

# ARE PUNCTUAL AND REGULAR TRAIN LINES EFFICIENT? RE-EVALUATING REGIONAL RAILWAY PERFORMANCE WITH A SERVICE QUALITY- INTEGRATED DEA FRAMEWORK

---

FABRE N., LAROCHE F., EBOLI L., MAZZULLA G., BOUZOUINA L.

Une thèse entre le LAET et la Chambre régionale des comptes AURA (Cour des comptes) :

- Contrat COFRA
- Appui des équipes pour les contrôles « transport » dont l'epp ferroviaire AURA

Dans un contexte d'ouverture à la concurrence du ferroviaire régional français

- Depuis décembre 2023, tous les nouveaux contrats d'exploitation régionaux doivent être attribués après une mise en concurrence



**RAPPORT D'ÉVALUATION DÉFINITIF  
ET SES RÉPONSES**

**ÉVALUATION DE LA POLITIQUE  
DU MATÉRIEL ROULANT  
FERROVIAIRE DE  
LA RÉGION AUVERGNE-RHÔNE-  
ALPES**

Le présent document, qui a fait l'objet d'une contradiction avec les destinataires concernés,  
a été délibéré par la chambre le 30 janvier 2024.

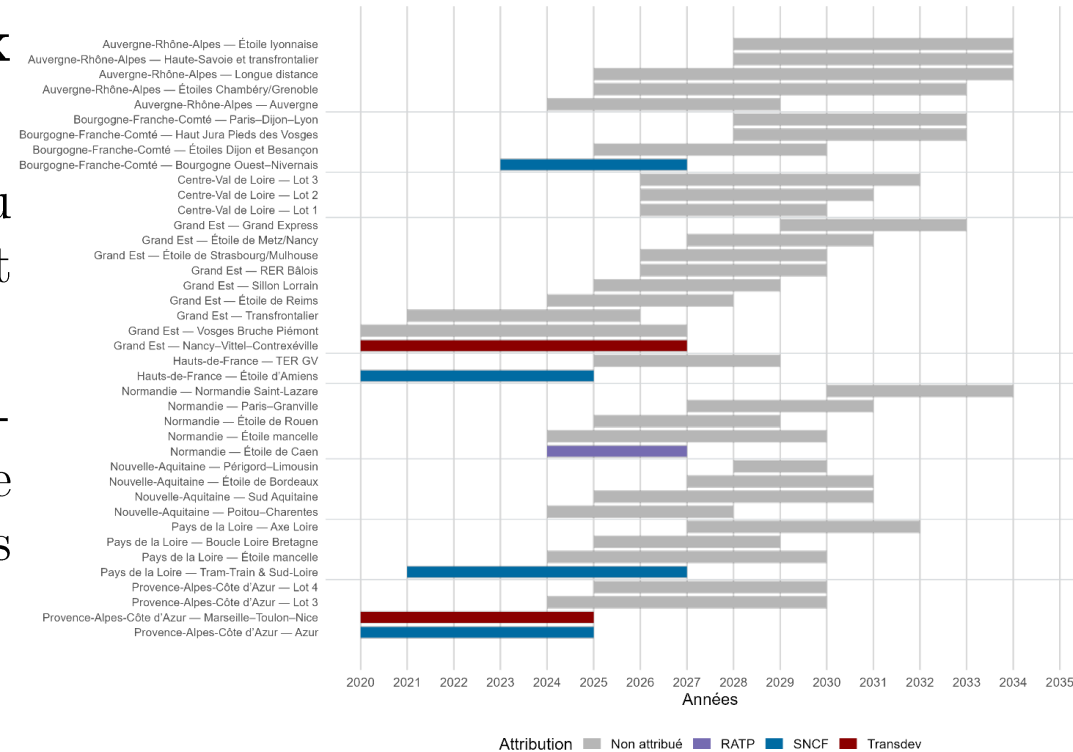
# LA PETITE RÉVOLUTION DE L'OUVERTURE À LA CONCURRENCE

L'ouverture à la concurrence implique deux grands changements dans le système ferroviaire :

