L'Hôpital's (Selection) Rule:

An Empirical Bayes Application to French Hospital Efficiency

Fu Zixuan

Supervised by Thierry Magnac

July 4, 2024

Abstract

There has been ongoing interest in comparing public and private hospitals. Inspired by the nascent study of Croiset and Gary-Bobo (2024), which investigates labor employment efficiency in four types of French hospitals, this article aims to identify the top-performing units in the same aspect. By borrowing strength from the compound decision framework and the non-parametric likelihood estimator (Kiefer and Wolfowitz, 1956), it derives the decision rule for selecting meritorious hospitals following the work of Gu and Koenker (2023). Selection outcomes are later compared across different scenarios. A preliminary conclusion can be drawn from the selection results. Though ranking and selection are both tempting to perform, such decisions need to be made with great caution.

Contents

1 Introduction

It is almost human nature to compare, rank, and select.And competition, be it good or bad, emerges in its wake. As invidious as ranking and selection may be, they often serve as key drivers of performance improvement. Society itself continuously constructs league tables, rewarding the meritorious while questioning or even punishing the unsatisfactory. The objects of interest for ranking vary widely, from teacher evaluations (Chetty et al., 2014) and community mobility indices (Chetty and Hendren, 2018) to assessments of firm discrimination (Kline et al., 2022).

This article extends such practices to the healthcare sector. Specifically, it examines labor efficiency across all hospitals in France. Utilizing a comprehensive database, *The Annual Statistics of Health Establishments (SAE)*, I first construct a measure of labor efficiency. Then, using estimates derived from this data, I compare public and private hospitals by selecting the top-performing units. I leverage recent developments in the Empirical Bayes method to facilitate this comparison.

I found that among the top 20% of best-performing hospitals, there are roughly six times more private units than public ones, adjusted for the number of hospitals in each category. The difference becomes even more pronounced when controlling for the expected number of incorrect selections. The conclusion is that public hospitals generally exhibit less efficiency than their private counterparts, aligning with the findings of Croiset and Gary-Bobo (2024). However, we now have a more granular perspective on the performance comparison.

The article bridges two fields of interest. The first focuses on productivity analysis, where the most popular methodologies are Data Envelopment Analysis (Charnes et al., 1978) and Stochastic Frontier Analysis (Aigner et al., 1977; Meeusen and van Den Broeck, 1977). However, I diverge from these approaches and adopt the *conditional input demand function* specification detailed in Croiset and Gary-Bobo (2024).[1] Simply put, we estimate a linear function that determines the necessary labor input to produce a given list of eight hospital outputs. My focus is specifically on the employment level of nurses, as unlike medical doctors, this category does not suffer from a shortage of labor supply.

The second area of interest is Empirical Bayes Methods. I draw upon a series of works by Jiaying Gu and Roger Koenker, primarily the following two papers: Gu and Koenker (2017) discusses the usefulness of estimating a prior distribution in baseball batting average prediction, while Gu and Koenker (2023) formally defines the selection problem as a compound decision, where the estimated prior can be helpful. Additionally, Kiefer and Wolfowitz (1956) demonstrates that non-parametric maximum likelihood estimation of the prior is feasible and consistent. The computation of NPMLE has been significantly improved by Koenker and Mizera (2014), leveraging recent developments in convex optimization (Andersen and Andersen, 2010). For my analysis, I will use the `REBayes` package (Koenker and Gu, 2017), which is based on the `MOSEK` software developed by Andersen and Andersen (2010).

In Croiset and Gary-Bobo (2024), the authors argue that public hospitals are less efficient than their private counterparts, in the sense that a public hospital would require a smaller

[1] I refer the reader to Croiset and Gary-Bobo (2024) for detailed reasons behind this choice.

workforce if it adopted the input demand function used by private hospitals—a key finding from their counterfactual analysis.

Having approximately replicated these results with an extended panel length, this paper differentiates itself by employing classical panel data methods for estimating the input demand function, specifically standard fixed-effect estimation and GMM. While incorporating individual fixed effects into the specification is straightforward, the estimation process is not without challenge. For instance, as Croiset and Gary-Bobo (2024) have noted, within-hospital variation is much smaller than between-group variation, which may be insufficient for obtaining reliable estimates. To address this issue, I extended the panel length, aiming to enhance the robustness of the estimates. Furthermore, the strict exogeneity assumption required by standard within-group estimation is debatable. A practical solution to relax this assumption is to use the first difference GMM estimator, as proposed by Arellano and Bond (1991). Another challenge arises from the high persistency of the regressors, which can lead to weak instruments. In response, I adapted and implemented the system GMM approach, as modified by Arellano and Bover (1995) and Blundell and Bond (1998), using the resulting estimates for subsequent analyses.

The advantage of the panel data estimator lies in its ability to capture the underlying heterogeneity, which opens the door to individual comparisons. The fixed effect estimates are generally noisy, rendering the ensuing decision-maker hand-wavy. The EB methods are proposed in an attempt to rectify the situation by empirically estimating the prior distribution of the fixed effects.

For example, in Gu and Koenker (2023), we are given the task of selecting the top 20% of fixed effects, denoted by θ_i. If θ_i follows a distribution G, this implies selecting those θ_i that exceed $G^{-1}(0.8)$. The decision rule for an individual i is represented by an indicator function δ_i, which determines whether i belongs to the selection set. This task aligns with the compound decision framework pioneered by **?**, especially when defining the loss function of the selection problem in a way that incorporates the outcomes of all individual decisions δ_i:

$$\delta^* = \arg\min_{\delta} \mathbb{E}_G \mathbb{E}_{\theta|\hat{\theta}} \left(L_n \right).$$

Since the true value of θ is unknown, we aim to minimize the expected compound loss L_n over the distribution of θ given the observed $\hat{\theta}$.

In addition to the capacity constraint of selecting the top 20%, Gu and Koenker (2023) further mitigates the number of false positive in the selection process. The false discovery rate (FDR) constraint is imposed to ensure that the expected number of incorrectly selected units remains below a specified threshold. The FDR is quantified as the proportion of false positives among all selected units, formally defined by the condition $\Pr(h_i = 0|\delta_i = 1) \leq \gamma$.

Being interested in the top-performing French hospitals, I define my selection problem as *Left tail selection* because the goal is to choose the bottom 20% of the hospital fixed effect θ_i. A smaller θ_i indicates that less labor input is required to produce the same amount of output, compared to hospitals with a higher θ_i.

It is worth mentioning that classical empirical Bayes method assumes a parametric form of the prior distribution G which is computationally more attractive. Yet thanks to fast convex optimization algorithms, the non-parametric maximum likelihood estimation is now both feasible and efficient. Nevertheless, we are completely free from imposing any parametric assumption. In fact, there are two *layers* of distribution. The lower hierarchy is the prior G with $\theta \sim G$ while the higher hierarchy is $\hat{\theta}|\theta \sim P_\theta$. It is when P_θ belongs to the exponential family that the Lindsay (1995) results hold. Usually in application, we need to impose assumptions or perform some transformation such that P_θ is normal. This kind of procedure is often questionable. Often times, researchers resort to asymptotics to justify the normality assumption, which may not be valid in small samples.

It is worth mentioning that the classical empirical Bayes method assumes a parametric form for the prior distribution G, which is computationally attractive. However, thanks to advancements in fast convex optimization algorithms, non-parametric maximum likelihood estimation (Kiefer and Wolfowitz, 1956) has become both feasible and efficient (Koenker and Mizera, 2014; Andersen and Andersen, 2010). Nonetheless, we are not entirely free from imposing any parametric assumptions. In fact, there are two layers of distribution: the lower hierarchy where $\theta \sim G$, and the higher hierarchy where $\hat{\theta}|\theta \sim P_\theta$. Lindsay (1995) has shown that a solution $\hat{G}$ exists and is a discrete probability measure, with no more than n mass points in the interval $[\min yi, \max yi]$. It is when P_θ belongs to the exponential family that the Lindsay (1995) results hold. Typically in applications, it is necessary to impose assumptions or perform transformations to make P_θ normal. This procedure is often questionable. Researchers frequently rely on asymptotic properties to justify the normality assumption, which may not be valid in small samples.

The remainder of the paper is organized as follows: **Section 2** briefly describes the data and outlines the reduced form estimation of the input demand function, where the number of nurses serves as the dependent variable, with nine output measures acting as regressors. It then applies classical panel data estimators to the same specification, examining and later relaxing the assumptions of strict exogeneity. In **Section 3**, I introduce the compound decision framework and describe the methodology for non-parametrically estimating G. **Section 4** specifically defines the selection problem following the framework proposed by Gu and Koenker (2023), and includes a comparison of different selection outcomes. Preliminary conclusions about the comparative performance of public versus private hospitals are drawn. **Section 5** discusses potential issues and concludes the paper.

