

Economies d'échelle dans l'industrie du logiciel

Préprint – non évalué par les pairs

Fried-Junior SABAYE¹

Laboratoire IRJI - Université de Tours - Version 1 – Juillet 2025

¹ORCID : 0009-0009-3466-0887

E-mail : sfriedjunior@gmail.com

DOI : 10.5281/zenodo.16312176

Résumé

Cet article évalue empiriquement, à partir de quatorze années de données comptables trimestrielles (2011-2024), la réalité de l'idée du « coût marginal nul » dans l'édition logicielle. Il mesure comment l'introduction, d'abord d'un SaaS hébergé puis d'une architecture Full web — en complément du mode on-premise historique — infléchit la structure des coûts unitaires et les rendements d'échelle d'un éditeur français.

La méthodologie combine quatre briques analytiques articulées : une modélisation log-log Cobb-Douglas régularisée par Ridge, qui livre des élasticités robustes par direction malgré la colinéarité comptable ; une détection statistique de ruptures fondée sur les tests de Chow et de Wald, corrigés de l'hétéroscédasticité et de l'autocorrélation (Newey-West, lag = 4), qui met en évidence deux césures majeures — 2015-T3 (essor du SaaS) et 2020-T1 (lancement du Full web) ; une ré-estimation des coefficients sur trois sous-périodes homogènes couplée à un suivi glissant de seize trimestres, où la somme des élasticités Σ_β passe de 0,70 à 0,25 sous l'effet du SaaS, puis rebondit à 1,12 avec le Full Web ; enfin, une extension Translog pénalisée par Elastic-Net ($\alpha = 0,50$) qui révèle les effets quadratiques et les interactions — synergies résiduelles entre les centres de coûts, friction croissante R&D-Assistance — masqués par la forme linéaire. Les résultats convergent : chaque bascule de mode de commercialisation recompose la hiérarchie des directions de coûts et modifie la pente des économies d'échelle, jusqu'à inverser le signe du rendement global lorsque les investissements Full web excèdent la capacité d'absorption du volume.

Ce dispositif fournit un cadre opérationnel reproductible : il permet de fixer le pricing sur le coût marginal réel, d'adapter le staffing aux élasticités des centres de coûts, d'arbitrer les choix d'investissements en R&D ou en infrastructure sur la base de capacités d'investissement mesuré, et de fonder les stratégies de croissance sur une estimation précise des économies — ou déséconomies — d'échelle.

Mots clés : économie du logiciel ; coûts unitaires ; rendements d'échelle ; fonctions de coût ; régression Ridge ; Elastic-Net ; bootstrap blocs mobiles.

1 Introduction

Dans la littérature consacrée aux biens numériques, on souligne souvent que les **coûts marginaux** d'un éditeur de logiciels deviennent négligeables une fois le produit développé : la duplication du code ne requiert que des ressources de stockage ou de bande passante (Brynjolfsson, Hitt, 1997). Cette apparente gratuité de reproduction laisse entrevoir des rendements d'échelle substantiels. L'expérience des éditeurs montre pourtant une réalité plus nuancée : pour acquérir et fidéliser une clientèle, il faut financer marketing, force de vente, infrastructure d'assistance et fonctions support. À mesure que la base d'utilisateurs s'élargit et que les offres se diversifient, ces dépenses croissent et viennent contrebalancer les économies liées à la réplication du code.

Depuis une dizaine d'années, l'économie du logiciel a connu deux **transitions de modèle de commercialisation successives**. D'abord, les éditeurs ont ajouté des abonnements **SaaS** : le programme reste le même, mais il tourne sur leurs serveurs et l'utilisateur y accède via Internet, ce qui facilite les mises à jour et transforme l'achat ponctuel et l'incertitude du renouvellement, en paiement récurrent. Puis est venu le **Full Web** : l'application est réécrite pour fonctionner entièrement dans le navigateur, en micro-services et sans installation, afin d'offrir la même expérience sur ordinateur, tablette voire mobile. Chaque nouveau mode de commercialisation reconfigure la part de coûts fixes et variables des éditeurs ; mesurer l'incidence de ces bascules sur les économies d'échelle et la structure de coûts constitue dès lors un enjeu analytique et managérial majeur.

Limites des approches existantes

Dans l'industrie du logiciel, les rares travaux académiques consacrés aux coûts se concentrent presque exclusivement sur la phase de développement. Les modèles d'estimation de l'effort – de COCOMO (Boehm, 1981) à SLIM (Putnam, 1978) – décrivent la productivité initiale des équipes, mais laissent de côté les dépenses post-lancement, en particulier les coûts récurrents de maintenance, de support ou de mise à l'échelle de l'infrastructure. Les travaux récents d'Activity-Based Costing (ABC) appliqués aux suites logicielles montrent que la précision attendue devient vite impossible à exploiter dès que les données de traçabilité sont partielles (Kaplan & Anderson, 2004). De leur côté, les analyses de Total Cost of Ownership (TCO) menées après déploiement révèlent que le « coût marginal nul » théorique est neutralisé par l'augmentation des dépenses de support, des correctifs de sécurité et de la maintenance post-mise en production, lesquels peuvent absorber jusqu'à 70 % de la dépense totale (Baytech Consulting, 2020). Toutefois, ces études ne fournissent toujours pas de fonction de coût paramétrique intégrant simultanément toutes les directions opérationnelles, n'intègrent pas la dimension des modes de commercialisation et ne traitent pas de l'évolution des économies d'échelle. Dès lors, la littérature manque d'un cadre théorique capable de suivre l'évolution structurelle des coûts lorsqu'un éditeur passe successivement du *Desk*, au *SaaS*, puis au *Full web* et de répondre de façon certaine au mythe du « coût marginal nul ».

Vers une compréhension fine des économies d'échelle dans l'industrie du logiciel

L'étude poursuit trois objectifs complémentaires :

1. **Quantifier la structure globale des élasticités de coût** d'un éditeur multicanal, en tenant compte de la colinéarité naturelle entre directions opérationnelles ;
2. **Identifier les ruptures temporelles majeures** induites par l'essor du SaaS puis du Full web, afin de distinguer les évolutions conjoncturelles et les effets des changements de régime ;

3. **Évaluer, pour chacune des phases stratégiques isolées**, la robustesse des économies d'échelle et la sensibilité des postes de coût à l'augmentation du volume et aux transitions technologiques, en estimant un modèle log-log régularisé (Ridge) et une forme flexible Translog pénalisée (Elastic-Net).

L'entreprise étudiée est un éditeur de logiciels de gestion qui opère en France, compte 550 salariés et un portefeuille *B to B* mixte TPE/PE/PME. Les conclusions valent pour tout éditeur de logiciels (i) disposant d'une comptabilité analytique permettant de distinguer les coûts par direction, (ii) avec un mode de commercialisation unique ou dans une superposition multiple (Desk / SaaS / Full web) et (iii) facturant, ou non, une variété de produits ou services très hétérogènes. Les firmes "cloud native" pourraient présenter un profil de coûts différent ; nous y revenons en conclusion.

2 Revue de littérature

2.1 Du « coding effort » aux dépenses de cycle de vie

Les modèles d'estimation d'effort

Les tout premiers travaux économiques sur le logiciel se sont concentrés sur la phase de **développement**. Les modèles d'estimation d'effort – COCOMO (Boehm, 1981) ou SLIM (Putnam, 1978) – traduisent la taille du code ou la complexité fonctionnelle en *homme-mois*. Leur succès – ils restent aujourd'hui la référence des départements PMO – tient à leur simplicité d'implémentation : ils n'exigent, au fond, qu'un métrique de volume logiciel et quelques pondérations expertes. Cependant, cette focalisation sur l'effort « build » laisse dans l'ombre les autres composantes du coût total : l'infrastructure d'hébergement, les correctifs après mise en marché, l'assistance, le marketing et l'administration générale. Autrement dit, un état-de-l'art dominé par l'ingénierie logicielle ne peut, à lui seul, répondre à la question économique centrale : *combien coûte réellement la fourniture d'un service logiciel sur tout son cycle de vie ?*

Les chercheurs ont donc cherché, à partir des années 1990, à importer dans l'édition logicielle les fonctions de coût traditionnelles de l'économie industrielle, capables d'englober la totalité des ressources consommées.