- **Une fragmentation** → Perte de la nature réticulaire du ferroviaire. 11 contrats régionaux (hors Idf et Corse) vont laisser place à plus de 40 lots
- **Un rééquilibrage des pouvoirs** dans le couple AOM-opérateur. Les AOM doivent définir des lots, s'assurer de leur cohérence et de la bonne exécution du service sans l'expertise de l'opérateur historique

Un besoin d'évaluer « autrement » :

- **A une échelle plus fine** → Il faut descendre de la région jusqu'au lot ou à la ligne (comme lors de la régionalisation de 2002)
- **Pour piloter « efficacement » le service** → Il s'agit de comprendre d'une part les déterminants de l'efficience et d'autre part inclure le principal concerné : l'utilisateur



## Objectifs de la recherche :

- Produire un indicateur de performance économique au niveau de la ligne pour évaluer un système qui sera fragmenté (le problème du contribuable)
- Ajouter à la performance économique, la performance en termes de qualité du service (le problème de l'utilisateur)
- Il s'agit d'identifier grâce à un indicateur composite et « efficiency-based » les lignes performantes dans la dimension économique et dans la dimension « qualitative »
- C'est une première brique vers un pilotage/régulation par la qualité
- **Problématique :**
  - Comment intégrer la qualité du service à l'analyse de l'efficience ?

## Une définition de l'efficacité dans le ferroviaire :

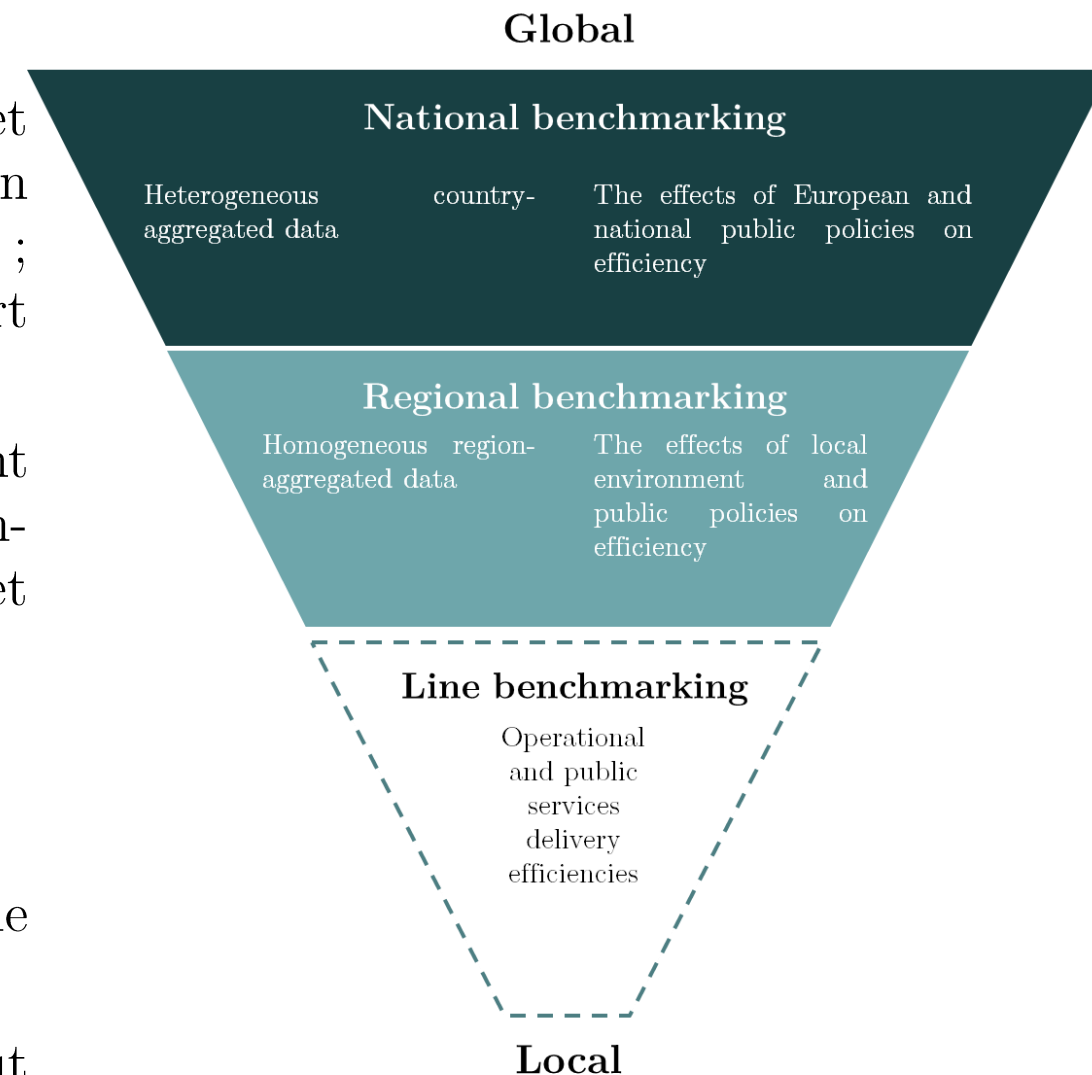
- Les analyses d'efficacité comparent plusieurs unités statistiques (des pays, des régions ou des lignes). Il s'agit de mesurer quelles sont les unités qui optimisent le mieux la production d'outputs (trains-km, voyageurs-km) en fonction d'une série d'inputs (nombre ou coûts de personnels, de matériel roulant, d'énergie etc.). C'est une mesure relative et chaque unité se voit attribuer un score d'efficacité entre 0 et 1. Si une unité produit plus de train-km pour le même volume d'inputs, elle aura un score d'efficacité plus élevé
- Il y a deux méthodes principales pour construire ces scores, les SFA (paramétrique) et les DEA (non-paramétrique)
- Les SFA sont dérivées d'un modèle de régression avec  $y$  un output  $f(\text{inputs})$ . Très interprétable, mais un seul output
- Les DEA permettent d'inclure facilement plusieurs inputs et plusieurs outputs, ce qui nous incite à choisir cette spécification