2 Data and Estimation

2.1 Data

The data utilized in this study is from The Annual Statistics of Health Establishments (SAE), a comprehensive and mandatory administrative survey that serves as the primary source of data on all health establishments in France [2]. Our analysis primarily focuses on healthcare outputs (across ten different measures) and labor inputs (specifically registered and assistant nurses). The panel data spans nine years, from 2013 to 2022, excluding 2020

[2] La Statistique annuelle des établissements (SAE)

due to disruptions caused by the pandemic. The SAE data categorizes establishments into three types based on their legal status:

1. Public hospitals,
2. Private for-profit hospitals,
3. Private non-profit hospitals.

In alignment with the categorization used by Croiset and Gary-Bobo (2024), this study also distinguishes public teaching hospitals as a separate category, given their unique role within the French healthcare system.

Year	Teaching	Normal Public	Private For Profit	Private Non Profit	Total
2013	198	1312	1305	1382	4197
2014	201	1274	1293	1349	4117
2015	211	1275	1297	1349	4132
2016	212	1266	1297	1313	4088
2017	211	1249	1297	1306	4063
2018	214	1247	1296	1288	4045
2019	214	1236	1287	1281	4018
2021	219	1222	1293	1264	3998
2022	220	1220	1296	1259	3995

Table 1: Number of hospitals in each category, 2013-2022

As illustrated in Table 1, the distribution of hospitals—categorized as normal public, private for-profit, and private non-profit—is relatively equal and remains stable over the years. It is important to highlight the unique role of teaching hospitals. Unlike other types of hospitals, teaching hospitals allocate significant resources to doctor training and research activities. This additional commitment to educational and research missions generally results in larger institutions. This aspect becomes evident when examining the output shares of hospitals. Despite their limited number, teaching hospitals contribute a disproportionately large share of overall healthcare output, a disparity that becomes more pronounced when adjusted for the number of facilities, as detailed in Table 3.

Furthermore, the analysis reveals distinct differences in the mix of service provided by each type of hospital. Emergency care is mostly taken care of by public hospitals and private hospitals are strong in medical sessions.

2.2 Estimation

Regression without individual fixed effect First, having performed the regression separately for each type of hospital, it is without surprise that teaching hospitals have very

Output	Teaching	Normal Public	Private For Profit	Private Non Profit	Total
STAC inpatient	25.17%	43.09%	23.64%	8.1%	100%
STAC oupatient	18.4%	19.46%	52.95%	9.18%	100%
Sessions	14.49%	21.96%	34.4%	29.16%	100%
Outpatient Consultations	36.8%	52.45%	0.23%	10.52%	100%
Emergency	21.4%	60.06%	13.37%	5.17%	100%
Follow-up care and Long-term care	7.6%	19.47%	37.95%	34.98%	100%
Home hospitalization	13%	17.38%	12.4%	57.22%	100%
Psychiatry stays	6.53%	62.26%	12.93%	18.28%	100%

Table 2: Hospital share of output, 2013-2022

Output	Teaching	Normal Public	Private For Profit	Private Non Profit	Total
STAC inpatient	66.98%	19.29%	10.25%	3.48%	100%
STAC oupatient	57.91%	10.29%	27.13%	4.67%	100%
Sessions	50.12%	12.7%	20.18%	16.99%	100%
Outpatient Consultations	77.69%	18.64%	0.08%	3.59%	100%
Emergency	62.02%	29.26%	6.31%	2.41%	100%
Follow-up care and Long-term care	33.5%	14.37%	27.31%	24.82%	100%
Home hospitalization	47.83%	10.75%	7.46%	33.96%	100%
Psychiatry stays	29.65%	47.38%	9.6%	13.37%	100%

Table 3: Hospital share of output weighted by the number of hospitals, 2013-2022

Note: For example, the value a_{ij} where i is STAC inpatient and j is teaching hospitals, is calculated by $a_{ij} = \frac{\text{Number of STAC inpatient in teaching hospitals}}{\text{Share of teaching hospitals} \times \text{Total number of STAC inpatient}}$.

different coefficients, as shown in Table 4. In addition to the differences in descriptive statistics from the last section, this intrinsic difference in input demand functions or equivalently in production function is another sign that teaching hospitals may not be directly comparable to other types of hospitals. For this reason, I will exclude teaching hospitals from the subsequent analysis.

In the analysis, the relationship between the number of nurses and output levels in hospitals is quantified via a log-log regression model as follows:

$$\log(x_{it}) = \beta_0 + \beta_1 \log(y_{it}) + \varepsilon_{it} \tag{1}$$

where $\log(x_{it})$ represents the log of the number of nurses at hospital i in time t, and $\log(y_{it})$ denotes the log of a vector of output levels for the same hospital at the same time.

The regression analysis was conducted separately for each type of hospital, as detailed in Table 4. The results reveal that teaching hospitals have significantly different coefficients compared to other hospital categories. This variance is consistent with the unique characteristics of teaching hospitals, which were previously noted in their descriptive statistics and operational roles.

Specifically, the deviation in coefficients suggests fundamental differences in the input demand functions or, equivalently, in their production functions. These differences underscore the distinct operational and functional framework within which teaching hospitals operate, further evidenced by their dual focus on healthcare delivery and educational responsibilities.

For this reason, teaching hospitals will be excluded from subsequent analyses. The focus will be on the remaining three categories of hospitals: ordinary public hospitals, private for-profit hospitals, and private non-profit hospitals.

Dependent Variable:	Nurses			
	Teaching	Public	Forprofit	Nonprofit
Model:	(1)	(2)	(3)	(4)
Variables				
Constant	3.28***	1.40***	1.41***	1.00***
	(0.123)	(0.099)	(0.041)	(0.059)
STAC inpatient	0.108***	0.328***	0.261***	0.343***
	(0.016)	(0.018)	(0.007)	(0.012)
STAC outpatient	0.131***	0.078***	0.048***	0.046***
	(0.013)	(0.005)	(0.005)	(0.010)
Medical sessions	0.058***	0.049***	0.075***	0.093***
	(0.008)	(0.003)	(0.002)	(0.006)
External consultations	0.018**	0.025***	-0.003	0.002
	(0.009)	(0.004)	(0.006)	(0.005)
Emergency	0.049***	-0.008**	0.034***	0.024***
	(0.004)	(0.003)	(0.002)	(0.004)
Long-term & follow-up	0.058***	0.051***	0.057***	0.117***
	(0.005)	(0.003)	(0.003)	(0.007)
Home care	0.020**	0.029***	0.049***	-0.012
	(0.010)	(0.003)	(0.007)	(0.009)
Psychiatric care	0.029***	0.071***	0.076***	0.049***
	(0.005)	(0.004)	(0.007)	(0.017)
Fit statistics				
Observations	1,123	5,260	4,415	2,604
R^2	0.780	0.863	0.742	0.754

Heteroskedasticity-robust standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 4: Separate estimation of input demand function, lagged value as IV, 2013-2022

By excluding teaching hospitals from the analysis, we can more reasonably assume that the remaining hospitals share a homogeneous set of coefficients. This assumption underpins the pooled regression model, the results of which are presented in Table 5. The model incorporates dummy variables and uses lagged values as instrumental variables.

At first glance, it appears that private sectors are indeed more efficient in labor use. In the next section, I will delve into comparisons between individual hospitals and investigate

Dependent Variable:	Nurses	
	Dummy	Dummy IV
Model:	(1)	(2)
Variables		
Constant	1.51***	1.50***
	(0.025)	(0.028)
STAC inpatient	0.291***	0.290***
	(0.004)	(0.005)
STAC outpatient	0.048***	0.048***
	(0.003)	(0.004)
Medical sessions	0.068***	0.068***
	(0.002)	(0.002)
External consultations	0.025***	0.028***
	(0.002)	(0.002)
Emergency	0.019***	0.018***
	(0.001)	(0.001)
Long-term & follow-up	0.066***	0.067***
	(0.002)	(0.002)
Home care	0.026***	0.025***
	(0.002)	(0.003)
Psychiatric care	0.072***	0.071***
	(0.003)	(0.004)
Private Forprofit	-0.258***	-0.245***
	(0.024)	(0.027)
Private Nonprofit	-0.178***	-0.160***
	(0.020)	(0.022)
Fit statistics		
Observations	14,067	12,279
R^2	0.820	0.821

Heteroskedasticity-robust standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 5: Pooled regression with dummy variables, lagged value as IV, 2013-2022

their selection outcomes.

