	0.000	0.016	0.028

Omnibus:	15330.467	**Durbin-Watson:**	1.743
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	18751.134
Skew:	1.143	**Prob(JB):**	0.00
Kurtosis:	2.310	**Cond. No.**	2.62

Notes:
[1] Standard Errors are heteroscedasticity robust (HC1)

Below is the result for a job-clustered OLS regression. Observe that the regression clusters the covariates by `job_id`:

Dep. Variable:	callback	**R-squared:**	0.001
Model:	OLS	**Adj. R-squared:**	0.001
Method:	Least Squares	**F-statistic:**	152.4
Date:	Fri, 06 Dec 2024	**Prob (F-statistic):**	8.89e-35
Time:	18:25:39	**Log-Likelihood:**	-46075.
No. Observations:	78910	**AIC:**	9.215e+04
Df Residuals:	78908	**BIC:**	9.217e+04
Df Model:	1		
Covariance Type:	cluster		

	coef	**std err**	**z**	**P> |z|**	**[0.025**	**0.975]**
const	0.2408	0.004	65.473	0.000	0.234	0.248
white	0.0219	0.002	12.346	0.000	0.018	0.025

Omnibus:	15330.467	**Durbin-Watson:**	1.743
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	18751.134
Skew:	1.143	**Prob(JB):**	0.00
Kurtosis:	2.310	**Cond. No.**	2.62

Notes:
[1] Standard Errors are robust to cluster correlation (cluster)

4.2 List of Rejected Firms from Section 3.4

I have included the list of firms rejected under the rule from section 4.2 below, ranked by increasing q_f:

- "hat_theta_f": $\hat{\theta}_f$
- "s_f": s_f
- "p_f": p_f
- "q_f": q_f
- "hat_theta_f_star": $\hat{\theta}^*_f$

firm_name	hat_theta_f	s_f	p_f	q_f	hat_theta_f_star
Genuine Parts (Napa Auto)	0.098171	0.020480	8.193033e-07	0.000031	0.051989
AutoNation	0.052667	0.015000	2.230541e-04	0.004238	0.038620
Advance Auto Parts	0.073961	0.022226	4.378415e-04	0.005546	0.040249
O'Reilly Automotive	0.078701	0.024413	6.328117e-04	0.006012	0.039538
Rite Aid	0.046957	0.016194	1.867380e-03	0.011827	0.034468
Ascena (Ann Taylor / Loft)	0.067797	0.022934	1.557643e-03	0.011838	0.037347
Gap	0.052209	0.018767	2.701431e-03	0.014665	0.034897
Dick's	0.050565	0.020628	7.117912e-03	0.015027	0.032812
Murphy USA	0.048992	0.019932	6.986690e-03	0.015617	0.032649
Dillard's	0.045464	0.016804	3.410006e-03	0.016198	0.033246
Tractor Supply	0.039917	0.016216	6.915316e-03	0.016424	0.030817
AutoZone	0.054000	0.021560	6.129713e-03	0.016638	0.033487
TJX	0.047868	0.019305	6.577897e-03	0.016664	0.032600
Bed Bath & Beyond	0.040080	0.015981	6.070747e-03	0.017745	0.031040
Pizza Hut	0.056000	0.022166	5.762001e-03	0.018246	0.033789
VFC (North Face / Vans)	0.037540	0.014496	4.802405e-03	0.018249	0.030503
Pilot Flying J	0.060375	0.023896	5.758349e-03	0.019892	0.034044
Walgreens	0.061017	0.023519	4.738887e-03	0.020009	0.034537
Universal Health	0.054608	0.024459	1.278747e-02	0.024296	0.031802
Sherwin-Williams	0.034694	0.015463	1.242551e-02	0.024851	0.028443
Bath & Body Works	0.040404	0.018461	1.431308e-02	0.025900	0.029821
Olive Garden	0.034000	0.015907	1.628088e-02	0.028122	0.027887
J.C. Penney	0.030181	0.015810	2.813249e-02	0.044543	0.025907
Nationwide	0.026084	0.013545	2.706557e-02	0.044717	0.024118
Dollar General	0.039460	0.021649	3.417256e-02	0.048095	0.027981
Publix	0.047002	0.025607	3.321386e-02	0.048543	0.028868
Marriott	0.029838	0.016856	3.834436e-02	0.048570	0.025447
Republic Services	0.019522	0.011014	3.815523e-02	0.049997	0.020002