## Deux familles d'analyse d'efficience

- Les analyses macro qui comparent des pays et notamment les avancées en termes de réformes (Gathon and Perlman, 1992 ; Cantos, 2001 ; Lan and Lin, 2003 ; Couto and Graham, 2009 ; Asmild et al. 2009 ; Merkert et al., 2009 ; Friebe et al., 2010 ; Cantos et al., 2012)
- Les analyses méso, à l'échelle régionale qui comparent des systèmes homogènes avec plus de variables non-discrétionnaires (Link, 2016, 2019 ; Desmaris et Monchambert, 2023 ; Lévêque, 2005)

## Deux « gaps » dans la littérature :

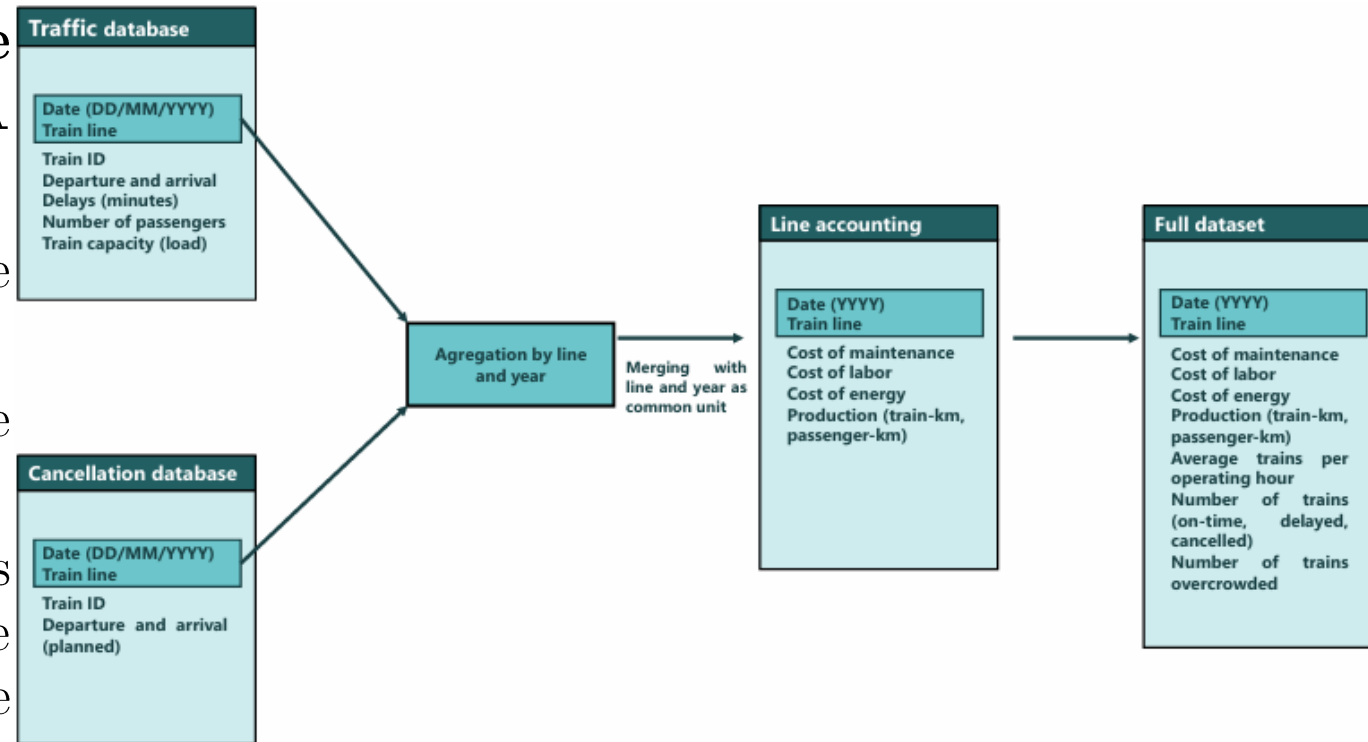
- Des analyses agrégées à la région ou au pays qui ne permettent pas de capter toute l'hétérogénéité
- Peu de littérature intégrant la QS comme output « désirable » sauf Link (2019) sur données régionales



# ÉTUDE DE CAS ET DONNÉES

Un jeu de données complet à l'échelle de la ligne pour la région AURA (2021-2023) :

- Une base de circulation avec le train comme individu statistique
- Une base d'annulation avec le train comme individu statistique
- Une base de comptabilité financière (coûts pour différents postes) et métriques de production (tkm, vkm) avec la ligne et l'année comme individu statistique
- Calcul d'un indicateur de ponctualité, d'annulation, du respect de l'emport à la ligne et l'année
- Jointage avec la base comptabilité