Regression with individual fixed effect Let $\log(x_{it})$ and $\log(y_{it})$ be defined as previously. Additionally, let θ_i represent the fixed effect for hospital i, where θ_i can be interpreted as a measure of labor inefficiency. Specifically, a smaller θ_i indicates greater efficiency in labor use by the hospital. These fixed effects, θ_i, will be instrumental in ranking and selecting hospitals in the subsequent analysis. The model can be specified as follows:

$$\log(x_{it}) = \beta_0 + \beta_1 \log(y_{it}) + \theta_i + \varepsilon_{it} \tag{2}$$

I considered four types of estimator, within-group, first difference, fist difference GMM, system GMM. For expository purposes, the linear specification takes the general form of

$$y_{it} = x_{it}\beta + \theta_i + \epsilon_{it} \quad \text{where} \quad E[\epsilon_{it}|x_{i1}, \dots, x_{it-1}, \theta_i] = 0.$$

The system GMM estimator utilizes two types of moment conditions. The first mirrors those used in the first difference GMM estimator:

$$E[x_{i,t-2}(\Delta y_{it} - \beta \Delta x_{it})]$$

with lagged $x_{i,t-2}$ serving as an instrument for Δx_{it}.

However, if there is high persistency in x_{it}, such that $x_{it} = \alpha x_{i,t-1} + \eta_{it}$ with α close to 1, this affects the strength of the instruments used in the estimation. In this scenario, the reduced form relationship between Δx_{it} and $x_{i,t-2}$ can be expressed as:

$$\Delta x_{it} = (\alpha - 1)\alpha x_{i,t-2} + \alpha \eta_{i,t-1} + \eta_{it}$$

This relationship posits a challenge in GMM estimation due to the potential issue of weak instruments.

The second moment condition makes another assumption, requiring that the correlation between x_{it} and θ_i is the same as that between $x_{i,t-1}$ and θ_i,

$$\mathbb{E}[\Delta x_{i,t-1}(y_{it} - \beta x_{it})] \quad \text{if} \quad \mathbb{E}\left[\Delta x_{i,t-1}(\theta_i + \varepsilon_{i,t})\right] = 0$$

where the current level x_{it} is instrumented by lagged first difference $\Delta x_{i,t-1}$.

Table 6 shows that there's a large difference between the first difference GMM and system GMM, a sign that the second moment condition needs more investigation. Though the estimate from first difference GMM looks more hopeful, the Sargan-Hansen test almost rejects the over-identification null hypothesis for sure, indicating that some moment conditions are not in accordance with each other. Though the issues of weak instruments and rejection of over-identification are intriguing problems, I will set them aside for future investigation since the focus of this paper is more on empirical Bayes application. I will take as given the estimation results from the third column of Table 6 and proceed to the next section.

3 Compound Decision and Empirical Bayes

3.1 Compound decision framework

The idea of compound decision theory was pioneered by Robbins (1956). This approach considers the consequences of all individual decisions collectively. Consider the scenario

Dependent Variable:	Nurses			
	Within Group	First Difference	FD GMM	SYS GMM
Model:	(1)	(2)	(3)	(4)
Variables				
STAC inpatient	0.10***	0.07***	0.13***	0.48***
	(0.00)	(0.01)	(0.03)	(0.02)
STAC outpatient	0.02***	0.01***	0.02	0.05*
	(0.00)	(0.00)	(0.01)	(0.02)
Medical sessions	0.02***	0.02***	0.02***	0.06***
	(0.00)	(0.00)	(0.00)	(0.01)
External consultations	0.00	0.00	0.01	0.07***
	(0.00)	(0.00)	(0.01)	(0.01)
Emergency	0.01***	0.01	0.05	−0.04**
	(0.00)	(0.00)	(0.04)	(0.01)
Long-term & follow-up	0.01***	0.01***	−0.06**	0.07***
	(0.00)	(0.00)	(0.02)	(0.01)
Home care	0.01***	0.02**	−0.00	0.01
	(0.00)	(0.01)	(0.02)	(0.02)
Psychiatric care	0.02***	0.01	0.01	0.07**
	(0.00)	(0.01)	(0.02)	(0.02)
Fit statistics				
Num. obs.	14067	12377	14067	14067
n	1690	1690	1690	1690
T	9	9	9	9
Num. obs. used			10929	23587
Sargan Test: chisq			128.32	360.97
Sargan Test: df			48.00	104.00

Heteroskedasticity-robust standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 6: Estimation of input demand function with individual fixed effect, 2013-2022

where each individual unit is associated with an unobserved parameter θ_i. We are provided with a list of estimates $\hat{\theta}_i$ for each θ_i.

$$\hat{\boldsymbol{\theta}} = (\hat{\theta}_1, \ldots, \hat{\theta}_n) \quad \text{where} \quad \hat{\theta}_i | \theta_i \sim P_{\theta_i}$$

For the moment, I will remain agnostic about the specific decision to make and denote the decision rule by δ.

$$\delta(\hat{\boldsymbol{\theta}}) = (\delta_1(\hat{\boldsymbol{\theta}}), \ldots, \delta_n(\hat{\boldsymbol{\theta}}))$$

The next step is to define the loss function as the objective function to minimize. Since I care about the **collective performance** of my decision, I will define the loss function such that it reflects attention to the compound decision. A natural choice would be to aggregate the individual losses. Therefore, the compound loss function is defined as:

$$L_n(\theta, \delta(\hat{\boldsymbol{\theta}})) = \sum_{i=1}^{n} L(\theta_i, \delta_i(\hat{\theta})).$$

Correspondingly, the compound risk is defined as the expectation of the compound loss.

$$R_n(\theta, \delta(\hat{\boldsymbol{\theta}})) = \mathbb{E}_{\theta|\hat{\theta}}[L_n(\theta, \delta(\hat{\boldsymbol{\theta}}))]$$

We further restrict our attention to the separable decision rule $\delta(\hat{\boldsymbol{\theta}}) = \{t(\hat{\theta}_1), \ldots, t(\hat{\theta}_n)\}$. In order to make the connection with the Bayesian view, under which we assume that $\theta \sim G$, we can rewrite the compound risk as:

$$R_n(\theta, \delta(\hat{\boldsymbol{\theta}})) = \int \int L(\theta_i, t(\hat{\theta}_i)) dP_{\theta_i}(\hat{\theta}_i) dG_n(\theta)$$

where $G_n(\theta)$ is the empirical distribution of θ. [3]

The Frequentist and Bayesian views differ slightly here in the definition of risk. The original compound decision formulation retains the empirical distribution G_n in the compound risk, while the Bayesian risk replaces it with the prior distribution G. On a side note, the two views are somewhat related to the fixed/random effect terminology, in the sense that the fixed effect view treats θ_i as fixed unknown parameters, while the random effect view treats $\boldsymbol{\theta}$ as a random draw from a distribution G. However, in our context, this distinction has nothing to do with whether θ_i is correlated with x_{it}.

The final step involves identifying the decision rule δ^* that minimizes the risk:

$$\delta^* = \arg\min_{\delta} R_n(\theta, \delta(\hat{\boldsymbol{\theta}})) \tag{3}$$

subject to any constraints that we might have. Since G is unknown, the choice between using G_n or G for assessing risk is not particularly critical. For the remainder of this section, I will use the Bayesian risk as the minimization objective and apply constraints that are pertinent to the selection problem. Next, I will focus on the non-parametric estimation of the prior distribution G.

[3] $E_{G_n}(f(x)) = 1/n \sum_i f(x_i)$

3.2 Estimate G

Parametric G Most literature has imposed a parametric form of G. In the case of a Gaussian G, recall the hierarchical model

$$\hat{\theta}_i|\theta_i, \sigma_i \sim P_{\theta_i}$$
$$\theta \sim \mathcal{N}(\mu_\theta, \sigma_\theta^2)$$

There are two hyperparameters to be estimated μ_θ and σ_θ.

If we compare the performance of the posterior mean estimator $\theta^* = \mathbb{E}[\theta|\hat{\theta}]$ with the original estimate $\hat{\theta}$, James and Stein (1992) has shown that there is always an improvement in average performance if we assume G is Gaussian and replace it with an estimate $\hat{G}$. If we relax the normality assumption on G and adopt a nonparametric maximum likelihood estimation (NPMLE) as established by Kiefer and Wolfowitz (1956), there could be further improvements. For example, Jiang and Zhang (2009) has proven that a plugged-in θ^* with a NPMLE $\hat{G}$ is asymptotically optimal among all separable estimators. A comparison between the parametric and non-parametric $\hat{G}$ is demonstrated in Gilraine et al. (2020) on their teacher value-added application.