Activity-Based Costing (ABC) et Total Cost of Ownership (TCO)

Des auteurs ont dépassé les simples modèles d'effort grâce à l'ABC : Ooi et Soh (2003) montrent que, en retraçant chaque activité de développement et de maintenance – correctifs, assistance et ressources cloud –, on peut reconstituer un TCO presque exhaustif. Cependant, ils soulignent aussi la lourdeur organisationnelle de la démarche : journaliser l'activité fine des équipes suppose des outils internes rarement disponibles. Surtout, l'ABC répond au *qui consomme quoi*, mais pas à la question macro : *comment le coût total réagit-il lorsque la production augmente de 10 % ?* Autrement dit, l'ABC éclaire la répartition mais pas la **courbure** de la fonction de coût, d'où la pertinence persistante des formes Cobb–Douglas ou Translog pour détecter les rendements d'échelle et expliquer les variations de coût.

2.2 Mesurer la production des services

Les premières difficultés, lorsqu'on cherche à estimer une fonction de coût dans les services, tiennent à la **définition même de l'output** et aux questions de **mesure des facteurs**

de production : c'est particulièrement vrai dans l'industrie du logiciel. Dans l'industrie manufacturière, la pièce produite, le mètre linéaire ou la tonne constituent des unités physiques évidentes ; rien de tel dans les services où le « produit » se confond avec une expérience, un droit d'usage ou une information. L'hétérogénéité des prestations et la coproduction avec le client empêchent d'aligner les comptabilités sur un dénominateur universel : un logiciel d'ERP très paramétré n'équivaut pas à un utilitaire de sauvegarde, pas plus qu'une visite guidée d'un musée n'équivaut à l'accès libre d'un après-midi dans ce même musée. Dans l'industrie du logiciel, la production se décline non seulement en tant que bien immatériel — le logiciel lui-même, livré sous forme de licences ou d'abonnements — mais aussi en service continu (maintenance, support technique, mises à jour, formations), le tout structuré selon des différenciations verticales (niveaux de gamme : de l'édition *standard* à l'offre *premium*) et des différenciations horizontales (modules complémentaires, options fonctionnelles, nombre d'utilisateurs, etc.).

Pour éclairer ce problème, il est utile de se tourner vers d'autres secteurs de services qui, eux aussi, peinent à mesurer leur production. Dans les **musées**, Jackson (1988) a retenu la fréquentation — le nombre de visiteurs — comme proxy unique : chaque entrée représente un acte de consommation culturelle et la palette relativement simple mais détaillée des postes de dépense (salaires, expositions, conservation, etc.) se prête à une Cobb–Douglas sans interactions majeures. Dans les **télécommunications**, Bagadeem (2020) s'est trouvé confronté à la coexistence de plusieurs filières (voix, data, services à valeur ajoutée) et a choisi un **double indicateur** chiffre d'affaires/abonnés. L'approche, modélisée par une Translog multiproduit - et des inputs agrégé : *capital*, *force de vente* et *autres* - saisit simultanément la valeur monétaire captée et la taille réelle du réseau d'utilisateurs. **Le transport aérien**, analysé par Gillen, Oum et Tretheway (1990), se situe à mi-chemin : les sorties sont immatérielles pour le passager, mais la combinaison « passagers-kilomètres » ou « tonnes-kilomètres » permet de reconstituer une unité quasi-physique et d'estimer une Translog incluant les interactions entre trafic passagers et fret avec des inputs agrégés également.

Ces trois études convergent vers un constat : **il n'existe pas de métrique unique transposable d'un service à l'autre**. Chaque secteur doit inventer un compromis entre observabilité comptable, pertinence économique et faisabilité statistique. On observe alors deux stratégies : soit réduire la complexité — un seul output (le visiteur) et une forme fonctionnelle simple (Cobb–Douglas) lorsque les postes de coût sont peu corrélés ; soit, à l'inverse, embrasser la pluralité des outputs et introduire une forme flexible (Translog) lorsqu'on veut rendre compte d'effets de synergie.

Dans l'édition logicielle, la mesure des coûts et de la performance doit, à la suite des travaux pionniers de Griliches (1979, 1986), tenir compte simultanément de la valeur des produits et du volume d'activité, afin de refléter l'hétérogénéité des offres et la variabilité des prix. À l'instar de Bagadeem (2020), nous combinons donc, dans un même modèle, un indicateur de valeur (chiffre d'affaires) et un indicateur de volume (nombre d'articles ou d'abonnés). Cette sélection bivalente s'inscrit dans la tradition des fonctions de coûts multi-outputs — TMCF, CES imbriquée, Generalized Leontief, etc. — qui montrent qu'il suffit d'une variable-clé par ligne de service pour tester les économies de portée sans exploser le nombre de paramètres.

La transition cloud native (Kubernetes puis serverless) renforce ce cadre : le coût d'hébergement suit désormais le volume d'usage et, dans un modèle SaaS, ce volume est corrélé au chiffre d'affaires récurrent. Jonas et al. (2019) rappellent que le serverless est facturé à la requête, tandis qu'Adzic & Chatley (2017) documentent une réduction de 66 à 95 % des coûts après migration vers AWS Lambda, précisément parce que la dépense épouse enfin la demande effective. Autrement dit, le couple unités servies – chiffre d'affaires capte l'essentiel de la dynamique de coûts observée dans un environnement cloud-native ; l'optimisation fine du code

joue, mais en second ordre. Cette évidence conforte donc la pertinence de notre modélisation fondée sur ce binôme.

2.3 Formes fonctionnelles

Pour ce qui est de la forme fonctionnelle à retenir, la transition des modes de commercialisation - amorcée depuis plus d'une décennie, du *Desk* vers le *SaaS*, puis vers le *Full web* - a déplacé le centre de gravité des coûts : la R&D et l'infrastructure cloud montent en charge, tandis que le support interne et le commerce évoluent avec la base installée et que le marketing s'intensifie aux périodes deancements : conduisant à la nécessité d'observer l'évolution des directions au fil du temps et les interactions lors des différentes phases. La Translog, enrichie de termes quadratiques et d'interactions, offre la souplesse indispensable pour capturer l'effet conjugué de la migration technologique et de la coexistence de plusieurs gammes. En la pénalisant via un Elastic-Net, on atténue simultanément le risque de multicollinéarité et l'on conserve, grâce au shrinkage, une interprétation économique des coefficients — un point crucial pour la prise de décision managériale. L'inconvénient étant le trop grand nombre de variables pour le trop petit nombre de données. Des formes plus simples devront être tester également.

C'est cette grille d'analyse que nous appliquerons dans la suite de l'article pour estimer, tour à tour, une Cobb–Douglas pénalisée et une Translog Elastic-Net, avant de réexaminer les économies d'échelle à la lumière de l'ajout successif du SaaS puis du Full web.

2.3.1 De la Cobb–Douglas à la translog

La forme Cobb–Douglas (Cobb & Douglas, 1928), pose que la variable à expliquer croît log-linéairement avec chaque input : un double-logarithme suffit pour mesurer les élasticités et statuer sur les rendements d'échelle. Son pouvoir évocateur – dans notre cas : un coefficient de 0,3 signifie qu'une hausse de 1% d'un facteur accroît le coût unitaire moyen de 0,3% – explique sa diffusion. Néanmoins, elle impose par construction une élasticité de substitution constante et ne prend pas en compte les interactions entre directions ; or, chez les éditeurs de logiciel : les coûts d'infrastructure, la force de vente, l'assistance et le marketing réagissent souvent *en chaîne* à l'évolution du parc installé et aux sorties de gammes ce qui induit une autocorrélation très forte et des effets de synergie. La forme Cobb–Douglas s'écrit, en logarithmes naturel :

$$\ln C = \ln \alpha_0 + \gamma_1 \ln articles + \gamma_2 \ln ca + \sum_{i \in \{RnD, comm, market, assis, aut\}} \alpha_i \ln D_i + \epsilon$$

où :

- α_0 est une constante représentant les coûts fixes,
- C est le coût unitaire moyen,
- γ_1, γ_2 et α_i sont les coefficients à estimer,
- *articles* et *ca* sont les outputs nombre d'articles vendus et chiffre d'affaires,
- D_i les directions : *RnD, commerciale, marketing, assistance technique* et *autres*.
- ϵ est le terme d'erreur aléatoire.