- Panel de 58 lignes (174 observations)



## Analyse par DEA :

- L’observation est la ligne-année (Ligne A\_2021; Ligne A\_2022; Ligne A\_2023)
- L’objet des DEA est de mesurer, pour chaque unité, un ratio d’efficience défini comme le rapport entre une somme pondérée d’outputs et une somme pondérée d’inputs. Les poids des inputs et des outputs sont déterminés par un programme d’optimisation linéaire, qui maximise ce ratio pour l’unité étudiée, sous la contrainte qu’avec ces mêmes poids, aucune unité n’obtienne un ratio supérieur à 1

## Proposition de deux modèles (M1&M2) pour une analyse en « scénario » :

Model	Description	Inputs	Outputs
M1	Standard Efficiency Model	<ul style="list-style-type: none"><li>• Energy Costs (K€)</li><li>• Maintenance Costs (K€)</li><li>• Labor Costs (K€)</li></ul>	<ul style="list-style-type: none"><li>• Train-km</li><li>• Passenger-km</li></ul>
M2	QoS-inclusive Model	<ul style="list-style-type: none"><li>• Energy Costs (K€)</li><li>• Maintenance Costs (K€)</li><li>• Labor Costs (K€)</li></ul>	<ul style="list-style-type: none"><li>• Train-km</li><li>• Passenger-km</li><li>• Average trains per operating hour</li><li>• Punctuality Rate (%)</li><li>• Regularity Rate (%)</li><li>• Non-Overcrowding Rate (%)</li></ul>



Une forte hétérogénéité au sein de lignes de la région (stats descriptives) :