Nonparametric G The initial NPMLE (Nonparametric Maximum Likelihood Estimator) as defined by Kiefer and Wolfowitz (1956) takes the following form:

$$\hat{G} = \arg\min_{G \in \mathcal{G}} \left\{ -\sum_{i=1}^{n} \log g(y_i) \mid g(y_i) = \int \mathbb{P}(y_i|\theta) dG(\theta) \right\}$$

where $\mathbb{P}(y_i|\theta)$ represents the probability density function of y_i conditional on the true parameter θ, and $g(y_i)$ is the marginal pdf of y_i.

This optimization problem is convex, featuring a strictly convex objective and a convex constraint set. However, it involves an infinite-dimensional parameter space. To tackle the primal problem, it is necessary to discretize it.

The algorithm proposed by Koenker and Mizera (2014) utilizes the interior point method, a technique frequently employed in convex optimization implemented by `MOSEK` (Andersen and Andersen, 2010). This approach significantly enhances computational efficiency compared to the fixed point EM iteration method previously suggested by Jiang and Zhang (2009). This advancement in computational strategies has made the implementation of the Nonparametric Maximum Likelihood Estimator (NPMLE) more practical and faster for applied statistical analysis, offering a robust alternative for dealing with complex data models where traditional methods may falter.

4 The selection problem

4.1 Definition

The definition of the selection problem is taken from the work of Gu and Koenker (2023). Instead of focusing on the right tail of the distribution, the top performers in my context correspond to the left tail. The task at hand is to select the bottom 20% of the θ_i and compare the share of public and private sectors in the meritorious group. This offers another perspective on the public and private sectors, different from that of Croiset and Gary-Bobo (2024).

In addition to the constraint on the size of the selected group (20%), I have further imposed a constraint on the number of false positive mistakes made in the selection process. This leads to the implementation of a false discovery constraint at level γ.

$$\frac{\mathbb{E}_G\left[h_i = 0, \delta_i = 1\right]}{\mathbb{E}_G\left[\delta_i\right]} \leq \gamma$$

where $h_i = 1\left\{\theta_i < \theta_\alpha\right\}$ is the indicator function of whether the unit i is truly below the threshold θ_α. And $\delta_i = 1$ when unit i is selected.

All in all, we can formally define the loss function of selection problem as

$$L(\delta, \theta) = \sum_i h_i(1-\delta_i) + \tau_1\left(\sum_i (1-h_i)\delta_i - \gamma\delta_i\right) + \tau_2\left(\sum_i \delta_i - \alpha n\right)$$

and the optimal decision rule is given by

$$\begin{aligned}
\delta^* &= \arg\min_\delta \mathbb{E}_G\mathbb{E}_{\theta|\hat{\theta}}\left[L(\delta, \theta)\right] \\
&= \mathbb{E}_G \sum_i \mathbb{E}_{\theta|\hat{\theta}}(h_i)(1-\delta_i) + \tau_1\left(\sum_i (1-\mathbb{E}_{\theta|\hat{\theta}}(h_i))\delta_i - \gamma\delta_i\right) + \tau_2\left(\sum_i \delta_i - \alpha n\right) \\
&= \mathbb{E}_G \sum_i v_\alpha(\hat{\theta})(1-\delta_i) + \tau_1\left(\sum_i (1-v_\alpha(\hat{\theta}))\delta_i - \gamma\delta_i\right) + \tau_2\left(\sum_i \delta_i - \alpha n\right)
\end{aligned} \tag{4}$$

Here, the term $\mathbb{E}_{\theta|\hat{\theta}}(h_i)$ is called **posterior tail probability**. It is the probability of i being truly in the bottom α% given the estimated $\hat{\theta}$. This is a posterior statistic different from the posterior mean $\mathbb{E}_{\theta|\hat{\theta}}(\theta_i)$ because the variable inside the expectation $h_i = 1\{\theta_i < G^{-1}(\alpha)\}$ is specific to the capacity constraint at the α level. From the previous section, we have obtained an estimate of the prior distribution G so that we can derive the posterior tail probability $v_\alpha(\hat{\theta}_i)$.

$$v_\alpha = P(\theta_i < \theta_\alpha | \hat{\theta})$$

If we know that $\hat{\theta}|\theta \sim P_\theta$ with density function p_θ, the posterior tail probability can be further expressed as:

$$v_\alpha(y_i) = \frac{\int_{-\infty}^{\theta_\alpha} p_{\theta_i}(y_i) dG(\theta_i)}{\int_{-\infty}^{\infty} p_{\theta_i}(y_i) dG(\theta_i)}$$

From now on, the notation $\hat{\theta}_i$ is replaced by y_i. For example, if $y_i|\theta_i \sim \mathcal{N}(\theta_i, \sigma_i^2)$, as is often the case in applications, v_α takes the explicit form:

$$= \frac{\int_{-\infty}^{\theta_\alpha} \varphi(y_i|\theta_i, \sigma_i^2) dG(\theta_i)}{\int_{-\infty}^{\infty} \varphi(y_i|\theta_i, \sigma_i^2) dG(\theta_i)}$$

where φ is the density function of y_i with conditional mean θ_i and variance σ_i^2.

Here, we have assumed that σ_i is known, meaning that P_θ only depends on θ_i. However, sometimes σ_i is unknown, meaning that P_θ depends on additional parameters. These two cases must be distinguished when defining the posterior tail probability and constraints.

The two constraints can be preliminarily written out as follows:

- Capacity constraint:

$$\mathbb{P}(v_\alpha > \lambda_1^*) \leq \alpha \Rightarrow \lambda_1^* = H^{-1}(1-\alpha).$$

  Empirically, λ_1^* is found by the inverse function of the empirical cumulative distribution H of v_α at $1-\alpha$.

- False Discovery constraint:

$$\mathbb{P}(\theta < \theta_\alpha | v_\alpha > \lambda_2^*) \leq \gamma \Rightarrow \frac{\sum_i \mathbb{E}[(1-v_{\alpha,i})\delta_i]}{\sum_i E[\delta_i]} \leq \gamma.$$

  First, We approximate $\mathbb{P}(\theta < \theta_\alpha | v_\alpha > \lambda_2^*)$ by $\frac{\sum_i E[(1-h_i)\delta_i]}{\sum_i E[\delta_i]}$. Then it needs to be shown that $\mathbb{E}[(1-h_i)\delta_i]$ is equivalent to $\mathbb{E}[(1-v_{\alpha,i})\delta_i]$. This is straightforward by the law of iterated expectation. Let $D_i = (Y_i, \sigma_i^2)$ when σ_i^2 is known and $D_i = (Y_i, S_i)$ when σ_i^2 is unknown.

$$\mathbb{E}[(1-h_i)\delta_i] = \mathbb{E}[\mathbb{E}[(1-h_i)\delta_i | D_i]] = \mathbb{E}[\delta_i(1-\mathbb{E}[h_i|D_i])] = \mathbb{E}[\delta_i(1-v_{\alpha,i})]$$

These formulations allow us to control the size of the selected group while also maintaining the expected rate of false discoveries within an acceptable level γ.

Now, we will critically distinguish between the two scenarios where σ_i is known and unknown. The posterior tail probability and constraints are defined accordingly.