Pour pallier la rigidité inhérente à la forme Cobb–Douglas – en particulier l'hypothèse d'élasticités de substitution constantes et l'impossibilité de rendre compte d'effets d'interaction non linéaires entre inputs et outputs – Christensen, Jorgenson & Lau (1973) ont introduit

la forme Translog, qui se veut une approximation locale de toute fonction de coût deux fois différentiable. En autorisant explicitement, dans l'expression du logarithme du coût, des termes quadratiques et croisés entre inputs et outputs : cette spécification capte non seulement l'effet purement « taille » (ou « échelle ») de chaque variable, mais aussi la façon dont cet effet peut varier selon le niveau des autres variables. Autrement dit, elle permet d'estimer des élasticités de substitution variables, de modéliser la courbure de la fonction de coût et de tester plus finement les hypothèses d'homogénéité et de concavité.

La contrepartie de cette plus grande souplesse est un accroissement rapide du nombre de paramètres à calibrer – notamment δ_{kl} pour les interactions entre outputs, β_{ij} pour celles entre inputs, et θ_{ik} pour celles entre inputs et outputs – ce qui peut conduire, en présence de colinéarité forte et de jeux de données de taille limitée, à des variances estimées très élevées et à des coefficients peu stables. C'est ce phénomène que Gillen, Oum & Tretheway (1990) ont mis en évidence dans le cas du transport aérien, où l'on observe une instabilité des signes et de l'ampleur des effets second ordre selon les jeux de données et les méthodes d'estimation.

La forme Translog s'écrit alors :

$$\begin{aligned} \ln C = & \alpha_0 + \sum_{k=1}^2 \gamma_k \ln Y_k + \sum_{i \in \mathcal{D}} \alpha_i \ln D_i + \frac{1}{2} \sum_{k,l} \delta_{kl} \ln Y_k \ln Y_l \\ & + \frac{1}{2} \sum_{i,j} \beta_{ij} \ln D_i \ln D_j + \sum_{i,k} \theta_{ik} \ln D_i \ln Y_k + \epsilon \end{aligned}$$

avec $Y_1 = \text{articles}$, $Y_2 = \text{ca}$ et $\mathcal{D} = \{R\&D, \text{commerce}, \text{marketing}, \text{assistance}, \text{autres}\}$.

2.3.2 Les régressions pénalisées comme remède empirique à la multicolinéarité

La structure d'un **compte de résultat logiciel** crée d'emblée des corrélations fortes : une nouvelle version majeure déclenche (i) un surcroît de dépenses marketing pour la promotion, (ii) un renforcement du support et de l'assistance, (iii) une montée en charge des ressources cloud. Ces effets de synergies impliquent une autocorrélation forte entre les variables. Dans notre échantillon, le couple *commerce* – *articles* atteint un facteur d'inflation de la variance (VIF) de 42 ; au moindre changement d'intervalle temporel, le signe même de l'élasticité peut s'inverser, rendant toute interprétation micro-économique instable.

La **pénalisation statistique** offre une réponse élégante à ce problème : au lieu de supprimer des variables (procédures pas-à-pas classiques), on conserve l'information complète mais on contraint l'amplitude des coefficients, évitant ainsi l'« explosion » des estimateurs. Concrètement, la pénalisation agit comme un filtre : elle écarte la composante des régresseurs qui n'apporte qu'une variance redondante et privilégie la part réellement orthogonale, ce qui résout le paradoxe d'un R^2 élevé mais d'une faible transférabilité hors échantillon. Dans un environnement où les directions R&D, marketing, commerce et assistance évoluent souvent de concert, le Ridge puis l'Elastic-Net jouent donc le rôle de **réducteurs de colinéarité** : toutes les directions restent présentes dans le modèle, mais leurs élasticités se stabilisent autour de valeurs économiquement plausibles et statistiquement fiables.

Ridge : un filtre de Tikhonov pour stabiliser les élasticités

Proposée par Hoerl & Kennard (1970), la régression Ridge consiste à résoudre :

$$\min_{\beta} \|\mathbf{y} - \mathbf{X}\beta\|_2^2 + \lambda \|\beta\|_2^2,$$

où le paramètre λ contrôle le **rétrécissement des coefficients**. Il s’agit d’une forme de régularisation de Tikhonov, qui revient à imposer a priori que les coefficients ne doivent pas croître de manière excessive. Ce mécanisme est particulièrement utile pour stabiliser les estimateurs dans les contextes de colinéarité forte, comme c’est le cas ici entre les budgets *R&D*, *marketing*, *commerce* et *assistance technique*. La Ridge préserve toutes les variables, mais réduit l’influence des redondances structurelles. Elle permet ainsi de conserver l’intégralité de l’information opérationnelle, tout en limitant les variations erratiques dues au sur-ajustement.

LASSO : la parcimonie structurelle par annulation automatique

Tibshirani (1996) a introduit la pénalisation LASSO, fondée sur la norme L1 :

$$\min_{\beta} \|\mathbf{y} - \mathbf{X}\beta\|_2^2 + \lambda \|\beta\|_1.$$

Cette formulation favorise l’annulation de certains coefficients : la géométrie anguleuse du “ballon de contrainte” coupe des axes, mettant certains coefficients exactement à zéro. Ce comportement est très utile dans les sciences sociales pour extraire un noyau explicatif minimal. Toutefois, dans notre cas, le LASSO a tendance à éliminer des variables économiquement fondamentales comme *articles* (volume vendu), ce qui revient à mutiler le cœur même de la fonction de coût d’un éditeur. Le LASSO sert donc ici à tester la redondance potentielle entre directions ou effets croisés, mais n’est pas retenu comme modèle principal.

Elastic Net : arbitrage entre robustesse et parcimonie

Zou & Hastie (2005) ont proposé une combinaison des approches Ridge et LASSO :

$$\min_{\beta} \|\mathbf{y} - \mathbf{X}\beta\|_2^2 + \lambda \left[(1 - \alpha) \|\beta\|_2^2 + \alpha \|\beta\|_1 \right].$$

Ce mélange convexe permet d’exploiter les vertus des deux méthodes : la stabilité de la Ridge face à la multicolinéarité, et la capacité du LASSO à forcer des coefficients à zéro. Dans notre étude (section 5.4), nous fixons $\alpha = 0.5$, de manière à obtenir un compromis robuste : les variables structurelles sont conservées, tandis que les interactions superflues ou bruitées sont automatiquement éliminées. Ce choix s’avère particulièrement adapté à la spécification Translog, dont la richesse combinatoire nécessite un filtrage sans sacrifier l’interprétabilité.

Choix pratique de λ et validation croisée temporelle

Le réglage du paramètre de pénalisation λ nécessite un équilibre délicat : une valeur trop faible laisse persister la multicolinéarité, tandis qu’une valeur trop élevée peut estomper les relations économiques sous-jacentes. Pour nos données trimestrielles, nous adoptons une **validation croisée à fenêtre glissante**, une approche adaptée aux séries temporelles. Concrètement, à chaque itération, un bloc continu de quatre trimestres est exclu de l’échantillon d’entraînement pour servir de validation, garantissant ainsi que les données futures n’influencent pas l’estimation des modèles.

La sélection de λ s’appuie sur la règle du “**one-standard-error**”, telle que décrite par Hastie, Tibshirani et Friedman (2009) . Cette méthode consiste à choisir la valeur de λ la plus simple (c’est-à-dire la plus grande) dont l’erreur moyenne de validation est à moins d’un écart-type de l’erreur minimale observée. Cette approche favorise des modèles plus parcimonieux, réduisant ainsi la variance des estimateurs : cruciale lorsque l’on compare les phases *Desk*, *SaaS* et *Full Web*, chacune disposant d’environ 20 trimestres d’observations.

Portée méthodologique

En synthèse, la régularisation ne se limite pas à un « tour de main » statistique ; elle constitue une **condition de validité** pour estimer des fonctions de coût ayant des inputs corrélés, riches en interactions, sur des jeux de données typiques des entreprises (dans notre cas, $T=56$, avec des sous-périodes d'environ 20 trimestres). Sans elle, la mesure des rendements d'échelle se diluerait dans l'instabilité des coefficients. Avec elle, nous obtenons des élasticités interprétables, stables par validation croisée, et un socle robuste pour l'analyse dynamique des transitions *Desk* \rightarrow *SaaS* \rightarrow *Full web* qui suit dans l'article.

En résumé, la régularisation combine trois vertus : elle lisse la variance des coefficients corrélés, limite le risque de sur-ajustement en maîtrisant le compromis biais-variance, et stabilise l'estimation des termes quadratiques tout en préservant la souplesse de la fonction Translog. Ces propriétés, bien établies en économétrie pénalisée, se révèlent particulièrement utiles dans les configurations empiriques où le nombre d'observations est restreint par la segmentation temporelle ou stratégique. Enfin, la nomenclature retenue (outputs : volume + valeur ; inputs : décomposition fonctionnelle) est transposable à d'autres éditeurs de logiciels ou, plus largement, à tout service numérique distribué en mode Desk / SaaS / Full web.