Variable	Mean	Min	Max	Std. Dev.	Database Origin
Maintenance Costs (K€ 2021)	2208.27	41.69	10917.70	2183.69	Line Accounting
Labor Costs (K€ 2021)	2407.24	98.28	10103.71	1999.32	Line Accounting
Energy Costs (K€ 2021)	677.43	13.52	3789.56	714.76	Line Accounting
Train-km (thousands)	494.90	15.52	2128.64	456.26	Line Accounting
Passenger-km (thousands)	47699.31	281.08	332377.22	65315.19	Line Accounting
Average trains per operating hour	1.61	1.00	3.94	0.56	Traffic database
Punctuality Rate (%)	89.12	68.40	99.06	6.31	Traffic database
Regularity Rate (%)	97.44	79.57	99.67	2.22	Cancellation database
Non-Overcrowding Rate (%)	98.06	69.12	100.00	4.29	Traffic database

Scores d'efficience pour M1 (modèle productif) & M2 (modèle avec qs) :

Model Specification	Mean	Median	SD	Min	Max	N
Efficiency Scores from M1	0.62	0.59	0.21	0.28	1.00	174
Efficiency Scores from M2	0.74	0.73	0.23	0.31	1.00	174

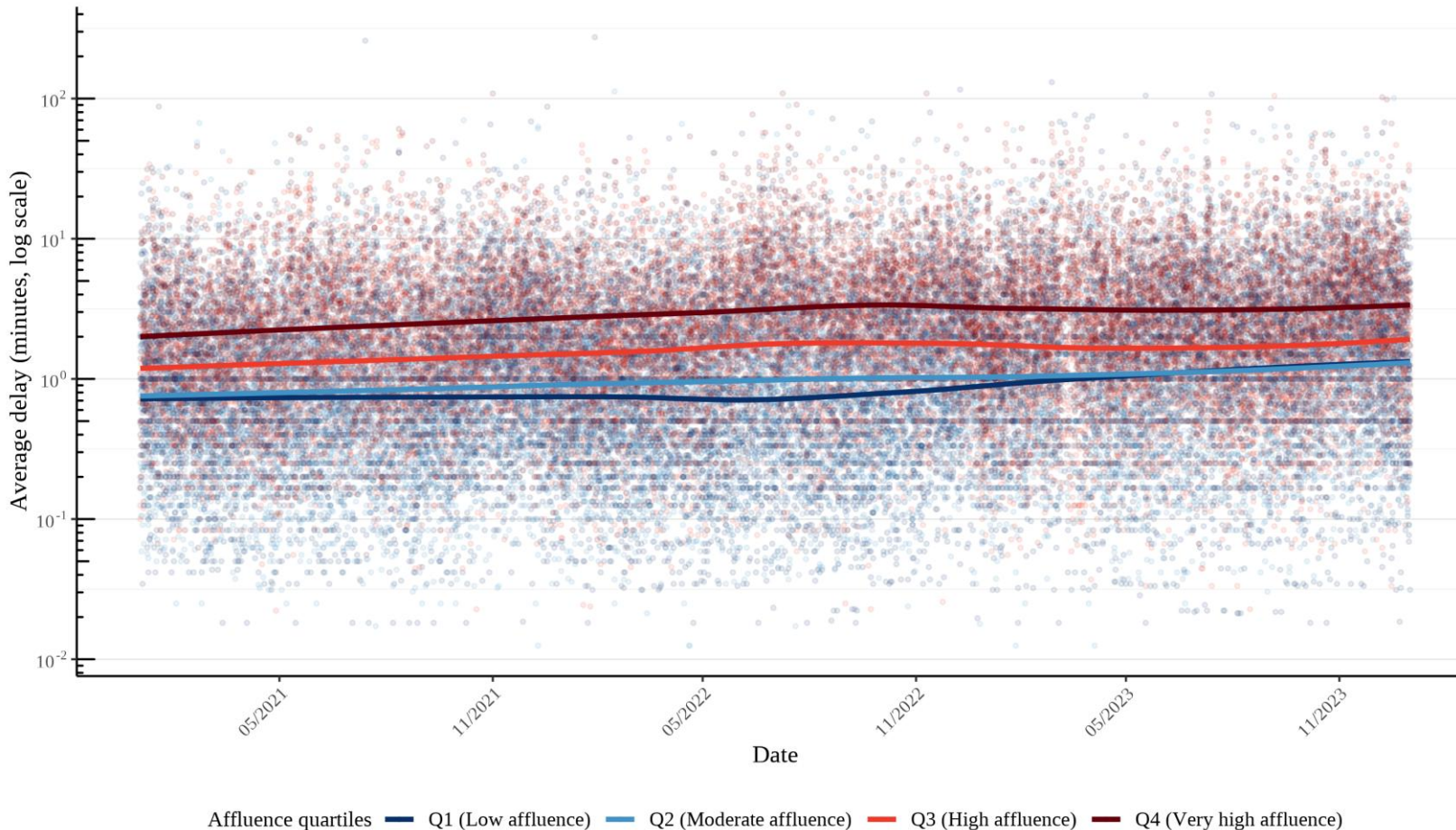
Une moyenne d'efficience de 0,62 signifie qu'en moyenne 62% des inputs auraient suffi à produire 100% des outputs dans un système optimisé