Known variance, $G(\theta)$ The true inefficiency value of hospital i is θ_i, We only observe a sequence of Y_i where

$$Y_i = \theta_i + \varepsilon_{it} \quad \varepsilon_{it} \sim \mathcal{N}(0, \sigma_i^2) \quad (\theta_i) \sim G$$

The tail probability v_α is a function of y_i only

$$v_\alpha(y_i) = \mathbb{P}(\theta_i < \theta_\alpha | y_i) = \frac{\int_{-\infty}^{\theta_\alpha} \varphi(y_i|\theta_i, \sigma_i^2) dG(\theta_i)}{\int \varphi(y_i|\theta_i, \sigma_i^2) dG(\theta_i)}$$

The cutoff λ^* is determined such that the constraints are satisfied

$$\mathbb{P}(v_\alpha > \lambda^*) \leq \alpha$$
$$\mathbb{P}(\theta < \theta_\alpha | v_\alpha > \lambda^*) \leq \gamma$$

Unknown variance, $G(\theta, \sigma)$ We only observe a sequence of Y_{it} where

$$Y_{it} = \theta_i + \varepsilon_{it} \quad \varepsilon_{it} \sim \mathcal{N}(0, \sigma_i^2) \quad (\theta_i, \sigma_i^2) \sim G$$

Neither θ_i nor σ_i^2 is known. But there exists two sufficient statistics for (θ_i, σ_i) such that

$$Y_i = \frac{1}{T_i}\sum_{t=1}^{T_i} Y_{it} \quad \text{where} \quad Y_i|\theta_i, \sigma_i^2 \sim \mathcal{N}(\theta_i, \sigma_i^2/T_i)$$

$$S_i = \frac{1}{T_i - 1}\sum_{t=1}^{T_i}(Y_{it} - Y_i)^2 \quad \text{where} \quad S_i|\sigma_i^2 \sim \Gamma(\frac{(T_i-1)}{2}, \frac{2\sigma_i^2}{T_i - 1})$$

Now the posterior tail probability v_α is a function of both y_i and s_i

$$\begin{aligned} v_\alpha(y_i, s_i) &= \mathbb{P}(\theta_i < \theta_\alpha | y_i, s_i) \\ &= \frac{\int\int_{-\infty}^{\theta_\alpha} \Gamma(s_i|\frac{(T_i-1)}{2}, \frac{2\sigma_i^2}{T_i-1})\varphi(y_i|\theta_i, \frac{\sigma_i^2}{T_i})dG(\theta, \sigma^2)}{\int\int \Gamma(s_i|\frac{(T_i-1)}{2}, \frac{2\sigma_i^2}{T_i-1})\varphi(y_i|\theta_i, \frac{\sigma_i^2}{T_i})dG(\theta, \sigma^2)} \end{aligned}$$

The cutoff λ^* is found in the same way as before.

4.2 Results

Having defined the selection problem, I will now present the results by incorporating empirical estimation of the prior distribution G.

To ensure the variability of the data, I have selected hospitals with than 6 years of observations. In the end, the sample contains 1661 hospitals, out of which 658 are public hospitals and 1003 are private ones. The Y_{it} in the last section is calculated as $\log(x_{it,\text{nurses}}) - \log(y_{it,\text{output}})\hat{\beta}$. Since the panel is too short to invoke central limit theorem, I am obliged to impose normality assumption on the error term $\varepsilon \sim \mathcal{N}(0, \sigma_i^2)$ in order to apply the results above. Thus, Y_{it} follows a normal distribution $\mathcal{N}(\theta_i, \sigma_i^2)$ as the number of hospitals N tends to infinity.

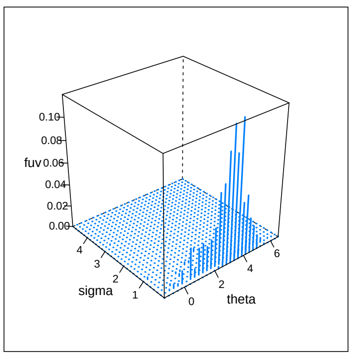
Instead of assuming that σ_i^2 (or the distribution of it) is known, in our setting it seems more reasonable to employ the estimates of it S_i in defining v_α. However, in empirical studies such as baseball batting averages (Gu and Koenker, 2017), teacher added value (Gilraine et al., 2020) and kidney dialysis center rating (Gu and Koenker, 2023), an estimate of the variance is taken to be the true value. A comparison of selection results under the two different assumptions will be presented in this section.

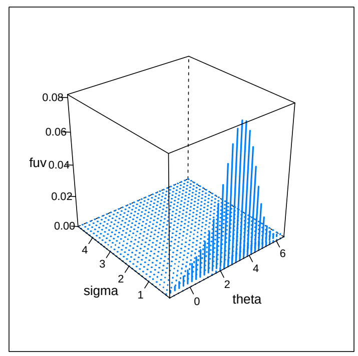
Unknown variance, and TP rules Since we only observe the sample mean Y_i and sample variance S_i, the prior G is a two-dimensional distribution on (θ, σ^2). Without assuming independence between the two parameters, we can utilize a two-dimensional gridding strategy in defining the convex objective function:

$$\hat{G} = \arg\min_{G \in \mathcal{G}} \left\{ -\sum_{i=1}^{n} \log g(y_i, s_i) | g(y_i, s_i) = \int \int \mathbb{P}(y_i, s_i | \theta, \sigma^2) dG(\theta, \sigma^2) \right\}$$

This can be solved using the interior point method, similar to the one-dimensional case. The solution is an atomic distribution with fewer than n atoms. It is worth mentioning that the NPMLE method is self-regularizing because the mass points are determined by the solution without the need for any tuning parameter. Further smoothing is justified by the fact that we have ignored the variability of G. The bandwidth of the biweight kernel for smoothing was chosen based on the mean absolute deviation from the median of the discrete $\hat{G}$.

The left-hand side of Figure **??** shows the estimated $\hat{G}(\theta, \sigma^2)$ before smoothing, while the right-hand side displays the smoothed version.





With an estimated prior, the tail probability function as well as the constraints are well-defined. However, given the discrete nature of selection, it resembles knapsack discrete optimization problem similar. I follow the approach described in Basu et al. (2018) and thus consider only sequentially selecting the units until one constraint is violated.

In Figure 1, I present the results of selecting the top 20% of hospitals with or without the FDR constraint set at 20%. The selection rule is the posterior tail probability, which is explained in the sections above, that is, the solution to the problem defined in 4. The prior G is taken to be the smoothed Kiefer-Wolfowitz estimate.

The left-hand side corresponds to the selection outcome without imposing the FDR constraints, while the right-hand side controls the expected FDR at 20%. In the first case, there are around 10 times more private hospitals in the top 20%, while the total number of hospitals is less than twice that of the public.

The FDR seems to have impacted only the private hospitals, leaving 18 out of the selection set. A more stringent FDR constraint would lead to a smaller set, as shown in 2.

For an overview of other selection rules using different ranking statistics, such as the posterior mean, MLE face value, and James-Stein linear shrinkage, see Appendix A.4.1.