3 Les données

L'étude repose sur une série chronologique de cinquante-six trimestres, allant du premier trimestre 2011 au quatrième trimestre 2024. Toutes les observations proviennent directement des systèmes comptables et de contrôle de gestion d'EBP Informatique, éditeur de logiciel de gestion français, ce qui garantit la cohérence entre les états financiers publiés et la base d'analyse. Chaque trimestre agrège, d'une part, la comptabilité générale — c'est-à-dire l'ensemble des charges directes et indirectes enregistrées au grand livre — et, d'autre part, les écritures analytiques qui ventilent ces charges par direction fonctionnelle. Les états de paie et les historiques des ressources humaines ont été raccordés à ces données financières ; ils fournissent pour chaque direction le nombre de collaborateurs, la masse salariale et les éventuelles variations de périmètre.

Les données ont été divisées en deux sous-ensembles. Les cinquante-deux trimestres compris allant de 2011-T1 à 2023-T4 forment l'échantillon d'estimation. Les quatre trimestres de 2024, qui n'ont pas servi à la mise au point des modèles, constituent un échantillon-test indépendant ; il servira à mesurer la précision hors échantillon (MAPE, RMSE, indicateur de précision globale).

La profondeur temporelle de la base, jointe à la ventilation fine des postes comptables et à l'homogénéité des procédures de collecte depuis plus d'une décennie, fournit ainsi un socle empirique solide pour tester les différentes formes fonctionnelles et les méthodes d'imputation présentées dans les sections qui suivent.

Les extractions ont été conduites sous NDA interne ; les agrégats trimestriels sont anonymisés et chiffrés pour publication. Aucun identifiant client ni salaire individuel n'apparaît dans les jeux publics fournis aux reviewers.

4 Construction du modèle généralisé

L'objectif de cette section est de développer un modèle généralisé pour analyser la structure des coûts d'un éditeur de logiciel selon ses différentes directions opérationnelles, vérifier la présence ou l'absence d'économies d'échelle, et prédire ces coûts de façon fiable. Trois approches

économétriques ont été mises en concurrence : un modèle de type Cobb-Douglas, et deux variantes de régularisation, Ridge et Lasso. D’un point de vue économique, les questions portent sur (1) la somme des coefficients pour évaluer les rendements d’échelle, (2) le poids de chaque direction dans la structure des coûts, et (3) la qualité des prévisions sur la période de test.

4.1 Variables et périmètre

Comme discuté dans les sections précédentes, la mise au point d’une fonction de coût adaptée au secteur du logiciel se heurte à la fois à la diversité des postes de dépenses et à la pluralité des formes de production. Pour mieux tenir compte de cette réalité, nous proposons un modèle généralisé où la variable à expliquer est le **coût unitaire moyen** – soit le coût total divisé par le nombre d’articles vendus.

Dans la pratique, un simple décompte des ventes (*‘articles’*) ne rend pas compte de l’hétérogénéité des offres, une option “ajout d’utilisateur” pouvant valoir quelques euros là où un “logiciel PME” se facture plusieurs milliers. Inversement, le chiffre d’affaires (*‘ca’*) reflète la disparité de valeur, mais masque la dynamique purement quantitative. En combinant ces deux indicateurs, on capture ainsi les deux dimensions fondamentales de l’activité.

Du côté des **inputs**, on retient une décomposition par direction, chacune étant normalisée par le nombre d’articles vendus : les directions R&D (*‘RnD’*), Marketing (*‘marketing’*), Commerce (*‘commerce’*), Assistance technique (*‘AT’*), et les autres directions (RH, secrétariat général, direction générale, BU en *‘Autres’*). Pour chaque direction seront inclus : les salaires, les frais généraux et frais divers par employés et les coûts de sous-traitance de chaque direction. Ce degré de finesse s’inspire des pratiques de contrôle de gestion, où les coûts de fonctionnement sont réaffectés selon la proportion d’employés dans chaque direction, offrant une vision plus granulaire du lien entre un produit vendu et la ressource qu’il a mobilisée. Le détail précis des comptes rattachés à chaque direction, ainsi que les règles d’allocation des charges communes (loyer, cloud, frais divers), figure dans l’Annexe A afin de faciliter la répliquabilité.

Nous travaillons systématiquement en logarithmes naturels. Cette transformation garantit (i) l’homogénéité d’ordre 0 de la fonction —chaque coefficient s’interprète comme une élasticité constante—, (ii) la réduction de l’asymétrie des distributions de coûts, et (iii) l’absorption mécanique des effets de niveau. Toutes les composantes de coût sont également divisées par le nombre d’articles vendus, de sorte qu’elles mesurent un coût moyen par article vendu : l’intercept n’agrége plus que les fragments fixes et variables non imputables à une direction ni aux employés à savoir : les coûts Saas (serveurs, hébergement et bandes passantes), les redevances ainsi que les impôts et taxes.

4.2 Estimations

Test de stationnarité

Avant d’estimer des modèles dynamiques, il faut s’assurer que les séries temporelles ne contiennent pas de racine unitaire : faute de stationnarité, les relations estimées risquent d’être « spurielles » et les statistiques asymptotiques (tests t, F, etc.) deviennent non valides. Nous appliquons donc trois procédures complémentaires : le test Augmented Dickey-Fuller (Dickey & Fuller 1979), le test de Phillips-Perron (Phillips & Perron 1988) et le test KPSS qui, à l’inverse, prend la stationnarité pour hypothèse nulle (Kwiatkowski et al. 1992). Les résultats détaillés en Annexe B montrent que, après passage en logarithme puis différenciation d’ordre 1, toutes les séries sont intégrées d’ordre 0. Nous pouvons donc estimer l’ensemble des

modèles présentés ci-après (Cobb-Douglas log-log, régularisations Ridge et Lasso) en utilisant ces variables sans craindre la présence de racines unitaires résiduelles.

Résultats et discussion

Table 1: Résultats des modèles sur les données test

| Cobb-Douglas | | | Régressions pénalisés | |
|---------------|---------|---------|-----------------------|--------|
| Variabes | Coeff | Signif. | Ridge | Lasso |
| (Intercept) | 1.033 | . | 2.752 | 1.748 |
| ca | -0.0164 | | 0.001 | -0.008 |
| articles | 0.056 | | -0.074 | . |
| RnD | 0.427 | *** | 0.383 | 0.423 |
| marketing | 0.089 | *** | 0.112 | 0.089 |
| commerce | 0.237 | *** | 0.107 | 0.186 |
| Autres | 0.137 | *** | 0.126 | 0.131 |
| AT | 0.159 | *** | 0.161 | 0.159 |
| Var expliquée | 99.26 | | 98.96 | 99.32 |
| RMSE | | | 0.0217 | 0.1132 |

La table 1 présente, pour chacune des méthodes, les coefficients estimés sur la période d'apprentissage 2011-2023. Le modèle Cobb-Douglas affiche une variance expliquée très élevée (99 %) et des coefficients *commerce*, *RnD*, *Autres* et *AT* fortement significatifs ($p < 10^{-6}$). Cependant, les outputs (*articles* et *ca*) ne sont pas significatifs, et surtout les valeurs de ces deux coefficients s'inversent de signe par rapport à l'intuition : un accroissement du volume aurait un effet positif sur le coût unitaire, ce qui contredit l'idée même d'économies d'échelle.

Le Ridge conserve l'intégralité des variables et compresse légèrement leurs amplitudes ; la somme des élasticités ($\approx 0,96$) confirme un régime de rendements décroissants tout en stabilisant les estimations (RMSE = 0,022 sur l'échantillon d'apprentissage). Le Lasso, en revanche, met à zéro la variable *articles*. Pour un éditeur logiciel, ignorer le pur volume de ventes revient à négliger un déterminant essentiel de la dilution des frais fixes ; cette exclusion fragilise l'interprétation économique, même si le critère deviance (99,3 %) reste excellent.

La question de la multicollinéarité est éclairée par la table 2. Les VIF dépassent 40 pour *articles* et *commerce* ; ce niveau est nettement supérieur aux seuils usuels entre 5 et 10 (Gujarati & Porter, 2009) et explique la faible robustesse de leurs t-statistiques dans le Cobb-Douglas. En pratique, les deux indicateurs mesurent des phénomènes corrélés : le nombre d'articles vendus et l'activité commerciale évoluent quasi simultanément, d'où une instabilité que la pénalisation L2 du Ridge corrige, sans pour autant évincer des variables potentiellement structurantes.