Corrélations entre l'efficience des lignes et les variables de production & qualité :

Variable	M1		M2	
Punctuality Rate (%)	-0.20	***	0.07	
Regularity Rate (%)	0.25	***	0.27	***
Non-Overcrowding Rate (%)	-0.05		0.25	***
Train Cost-Efficiency (€/Train-km)	-0.66	***	-0.51	***
Passenger Cost-Efficiency (€/Passenger-km)	-0.70	***	-0.46	***
Frequency (Trains/hr)	0.08		0.18	*

- Le modèle productif « récompense » les lignes avec une efficience coût, mais pas la qualité du service (retards, fréquence, surcharge),
- Du point de vue de l'efficacité pure, le train à l'heure et le train vide coûtent autant que le train en retard et le train plein

De cette « décorrélation » entre la qualité du service et l'efficacité, on ne peut pas conclure que la mauvaise qualité implique une meilleure efficacité :

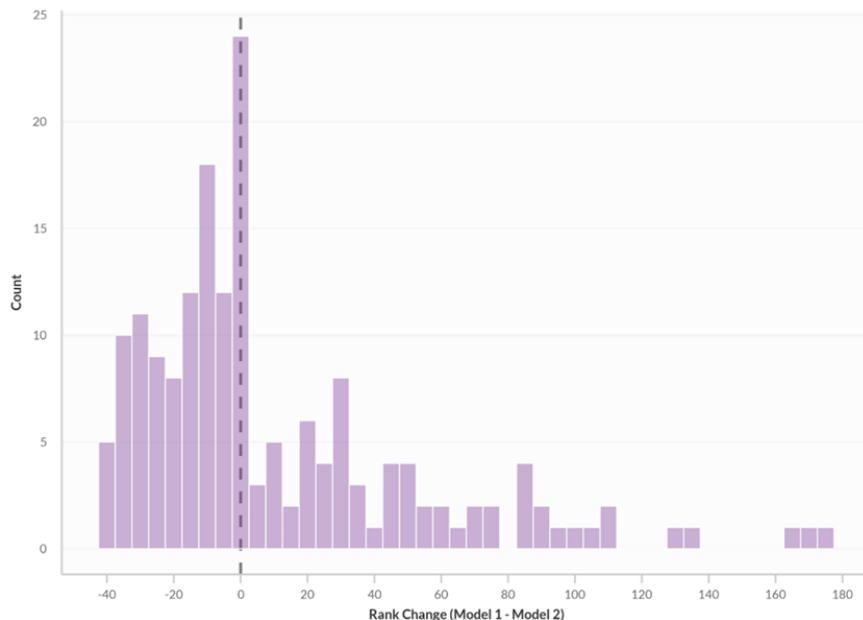


Note: Points represent line-day observations. Zero delays excluded.