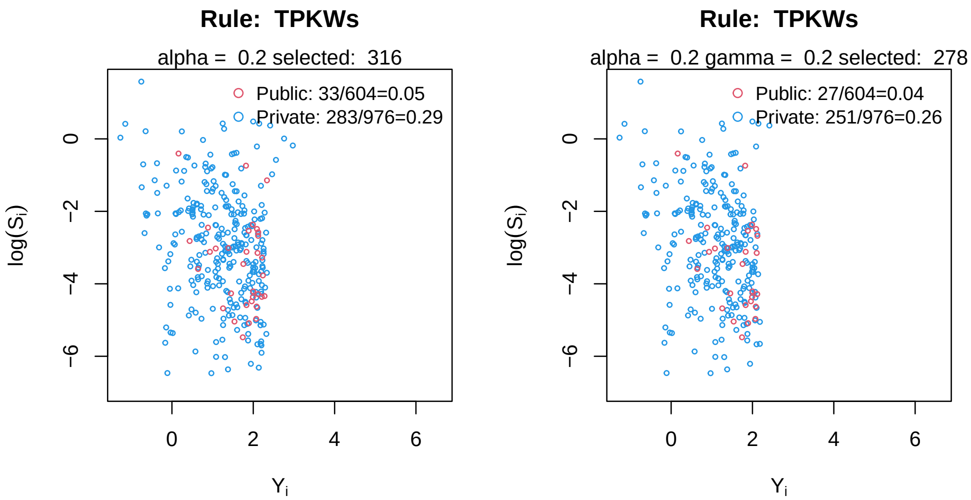


Figure 1: Tail probability rule, capacity 20%, FDR 20%, unknown variance

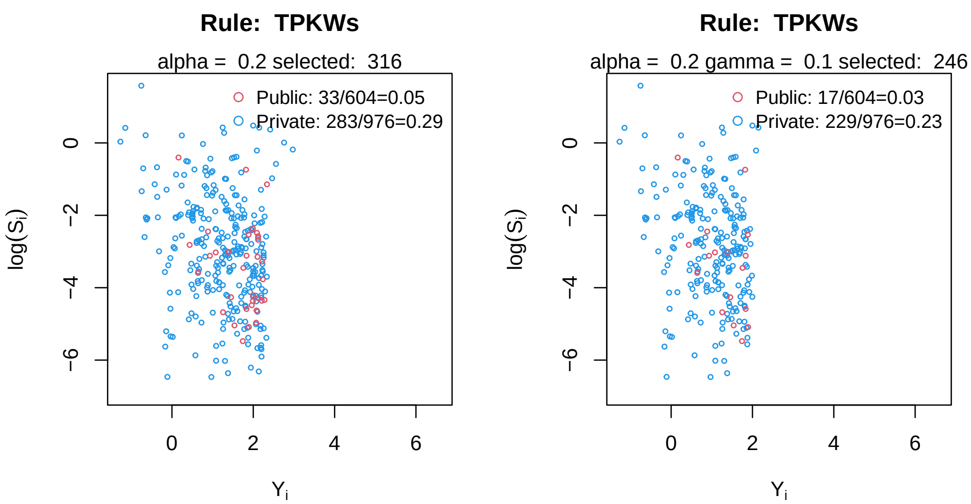
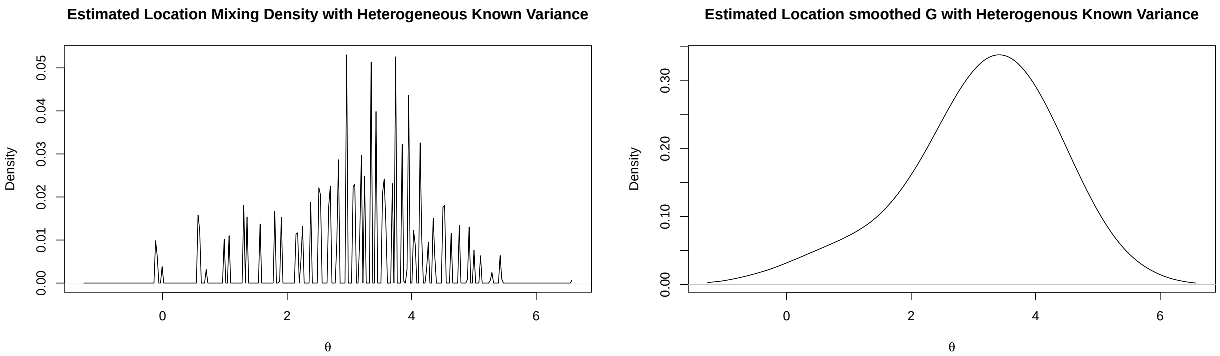


Figure 2: Tail probability rule, capacity 20%, FDR 10%, unknown variance

Known variance, and TP rules In Gu and Koenker (2023), the authors apply the newly proposed selection method to the selection of kidney dialysis centers, a topic previously studied by Lin et al. (2006, 2009). However, their focus is on the quality of service, specifically on the mortality rate. Furthermore, they assume that the predictions of expected mortality are sufficiently accurate such that the variance is known, independent of $\theta \sim G$.

Figure **??** shows the estimated $\hat{G}(\theta)$ before and after smoothing.

Though not desirable in the present setting, it would be interesting to see what the results would be if I take the S_i as the σ_i^2. For the moment, whether this assumption would lead



to a more stringent selection outcome is unclear. Figure 3 presents the outcome under the posterior tail probability rule with a smoothed estimated prior. It seems that with only the capacity constraint, the outcome does not differ much. However, when an FDR constraint is combined, the known variance assumption becomes too lenient to incorporate the newly imposed constraint. At the level $\alpha = 0.2$ and $\gamma = 0.2$, the FDR constraint is not binding. However, a more stringent FDR constraint at $\gamma = 10\%$ does bind, as shown in Figure 4.

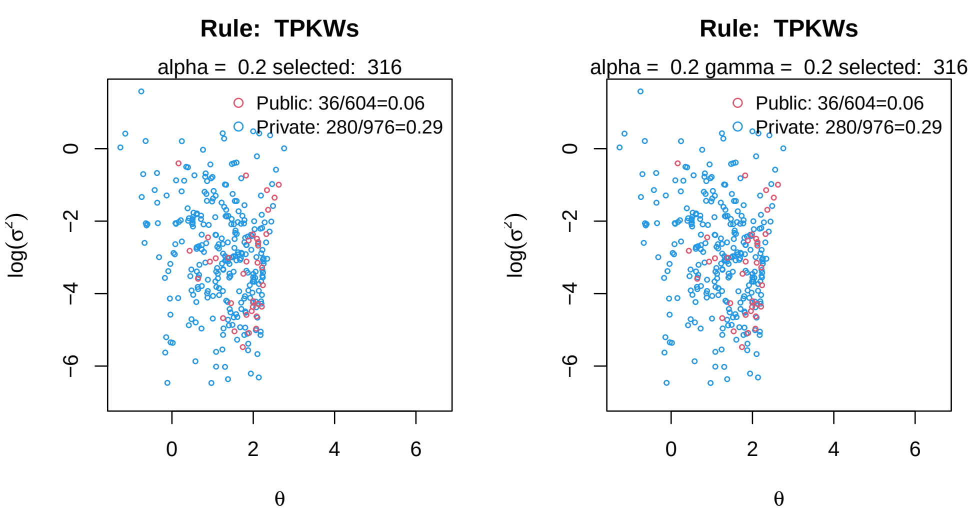


Figure 3: Tail probability rule, capacity 20%, FDR 20%, known variance

Appendix A.4.2 presents the results of other selection rules. In this case, a contour line can be drawn to highlight the differences between ranking statistics.

5 Conclusion

Exploiting the rich dataset covering all hospitals in France, I have attempted to estimate the fixed effect of individual hospitals, which has the interpretation of an *inefficiency index*.

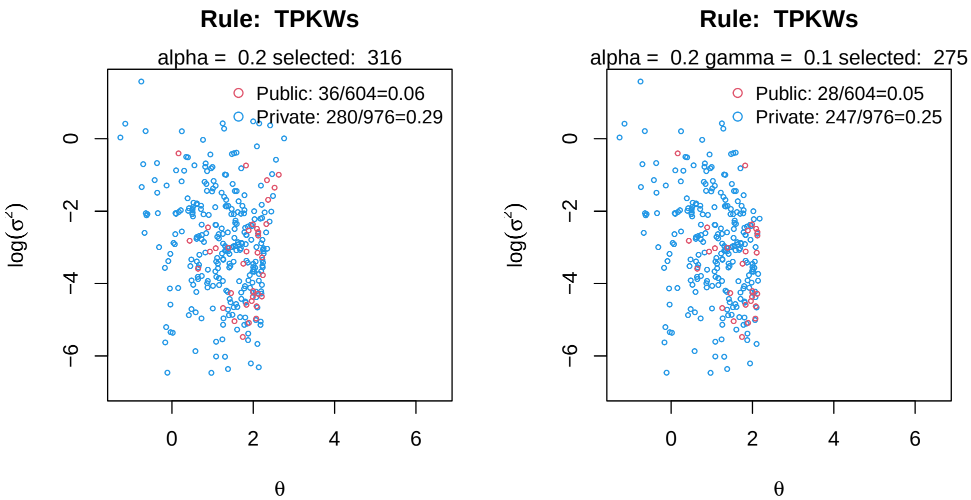


Figure 4: Tail probability rule, capacity 20%, FDR 10%, known variance

Based on an initial estimate, my goal is to select the top-performing hospitals and classify them by legal status, thereby having a more granular view of the efficiency comparison between public and private hospitals. I deliberately omit the teaching hospitals (approximately 200) from the dataset because they inherently differ in terms of objectives. To tackle the endogeneity issue, I used the lagged difference of the regressors as instruments for the current level, a simplified version of the system GMM method. The selection problem of interest falls naturally under the compound decision framework. The Empirical Bayes ideology, along with developments in nonparametric maximum likelihood estimation, has made the implementation more efficient. In addition to the artificial capacity constraint where only the top α% of units are of interest, it is an interesting and useful practice to incorporate another constraint called the False Discovery Rate (FDR) constraint such that the expected number of false positives is controlled at a certain level. From the application to French hospitals, it is clear that the FDR constraint shrinks the selection set by some amount. It is also intuitive that the larger the capacity (the larger the α), the less binding the FDR constraint is. The idea is that when the decision-maker can select more units, the probability of making mistakes decreases. Another observation comes from the assumption we make in the NPMLE of G. In Gu and Koenker (2023), the authors have pointed out that while the known variance assumption in $Y_i|\theta_i, \sigma_i$ may be plausible in some applications, it is more common to be faced with only an estimate of the variance. The two assumptions give rise to a different level of *stringency* in response to the constraints, especially when the decision-maker wants to control for the expected false discovery rate. Assuming an unknown variance treats the observation as noisier, thus increasing the probability of making mistakes. The same level of FDR constraint of 20% only binds in the unknown variance scenario. With respect to the private-public comparison, among the top 20% performers, there are around ten times more private than public hospitals, while the ratio of total number is 5 to 3. A preliminary conclusion is that in terms of labor employment efficiency, there are more efficient private hospitals among the top performers. It may be of interest to healthcare authorities to perform such selections and take corresponding actions with respect to the selection outcome. From my point of view, a related report based on the ranking and selection results will create an incentive for healthcare providers to ensure the completeness of data input. However, one cautionary note is the interpretation of the fixed effect estimate. Since the fixed effect captures all time-invariant components of the unit, whether it is only the unobserved heterogeneity of individual hospitals or an actual measure of inefficiency is questionable. This issue is discussed in Greene (2005). Lastly, despite the fact that it is human nature to construct rankings and make selections, every step of the procedure requires attention to specification, identification, and justifiable assumptions. Incorporating constraints such as FDR in defining the problem may be helpful, but the decision is still subject to great uncertainty and should be made with caution and justification.