Table 2: VIF modèle Cobb-Douglas

| ca | articles | RnD | marketing | commerce | Autres | AT |
|-----|----------|------|-----------|----------|--------|------|
| 4.4 | 47.8 | 3.01 | 2.45 | 44.25 | 2.03 | 3.12 |

Performances hors échantillon

La robustesse prédictive est évaluée sur les quatre trimestres 2024 laissés totalement hors estimation. Pour évaluer la performance hors-échantillon de nos trois modèles nous avons retenu deux indicateurs complémentaires sur les quatre trimestres de 2024 :

- **MAPE** (Mean Absolute Percentage Error), qui mesure l'écart moyen absolu en pourcentage entre les valeurs prédites \hat{y}_i et les valeurs réelles y_i :

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right| \times 100.$$

Un MAPE faible traduit une meilleure capacité du modèle à anticiper précisément l'ordre de grandeur des coûts.

- **Biais agrégé** (*Total Bias*) — écart signé, en %, entre la somme des prévisions et la somme observée :

$$\text{Précision}_{\text{global}} = \frac{\sum_i \hat{y}_i - \sum_i y_i}{\sum_i y_i} \times 100.$$

Un résultat < 0 indique une sous-estimation globale, > 0 une sur-estimation. Plus la valeur est proche de 0, en absolu, plus le modèle est précis.

Table 3: Performances sur les données test

| Modèle | MAPE | Biais_agrégé |
|--------------|------|--------------|
| Cobb-Douglas | 1.79 | -0.25 |
| Ridge | 2.40 | -1.70 |
| Lasso | 1.86 | -0.59 |

Ces résultats illustrent un trade-off classique : la parcimonie du Lasso séduit en théorie, mais devient discutable dès qu'une variable économiquement centrale (ici *articles*) est supprimée ; le Cobb-Douglas non régularisé exalte la performance statistique, mais succombe à la multicollinéarité et se révèle instable hors échantillon ; enfin, le Ridge offre un compromis robuste, limitant le gonflement des variances, conservant toutes les variables et délivrant une trajectoire de prévision fiable. Dans la suite de l'article, nous retiendrons donc la spécification Ridge comme référence opérationnelle ; les autres modèles serviront de point de comparaison pour évaluer la sensibilité des conclusions aux choix méthodologiques.

Diagnostics modèle Ridge

Nous appliquons deux diagnostics classiques avant de projeter le modèle sur l'année de test :

1. **Homoscédasticité** : Le test de Breusch & Pagan (1979) compare la variance conditionnelle des résidus à une variance constante. La statistique $\chi^2_7 = 5,21$ ($p = 0,63$) ne permet pas de rejeter l'hypothèse nulle d'homoscédasticité ; la dispersion des résidus du Ridge n'augmente donc pas systématiquement avec le niveau des prédictors.
2. **Autocorrélation** : Le test de Ljung & Box (1978) agrège les autocorrélations jusqu'au retard 4 (soit un an). Avec $Q = 18,11$ pour $df = 4$ ($p = 0,0012$), l'hypothèse nulle d'indépendance sérielle est rejetée : les résidus présentent une dynamique temporelle non capturée.

En pratique, l'absence d'hétéroscédasticité simplifie l'interprétation des élasticités (les variances ne dépendent pas du niveau d'activité), mais la présence d'autocorrélation invite à **prudence sur l'inférence**. Les coefficients restent convergents, mais leurs erreurs-types OLS sont biaisées : nous reportons, pour chaque coefficient, l'écart-type Newey–West (lag = 4) et la p-valeur associée;

5 Robustesse des estimations

Bien que le modèle Ridge (Table 1) offre une bonne performance prédictive moyenne et la meilleure lisibilité économique, il est estimé sur l'ensemble de la période, ce qui pourrait masquer (ou lisser) des modifications profondes de la relation entre coût unitaire et facteurs explicatifs et leur évolution au fil du temps. Afin de vérifier la solidité de nos conclusions, nous combinons trois familles de tests : (i) détection formelle de ruptures, (ii) ré-estimation sur sous-périodes homogènes, (iii) suivi dynamique glissant des rendements d'échelle.

5.1 Tests de rupture temporelle

Les résidus sont corrigés de l'hétéroscédasticité et de l'autocorrélation au moyen de la variance HAC de Newey-West. Le paramètre de troncature est fixé à **lag** = 4, choix couramment adopté pour une fréquence trimestrielle (\approx un an) et conforme à la règle automatique de Newey & West (1994) $L_{NW} = 4(N/100)^{2/9}$. Nous fixons le seuil de décision à 5 % pour le test de Chow et à 1 % pour le test de Wald global, ce dernier étant plus puissant lorsqu'on teste simultanément plusieurs restrictions.

- **Test de Chow**

À une date candidate t^* , on estime :

1. le modèle sur l'ensemble $t = 1, \dots, T$ (RSS_{glob}) ;
2. deux sous-modèles, avant et après t^* (RSS_1, RSS_2).

La statistique

$$F_{\text{Chow}} = \frac{(RSS_{\text{glob}} - RSS_1 - RSS_2)/(k+1)}{(RSS_1 + RSS_2)/(n_1 + n_2 - 2(k+1))}$$

suit une loi $F_{k+1, n_1+n_2-2(k+1)}$ sous l'hypothèse nulle d'égalité des paramètres (Chow, 1960), où $k+1$ est le nombre total de coefficients estimés, intercept compris. Nous appliquons le test à chaque trimestre compris entre 2013-T1 et 2022-T4 ; les segments 2011-T1 \rightarrow 2012-T4 et 2021-T4 \rightarrow 2023-T4 servent de bords de fenêtre. Les p -valeurs sont calculées avec la variance HAC (lag = 4) pour tenir compte des diagnostics d'autocorrélation présentés plus haut.

- **Test de Wald :**

On compare un modèle complet (tous paramètres libres) à un modèle restreint défini sous H_0 . La statistique :

$$F_{\text{Wald}} = (\hat{\beta} - \beta_0)^\top [\text{Var}_{\text{HAC}}(\hat{\beta})]^{-1} (\hat{\beta} - \beta_0)$$

suit asymptotiquement une loi χ_q^2 (ou $F_{q, n-q}$) sous H_0 , où q est le nombre de contraintes testées (Wald, 1943). Ce test permet d'évaluer **conjointement** l'ensemble des égalités « avant / après », là où un test t ne porterait que sur un seul coefficient à la fois.

Résultats et discussion

Le graphique trace l'évolution de la statistique F_{Chow} calculée à **chaque trimestre candidat** entre 2013-T1 et 2021-T3 : chaque point correspond à un test « avant / après t^* ». On observe une rupture modeste en 2015-T3, un palier plus net autour de 2017-T4, puis une envolée marquée à partir de 2019-T2, signe que la phase de réécriture Full web amplifie fortement la divergence entre sous-échantillons. Chaque test repose sur moins de 25 observations dans chacun des segments scindés. Dans un tel effectif, la distribution asymptotique $F_{k+1, n_1+n_2-2(k+1)}$ tend à sous-estimer la probabilité d'erreur de type I ; les p -valeurs associées doivent donc être lues comme indicatives plutôt que décisives.