- Les lignes les plus fréquentées sont aussi celles où il y a le plus de retards. Plus la densité de trains est élevée, plus la probabilité de retards l'est aussi (Laroche, 2014). Phénomène de déséconomies d'échelle (Coulombel & Monchambert, 2023)
- Ne pas intégrer la qualité du service aux analyses récompense certaines lignes, pas forcément représentatives du point de vue de l'utilisateur. Une observation partagée par Link (2019)

## Comparaison des modèles avec une analyse des rangs :

- On ne peut pas directement comparer les scores d'efficacité entre M1 & M2, puisque les variables utilisées diffèrent
- On regarde alors le rang des unités par modèle (ex : une unité avec un score d'efficacité de 1 aura un rang de 1 ; l'unité avec un score de 0 aura le rang 174)
- On compare les classements entre les deux modèles et on calcule une différence  $M1 - M2$ . Par exemple une ligne avec un rang  $M1 = 50$  et un rang  $M2 = 30$ , s'améliore de 20 places



- Entre M1 et M2, 68 s'améliorent, 20 ne bougent pas, 86 se dégradent
- Distribution étalée à droite : la qualité du service induit une (forte) amélioration mais sélective

Identification des lignes qui s’améliorent, se dégradent, restent les mêmes :

Variable	Improved A Lot	Improved	Unchanged	Worsened	Worsened A Lot	ANOVA Sig.	Kruskal- Wallis Sig.
Rank Categories N (observations)	≥+30 42	+1 to +29 26	0 20	-1 to -29 67	≤-30 19		
Frequency	1.71	1.50	2.04	1.51	1.45	**	
Total Costs* (k€)	4,61	5,11	6,13	5,33	5,98		
Train-km (thousands)	452	511	711	425	588		
Passenger-km (thousands)	37,82	49,46	105,44	31,35	63,96	***	
Punctuality Rate (%)	91.5	89.0	89.0	88.9	85.0	**	**
Regularity Rate (%)	97.8	97.3	97.9	97.3	96.8		
Non-Overcrowding Rate (%)	98.7	99.4	96.8	97.9	96.6		***
€/Train-km	11.8	10.9	9.10	12.7	10.4	***	***
€/Passenger-km	0.43	0.23	0.13	0.25	0.10	*	***

Notes: \*\*\* p < 0.001, \*\* p < 0.01, \* p < 0.05

- On divise les lignes en 5 groupes en fonction de l’intensité du changement
- On mesure les écarts de variance entre les groupes
- On regarde si ces écarts sont significatifs pour différentes variables

Unchanged

- Les lignes les plus efficientes dans les deux modèles, qui cumulent performance financière et de QS
- Regroupe les lignes les plus empruntées
- Mais aussi les lignes très cadencées

Improved

- Une moins bonne efficacité économique
- Mais une bonne qualité du service
- Lignes moyennes, voire rurales ayant moins de contraintes d’exploitation

Worsened

- Des lignes peu performantes en coûts et qualité (worsened)
- Ou des lignes efficaces en coûts, mais très peu en QS (worsened a lot)
- ‘Worsened’ regroupe les lignes moyennes et rurales avec des problèmes d’exploitation
- ‘Worsened a lot’ regroupe des lignes inter-régionales (longue distance et mauvaise qs)

- Les analyses à la ligne dévoilent mieux la diversité et il s'agirait de les systématiser dans le cadre de l'ouverture à la concurrence (ex ante, in itinere, ex post)
- Ne pas inclure la qualité du service dans les analyses de l'efficacité n'est pas forcément représentatif du point de vue de l'utilisateur, puisque les lignes efficaces sont aussi associées à une « mauvaise » QS
- Les DEA permettent de faire ce travail de production d'un indicateur « efficiency-based » qui tienne compte de la QS tout en étant facilement interprétable
- L'analyse des rangs permet d'identifier les lignes à « benchmarker » i.e., le groupe 'unchanged'. Il s'agit de voir dans quelles mesures est-il possible de diffuser les bonnes pratiques (cadencement, investissements préventifs ?) mais aussi d'identifier les causes des défaillances des lignes dégradées (ces lignes sont-elles inefficaces du fait de contraintes d'exploitation ou par manque de moyens ?)
- Un point d'alerte concernant les lignes inter-régionales, dont la QS peut souffrir d'un manque de coordination → le principe de l'allotissement va encore plus désagréger ce système. Si la concurrence apporte de l'efficacité productive (au bénéfice du contribuable) elle peut aussi dégrader la QS (au détriment de l'utilisateur)

MERCI