References

Aigner, D., Lovell, C. K., and Schmidt, P. (1977). Formulation and estimation of stochastic frontier production function models. *Journal of econometrics*, 6(1):21–37.

Andersen, E. D. and Andersen, K. D. (2010). The mosek optimization tools manual, version 6.0.

Arellano, M. and Bond, S. (1991). Some tests of specification for panel data: Monte carlo evidence and an application to employment equations. *The review of economic studies*, 58(2):277–297.

Arellano, M. and Bover, O. (1995). Another look at the instrumental variable estimation of error-components models. *Journal of econometrics*, 68(1):29–51.

Basu, P., Cai, T. T., Das, K., and Sun, W. (2018). Weighted false discovery rate control in large-scale multiple testing. *Journal of the American Statistical Association*, 113(523):1172–1183.

Blundell, R. and Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of econometrics*, 87(1):115–143.

Charnes, A., Cooper, W. W., and Rhodes, E. (1978). Measuring the efficiency of decision making units. *European journal of operational research*, 2(6):429–444.

Chetty, R., Friedman, J. N., and Rockoff, J. E. (2014). Measuring the impacts of teachers i: Evaluating bias in teacher value-added estimates. *American economic review*, 104(9):2593–2632.

Chetty, R. and Hendren, N. (2018). The impacts of neighborhoods on intergenerational mobility ii: County-level estimates. *The Quarterly Journal of Economics*, 133(3):1163–1228.

Croiset, S. and Gary-Bobo, R. (2024). Are public hospitals inefficient? an empirical study on french data.

Gilraine, M., Gu, J., and McMillan, R. (2020). A new method for estimating teacher value-added. Technical report, National Bureau of Economic Research.

Greene, W. (2005). Fixed and random effects in stochastic frontier models. *Journal of productivity analysis*, 23:7–32.

Gu, J. and Koenker, R. (2017). Empirical bayesball remixed: Empirical bayes methods for longitudinal data. *Journal of Applied Econometrics*, 32(3):575–599.

Gu, J. and Koenker, R. (2023). Invidious comparisons: Ranking and selection as compound decisions. *Econometrica*, 91(1):1–41.

James, W. and Stein, C. (1992). Estimation with quadratic loss. In *Breakthroughs in statistics: Foundations and basic theory*, pages 443–460. Springer.

Jiang, W. and Zhang, C.-H. (2009). General maximum likelihood empirical Bayes estimation of normal means. *The Annals of Statistics*, 37(4):1647 – 1684.

Kiefer, J. and Wolfowitz, J. (1956). Consistency of the maximum likelihood estimator in the presence of infinitely many incidental parameters. *The Annals of Mathematical Statistics*, pages 887–906.

Kline, P., Rose, E. K., and Walters, C. R. (2022). Systemic discrimination among large us employers. *The Quarterly Journal of Economics*, 137(4):1963–2036.

Koenker, R. and Gu, J. (2017). Rebayes: an r package for empirical bayes mixture methods. *Journal of Statistical Software*, 82:1–26.

Koenker, R. and Mizera, I. (2014). Convex optimization, shape constraints, compound decisions, and empirical bayes rules. *Journal of the American Statistical Association*, 109(506):674–685.

Lin, R., Louis, T. A., Paddock, S. M., and Ridgeway, G. (2006). Loss function based ranking in two-stage, hierarchical models. *Bayesian Analysis (Online)*, 1(4):915.

Lin, R., Louis, T. A., Paddock, S. M., and Ridgeway, G. (2009). Ranking usrds provider specific smrs from 1998–2001. *Health Services and Outcomes Research Methodology*, 9:22–38.

Lindsay, B. G. (1995). Mixture models: theory, geometry, and applications. Ims.

Meeusen, W. and van Den Broeck, J. (1977). Efficiency estimation from cobb-douglas production functions with composed error. *International economic review*, pages 435–444.

Robbins, H. E. (1956). An empirical bayes approach to statistics. In *Proceedings of the third berkeley symposium on mathematical statistics and probability*, volume 1, pages 157–163.

A Appendix

A.1 Data

The panel is first filtered by the following criteria

1. the number of nurses is positive,
2. at least one of STAC inpatient, STAC outpatient, Sessions is positive,
3. the number of observations is larger than 6

Second, I add one to every variable to avoid null value when taking log.

A.2 NPMLE G

Koenker and Mizera (2014) defined the primal problem as

$$\min_{f=dG}\left\{-\sum_i \log g(y_i)\middle| g(y_i)=T(f),\ K(f)=1,\ \forall i\right\}$$

where $T(f)=\int p(y_i|\theta) f d\theta$ and $K(f)=\int f d\theta$.
By discretizing the support,

$$\min_{f=dG}\left\{-\sum_i \log g(y_i)\middle| g=Af,\ 1^T f=1\right\}$$

where $A_{ij}=p(y_i|\theta_j)$ and $f=(f(\theta_1), f(\theta_2), \ldots, f(\theta_m))$.
It is straightforward to derive the dual problem

$$\max_{\lambda,\mu}\left\{\sum_i \log \lambda_1(i)\middle| A^T\lambda_1 < \lambda_2 1,\ (\lambda_1 > 0)\right\}$$

A.3 Assumption on $\hat{\theta}_i|\theta_i, \sigma_i$

If our specification and assumptions on exogeneity are correct, the consistency of $\hat{\beta}$ is guaranteed by N's asymptotic. However, our estimate of the fixed effect is

$$\begin{aligned}\hat{\theta}_i &= \frac{1}{T}\sum(\theta_i+\varepsilon_{it}+x_{it}(\beta-\hat{\beta}))\\ &\xrightarrow{N\to\infty} \theta_i+\frac{1}{T}\sum_t \varepsilon_{it}\end{aligned}$$

When T is relatively small (or even fixed), I am not in a good position to use central limit theorem to claim that $\hat{\theta}_i \xrightarrow{d} \mathcal{N}(\theta_i, \frac{\sigma_i^2}{T})$. A bold assumption that $\varepsilon_{it} \sim \mathcal{N}(0, \sigma_i^2)$ will save me from the T issue, which I will impose for the rest of the section (and abstract from whether that for each i is a testable/reasonable/feasible assumption).