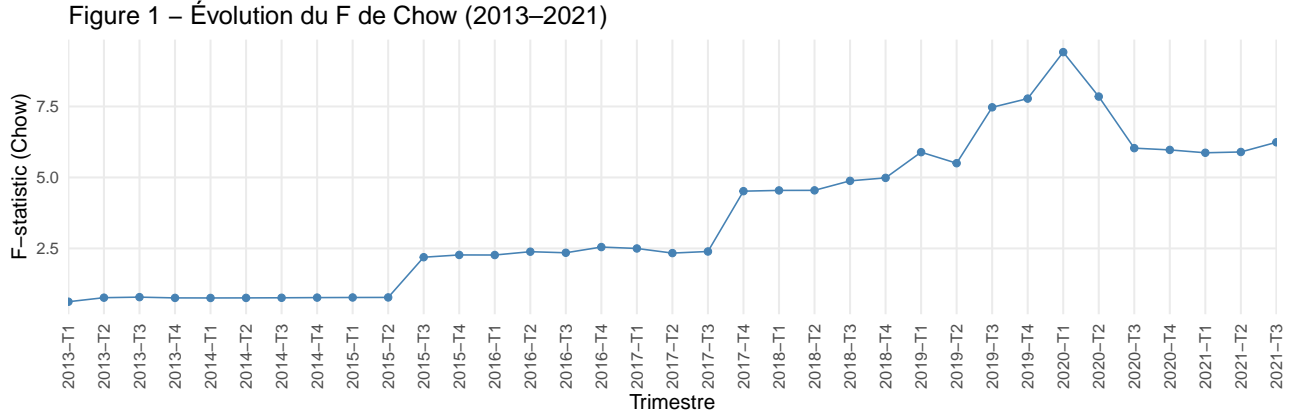


Table 4: Tests de rupture Chow et Wald aux points d'inflexion

| Période | Chow F | Chow p | Wald F | Wald p | Décision (Chow) |
|---------|--------|--------|---------|--------|-------------------|
| 2015-T3 | 2.6971 | 0.0196 | 18.3148 | 0 | Rupture confirmée |
| 2017-T4 | 5.6864 | 0.0001 | 21.5722 | 0 | Rupture confirmée |
| 2020-T1 | 9.8698 | 0.0000 | 23.6781 | 0 | Rupture confirmée |

Note : le test de Wald peut être appliqué à tout autre trimestre, mais, pour rester concis, nous l'avons limité aux trois maxima identifiés du F_{Chow} .

Le tracé de l'évolution du F_{Chow} (figure 1) et les résultats de la table 4—tous deux fondés sur une variance HAC de Newey–West (lag = 4)—mettent en évidence **trois ruptures nettes** : **2015-T3**, **2017-T4** et **2020-T1**.

Ces dates correspondent précisément aux jalons technico-commerciaux de l'éditeur (voir frise en Annexe C) : le premier pic (2015-T3) coïncide avec la première percée du SaaS et le réajustement organisationnel qu'elle a imposé, le second (2017-T4) suit la généralisation du modèle d'abonnement hébergé, et le troisième (2020-T1) marque le lancement du chantier **Full Web**, avec 3 % du budget R&D la première année – part qui grimpe à 27 % en 2024.

Afin d'étudier les économies d'échelle sans fragmenter l'échantillon, nous retenons ces trois fenêtres, qui allient cohérence stratégique et solidité économétrique :

1. **Phase Desk** (2011-T1 → 2015-T2) : offre 100 % on-premise.
2. **Phase Desk + SaaS** (2015-T3 → 2019-T4) : transition vers le cloud, incluant la rupture de 2017-T4 pour préserver la taille d'échantillon.
3. **Phase Desk + SaaS + Full web** (2020-T1 → 2024-T4) : mise en œuvre de l'architecture Full web dans un contexte d'accélération SaaS et de réorganisation post-Covid.

Ces trois segments servent de base aux estimations qui suivent, garantissant à la fois une correspondance avec le calendrier des investissements et la robustesse des ruptures détectées.

5.2 Réestimation des élasticités par sous-périodes

La table 5 ci-dessous compare les coefficients Ridge log-log des trois phases *Desk*, *Desk + SaaS* et *Desk + SaaS + Full web* ; chaque coefficient est assorti de son écart-type robuste Newey-West (SE *HAC*, lag = 4). La table 6 regroupe les métriques d'ajustement clés – somme des élasticités Σ_β , part de variance expliquée et taille d'échantillon.

Pour juger de la significativité en effectif réduit (18 – 20 trimestres) et sous autocorrélation, nous complétons l'inférence par un **bootstrap à blocs mobiles** : 1000 répliques de blocs contigus de quatre trimestres respectent la dépendance temporelle et fournissent une distribution empirique plus fiable que l'approximation asymptotique (HAC). Cette approche suit les recommandations de Künsch (1989) et Lahiri (2003) sur l'estimation de variance en série courte. Les *SE_boot* et *p_boot* obtenus confirment les conclusions tirées des écarts-type robuste ; leur détail figure en annexe D afin de ne pas surcharger le texte principal.

Ensemble, les deux tables permettent de suivre l'évolution des élasticités quand l'éditeur passe du Desk au SaaS puis au Full Web, et de vérifier que la qualité d'ajustement reste élevée malgré la segmentation temporelle.

Rendements d'échelle : trois régimes distincts

Table 5: Coefficients Ridge avec SE HAC

| | Desk | | Desk + SaaS | | Desk + SaaS + Full web | |
|-------------|--------|--------|-------------|--------|------------------------|--------|
| | Coef | SE_HAC | Coef | SE_HAC | Coef | SE_HAC |
| (Intercept) | 3.551 | 0.245 | 4.029 | 1.270 | 1.708 | 1.444 |
| ca | 0.003 | 0.002 | -0.001 | 0.008 | -0.022 | 0.038 |
| articles | -0.136 | 0.016 | -0.154 | 0.087 | 0.006 | 0.147 |
| RnD | 0.205 | 0.025 | 0.267 | 0.015 | 0.463 | 0.022 |
| marketing | 0.111 | 0.007 | 0.133 | 0.010 | 0.118 | 0.012 |
| commerce | 0.232 | 0.023 | 0.139 | 0.080 | 0.177 | 0.061 |
| Autres | 0.141 | 0.009 | 0.048 | 0.005 | 0.147 | 0.023 |
| AT | 0.157 | 0.006 | 0.195 | 0.013 | 0.151 | 0.078 |

Table 6: Métriques des estimations par sous-périodes

| Metric | Desk | Desk_SaaS | Full_Web |
|----------------|--------|-----------|----------|
| Σ_β | 0.712 | 0.626 | 1.039 |
| %Var | 99.798 | 99.675 | 98.324 |
| N | 18.000 | 18.000 | 20.000 |

Les écarts-types HAC montrent que, pour les trois périodes, les coefficients clés sont estimés avec une précision confortable : le rapport $|\beta|/\text{SE}$ dépasse 8 pour *RnD* et *Marketing* en phases *Desk* et *Desk + SaaS*, et reste supérieur à 4 en phase Full Web, ce qui fonde la significativité indiquée par les p-values (non reproduites ici, voir annexe E).

L'examen des coefficients met aussi en lumière la **redistribution progressive du poids des directions**. En phase Desk, le coût unitaire est largement tiré par *Commerce* ($\beta \approx 0,23$) et, dans une moindre mesure, par *RnD* (0,21) et *Marketing* (0,11). Lorsque le SaaS se généralise, *RnD* prend la tête (0,27) tandis que *Commerce* se replie (0,14) ; les investissements techniques

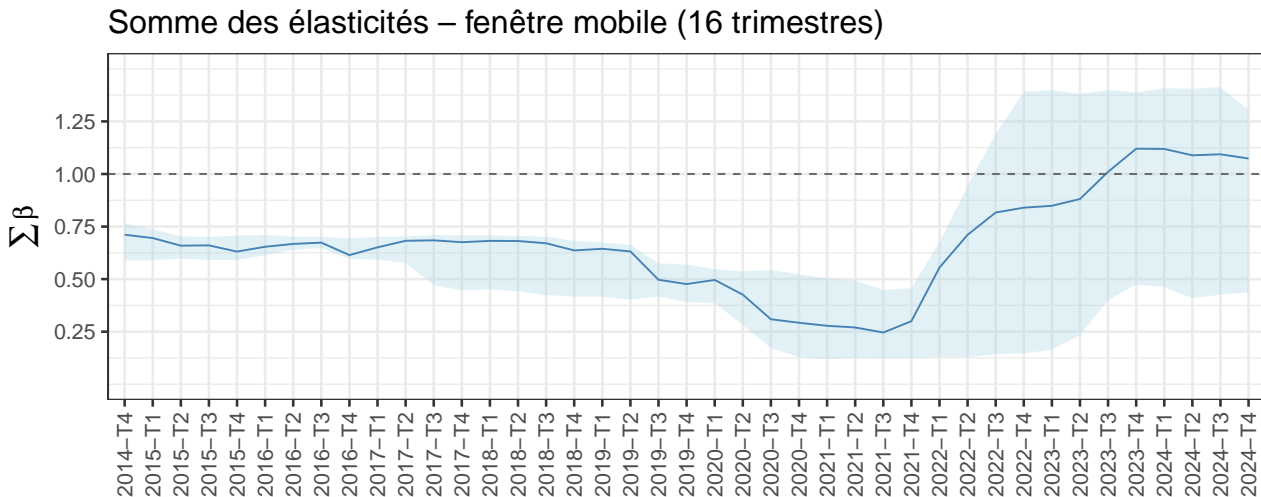
liés à l'hébergement cloud commencent déjà à dominer la formation du coût. Avec l'arrivée du *Full Web*, la hiérarchie se renverse nettement : l'élasticité *RnD* double pratiquement (0,46) et celle de *Commerce* se redresse (0,18), révélant l'effort conjoint de ré-écriture et de relance commerciale. Dans le même temps, l'effet volume (*articles*) devient nul et l'élasticité *AT* frôle la significativité, preuve que le support doit absorber une complexité accrue. En clair, les investissements massifs en Full Web déplacent la charge financière vers la R&D et ravivent le rôle du commerce, tandis que les économies d'échelle liées au volume s'épuisent.