A.4 Comparison of selection rules

A.4.1 Unknown variance

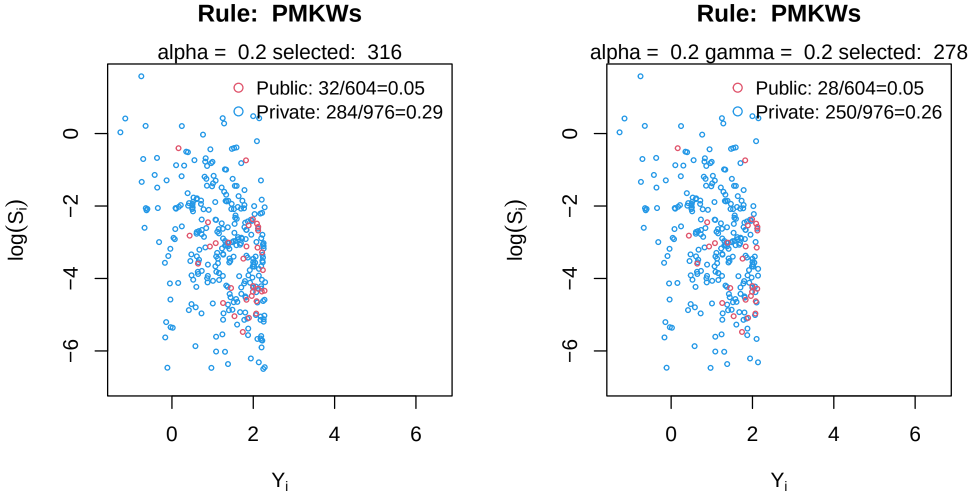


Figure 5: Posterior mean, capacity 20%, FDR 20%, unknown variance

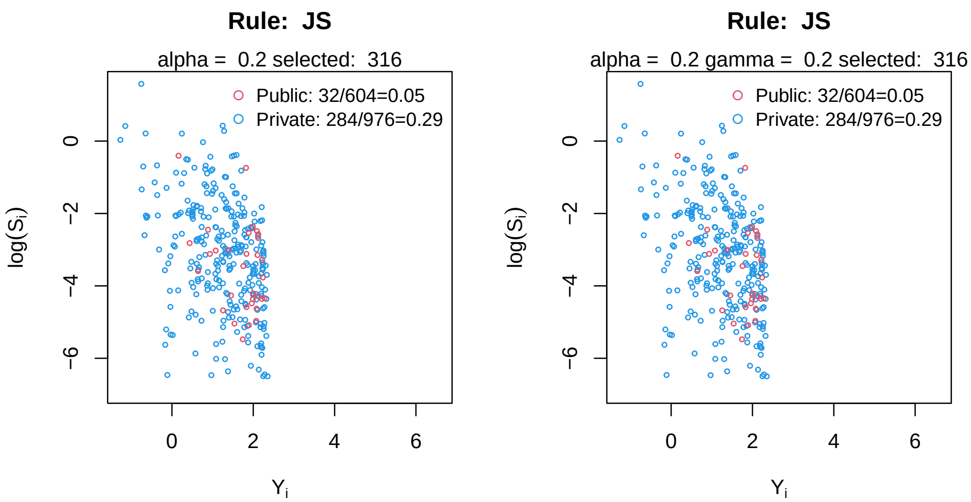


Figure 6: James-Stein Linear Shrinkage, capacity 20%, FDR 20%, unknown variance

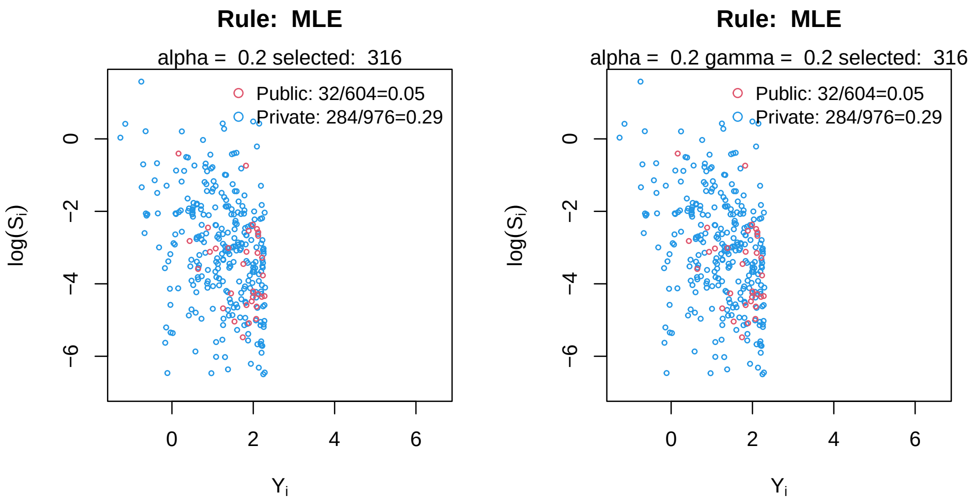


Figure 7: MLE, capacity 20%, FDR 20%, known variance

A.4.2 Known variance

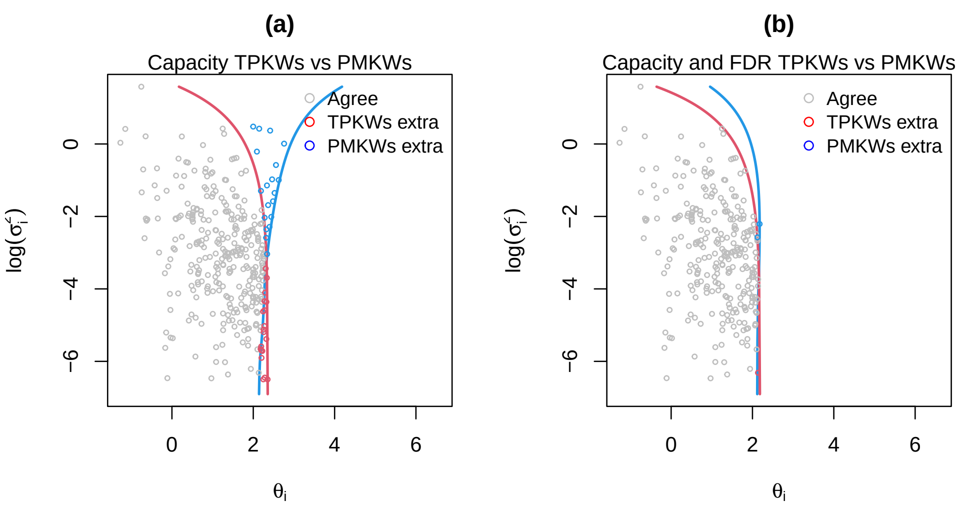


Figure 8: TP VS PM, capacity 20%, FDR 10%, known variance

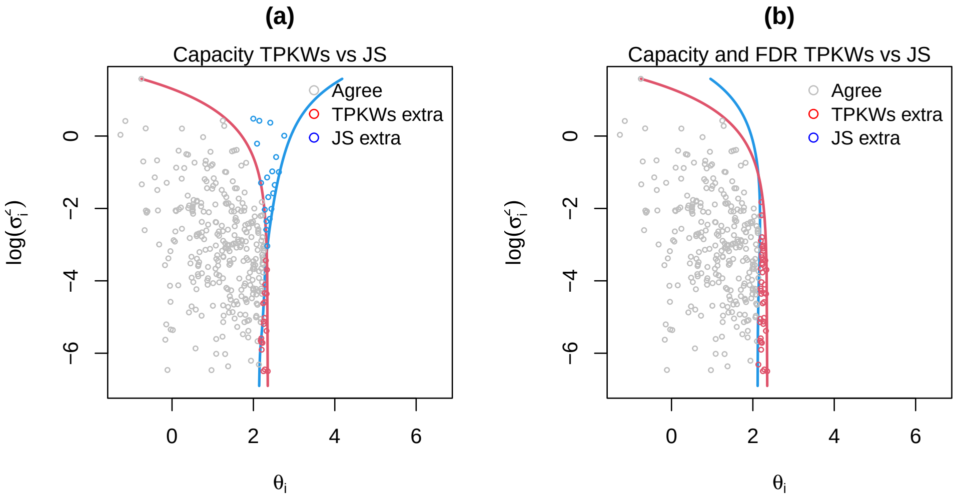


Figure 9: TP VS JS, capacity 20%, FDR 10%, known variance

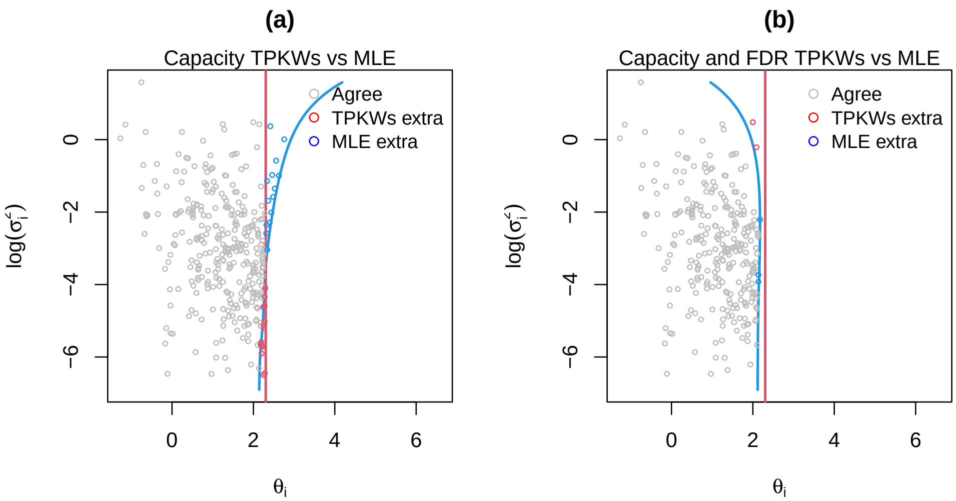


Figure 10: TP VS MLE, capacity 20%, FDR 10%, unknown variance