Du côté des métriques globales (table 6), la somme des élasticités Σ_β tombe de 0,71 à 0,63 quand le SaaS s'installe – signe d'économies d'échelle renforcées – puis rebondit à **1,04** avec le passage Full Web, indiquant des déséconomies. Malgré ces variations, la **variance expliquée** reste très élevée ($\approx 98\text{--}100\%$), preuve que la forme log-log capte l'essentiel de la dynamique de coûts même sur des segments courts ($N = 18, 18$ et 20).

En résumé, les erreurs-types HAC confirment la robustesse des coefficients, tandis que Σ_β et %Var révèlent un basculement net : économies d'échelle jusqu'en 2019, puis déséconomies dès que la R&D bascule vers le *Full Web*.

5.3 Évolution dynamique des rendements d'échelle

Pour dépasser les « photos fixes » des trois sous-périodes, nous suivons la **trajectoire glissante** des rendements d'échelle. Chaque trimestre t nous ré-estimons le modèle Ridge log-log sur la **fenêtre mobile de 16 trimestres** (soit 4 ans) précédent à t . Cette largeur garantit : (i) un échantillon suffisant pour la validation croisée, (ii) une lecture moyen-terme qui amortit les à-coups saisonniers. L'incertitude est mesurée par un **moving-block bootstrap** (1000 répliques, blocs = 4 trimestres) : pour chaque fenêtre on obtient la distribution empirique de Σ_β et on trace un intervalle de confiance à 95 %.



Entre 2014-T4 et 2021-T3, la somme glissante des élasticités (Σ_β) reste durablement en-dessous de 0,75 : l'entreprise évolue dans un **régime d'économies d'échelle**, renforcé par la montée progressive du SaaS (phase 1 et 2 de la section précédente). Le point bas de 0,25 atteint à 2021-T3 traduit le moment où le volume SaaS dilue le plus fortement les charges fixes.

À partir de 2022-T1, la courbe change brutalement de pente : Σ_β remonte à 0,80 en quatre trimestres, puis franchit le seuil de 1 à 2023-T3. Sur cette zone de transition, la bande de confiance s'élargit car la fenêtre mobile (16 trimestres) mélange encore des observations *Desk*

+ *SaaS* avec les premiers trimestres *Full Web* ; malgré cette dispersion, **tout l'intervalle à 95 % reste au-dessus de 1 après 2023-T3**. C'est le passage au **troisième régime**, celui des **déséconomies d'échelle** : les investissements massifs en R&D, le renforcement du support et le run cloud croissent désormais plus vite que le volume servi, l'idéal de « coût marginal nul » disparaît.

En 2024 la somme retombe légèrement vers 1,05, mais le ruban supérieur reste bien au-dessus de l'unité ; la firme ne retrouve donc pas les économies d'échelle antérieures. Autrement dit, après une période d'accélération de la dilution (Desk \rightarrow SaaS), l'arrivée du *Full Web* inverse la trajectoire : trois plates-formes coexistantes font passer Σ_β au-delà de 1 et placent l'entreprise dans un régime temporaire de rendements décroissants.

5.4 Forme flexible : Elastic-Net Translog pénalisée

Les régressions linéaires de la section 5.1 posaient l'hypothèse forte de rendements marginaux constants et ne permettaient pas de tester les synergies ou frictions entre directions. Pour détecter d'éventuels **seuils de saturation** (termes quadratiques) et **économies ou déséconomies de gamme** (interactions croisées), nous adoptons ici la fonction de coût translog — un développement de second ordre en logarithmes (Christensen, Jorgenson & Lau 1973) qui, localement, n'impose ni élasticités ni rendements fixes.

Spécification et sélection

Chaque input apparaît en niveau, au carré et, lorsque la littérature ou l'observation opérationnelle le justifie, dans des produits croisés. Compte tenu du faible effectif (18 – 20 trimestres par phase), nous ne retenons que trois interactions économiquement pertinentes :

- **RnD \times AT** – coût conjoint des correctifs et du support,
- **Marketing \times Commerce** – couplage lead-to-cash,
- **Commerce \times AT** – coordination front office / support client.

Le modèle est estimé par **Elastic-Net** ($\alpha = 0,50$), qui combine la robustesse à la colinéarité de la Ridge (L2) et la parcimonie du LASSO (L1). Le paramètre de pénalisation est choisi par validation croisée glissante : un bloc contigu de quatre trimestres est omis à chaque pli, respectant ainsi la dépendance temporelle ; la règle *one-standard-error* retient la solution la plus sobre. Pour pallier l'échantillon réduit et l'autocorrélation résiduelle, nous calculons enfin des intervalles de confiance par **bootstrap à blocs mobiles** (1 000 répliques, bloc = 4), conformément aux recommandations de Künsch (1989) et Lahiri (2003). Les résultats bootstrap confirment la significativité discutée ci-dessous ; ils figurent en annexe E.

Table 7: Coefficients Translog Elastic-Net par phase stratégique

| | Desk | Desk_SaaS | Full_Web |
|-------------|--------|-----------|----------|
| (Intercept) | 3.777 | 4.338 | 2.491 |
| ca | 0.000 | 0.005 | 0.000 |
| articles | -0.089 | -0.116 | 0.000 |
| RnD | 0.112 | 0.109 | 0.162 |
| marketing | 0.000 | 0.000 | 0.015 |
| commerce | 0.129 | 0.000 | 0.069 |

| | | | |
|--------------|--------|--------|--------|
| Autres | 0.066 | 0.000 | 0.040 |
| AT | 0.050 | 0.038 | 0.000 |
| RnD_sq | 0.010 | 0.025 | 0.022 |
| marketing_sq | 0.000 | 0.001 | 0.006 |
| commerce_sq | 0.014 | 0.000 | 0.016 |
| Autres_sq | 0.007 | 0.016 | 0.012 |
| AT_sq | 0.002 | 0.011 | 0.000 |
| mk_com | 0.023 | 0.029 | 0.021 |
| AT_com | 0.010 | 0.027 | 0.000 |
| RnD_AT | 0.012 | 0.008 | 0.058 |
| %Var | 99.750 | 99.900 | 98.630 |
| RMSE | 0.012 | 0.005 | 0.012 |

Les coefficients quadratiques confirment un **renforcement de la convexité** :

- le terme $\gamma_{\text{RnD},\text{RnD}}$ passe de 0,010 (Desk) à 0,022 (Full Web), indiquant une saturation plus rapide des équipes de développement ;
- marketing et commerce restent faiblement convexes, la montée du Web ne change pas leur profil de productivité marginale.

Les interactions révèlent l'évolution des **frictions organisationnelles** (toutes **positives**) :

- **marketing** \times **commerce** reste positif et sensiblement constant (0,02–0,03) ; la coordination des campagnes et du front-office génère un coût additionnel plutôt qu'une synergie.
- **commerce** \times **AT** est visible avant 2020 (0,010 \rightarrow 0,042) puis disparaît : la bascule vers un portail self-service réduit la charge conjointe vente–support.
- **RnD** \times **AT** explose de 0,012 à **0,062** ; l'entretien simultané de micro-services multiplie les tickets techniques et alourdit le coût partagé Développement–Support.

L'élasticité volume (*articles*) passe de $-0,089$ à $-0,059$ puis à zéro : les économies d'échelle pures s'éteignent, mais c'est bien la **hausse des frictions positives** qui propulse la somme des élasticités au-delà de l'unité.

Le passage au Full Web transforme donc la nature des rendements : il ne détruit pas uniquement l'effet volume, il renforce la convexité R&D et fait émerger une friction majeure avec le support. En parallèle, la seule synergie durable – marketing \times commerce – s'affaiblit. L'entreprise ne retrouvera un régime d'économies d'échelle qu'en réduisant la saturation R&D (automatisation pipeline) et la friction R & D–Support (outillage DevSecOps), tout en capitalisant sur les synergies commerciales restantes.

6 Conclusion

Suivre pendant quatorze ans la trajectoire d'un éditeur français — du modèle *Desk* initial à l'introduction du *SaaS*, puis à la mise en production d'une architecture *Full Web* — montre que la promesse d'un « coût marginal nul » n'a rien d'automatique. Les régressions Ridge mettent d'abord en évidence un régime d'économies d'échelle solide : tant que les offres Desk

et SaaS cohabitent, la somme des élasticités Σ_β oscille entre 0,25 et 0,70 et la dilution des charges fixes domine. Cette situation bascule dès qu'un seuil d'investissement dans la réécriture *Full Web* est franchi : Σ_β dépasse alors l'unité et l'entreprise entre temporairement dans un régime de rendements décroissants. La Translog pénalisée confirme ce diagnostic : l'effort R&D, conjugué à l'assistance technique, ainsi que la persistance des frictions *marketing* \times *commerce* neutralisent désormais les gains de volume.

Une économie d'échelle, même marquée, peut donc être annulée par une vague d'investissements trop lourds : lorsque la dépense R&D croît plus vite que le volume servi, les coûts unitaires remontent. La gouvernance doit donc ramener chaque projet à un **seuil d'équilibre** où l'économie d'échelle attendue compense exactement l'effort financier ; au-delà, les investissements — pour la refonte *Full Web* dans notre cas — accroissent le coût marginal.

Pour piloter ce compromis, cinq leviers apparaissent. (1) **Calibrer le budget R&D sur Σ_β** : si les projections placent Σ_β au-dessus de 1, différer ou phaser le chantier jusqu'à absorption de la dette technique ; (2) **Aligner les objectifs commerciaux sur la capacité d'investissement** : enrichir le produit à hauteur du volume prévisible — et non l'inverse ; (3) **Mettre un garde-fou de trésorerie** : plafonner l'investissement annuel au gain d'échelle réalisé l'année précédente, sous peine de financement externe ou de réduction de périmètre ; (4) **Surveiller les effets de saturation** : les termes quadratiques positifs de la Translog signalent le seuil au-delà duquel la productivité marginale décroît ; tout projet dépassant cette cap-table doit passer en revue de gouvernance ; (5) **Intégrer les interactions de coût dans les KPI** : le suivi en temps réel des interactions entre les directions permet de détecter la dérive avant qu'elle n'affecte le coût total.

Ainsi, la question n'est plus de savoir si le numérique offre un « coût marginal nul », mais dans quelles **conditions organisationnelles et financières** cette promesse tient : équilibre entre capacité d'investissement et gain d'échelle, visibilité sur les coûts croisés et pilotage dynamique des efforts techniques.

L'étude reste limitée par la taille des segments (18 à 20 trimestres) et la confidentialité des données, qui empêche une diffusion intégrale du jeu d'origine ; les intervalles bootstrap atténuent le biais sans l'annuler totalement. La portée externe devrait être testée sur un panel multi-entreprises, en intégrant la dynamique des prix et de la concurrence, ainsi qu'une analyse produit par produit. Il n'en demeure pas moins que l'industrie du logiciel connaît de fortes économies d'échelle, qu'un passage au SaaS peut renforcer mais qu'un investissement excessif — tel qu'une réécriture *Full Web* — peut entièrement annuler. Le *Full Web* ne supprime pas les coûts fixes ; il en déplace la frontière et en crée de nouveaux : la qualité de l'architecture logicielle et de la coordination inter-directions, plutôt que la simple croissance du volume, devient désormais la clé pour réaliser des économies d'échelle durables.

7 Références

- Brynjolfsson, E. & Hitt, L.M. (1997). “Information Technology and Internal Firm Organization: An Exploratory Analysis.” *Journal of Management Information Systems*, 14(2), 81–101.
- Boehm, B. W. (1981). *Software Engineering Economics*. Englewood Cliffs, NJ: Prentice-Hall.
- Putnam, L. H. (1978). A general empirical solution to the macro software sizing and estimating problem. *IEEE Transactions on Software Engineering*, 4(4), 345–361.
- Robert S. Kaplan, Steven R. Anderson. (2004). *Time-Driven Activity-Based Costing*. Harvard Business Review, 82(11), 131–138.
- Baytech Consulting. (2020). *Projecting Costs in Software Maintenance*
- Ooi, G., & Soh, C. (2003). “Developing an Activity-Based Costing Approach for System Development and Implementation”. ACM SIGMIS Database, 34(3), 54-71.
- Adzic, G., & Chatley, R. (2017). Serverless computing: Economic and architectural impact. In *Proceedings of the 11th Joint Meeting on Foundations of Software Engineering (ESEC/FSE '17)* (pp. 411-414).
- Bagadeem, S. (2020). Estimating the translog cost function for the telecommunications sector in Saudi Arabia: The case of Saudi Telecom Company. *International Journal of Financial Research*, 12(3), 19-30.
- Griliches, Z. (1979). Issues in assessing the contribution of research and development to productivity growth. *The Bell Journal of Economics*, 10(1), 92-116.
- Griliches, Z. (1986). Productivity, R&D, and basic research at the firm level in the 1970s. *American Economic Review*, 76(1), 141-154.
- Jonas, E., Schleier-Smith, J., Sreekanti, V., Tsai, C.-C., Khandelwal, A., Pu, Q., ... Patterson, D. A. (2019). *Cloud programming simplified: A Berkeley view on serverless computing* (Tech. Rep. No. UCB/EECS-2019-3). University of California, Berkeley.
- Jackson, R. (1988). “A Museum Cost Function.” *Journal of Cultural Economics*, 12(1), 41–50.
- David W. Gillen, Tae Hoon Oum and Michael W. Tretheway : “Airline Cost Structure and Policy Implications: A Multi-Product Approach for Canadian Airlines”, *Journal of Transport Economics and Policy*, Vol. 24, No. 1 (Jan., 1990), pp. 9-34
- Cobb, C. & Douglas, P.H. (1928). “A Theory of Production.” *The American Economic Review*, 18(1), 139–165.
- Christensen, L.R., Jorgenson, D.W. et Lau, L.J. (1973). “Transcendental Logarithmic Production Frontiers.” *The Review of Economics and Statistics*, 55(1), 28–45
- Hoerl, A.E. & Kennard, R.W. (1970). “Ridge Regression: Biased Estimation for Nonorthogonal Problems.” *Technometrics*, 12(1), 55–67.
- Tibshirani, R. (1996). “Regression Shrinkage and Selection via the Lasso.” *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1), 267–288.

- Zou, H. & Hastie, T. (2005). “Regularization and Variable Selection via the Elastic Net.” *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67(2), 301–320.
- Friedman, J., Hastie, T. & Tibshirani, R. (2010). “Regularization Paths for Generalized Linear Models via Coordinate Descent.” *Journal of Statistical Software*, 33(1), 1–22.
- Gujarati, D.N. and Porter, D.C. (2009) Basic Econometrics. 5th Edition, McGraw Hill Inc., New York.
- Hastie, T., Tibshirani, R. & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer.
- Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74(366), 427–431.
- Phillips, P. C. B., & Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika*, 75(2), 335–346.
- Kwiatkowski, D., Phillips, P. C. B., Schmidt, P., & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of Econometrics*, 54(1-3), 159–178.
- Breusch, T. S., & Pagan, A. R. (1979). A simple test for heteroscedasticity and random coefficient variation. *Econometrica*, 47(5), 1287–1294.
- Ljung, G. M., & Box, G. E. P. (1978). On a measure of lack of fit in time series models. *Biometrika*, 65(2), 297–303.
- Andrews, D. W. K. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica*, 59(3), 817–858.
- Newey, W. K., & West, K. D. (1994). Automatic lag selection in covariance matrix estimation. *The Review of Economic Studies*, 61(4), 631–653.
- Chow, G. C. (1960). Tests of equality between sets of coefficients in two linear regressions. *Econometrica*, 28(3), 591–605.
- Wald, A. (1943). Tests of statistical hypotheses concerning several parameters when the number of observations is large. *Transactions of the American Mathematical Society*, 54(3), 426–482.
- Künsch, H. R. (1989). The jackknife and the bootstrap for general stationary observations. *Annals of Statistics*, 17(3), 1217–1241.
- Lahiri, S. N. (2003). *Resampling Methods for Dependent Data*. Springer Series in Statistics.

8 Annexe

8.1 A. Allocation des coûts de fonctionnement par direction

1. Périmètre Les coûts de fonctionnement comprennent :

- Frais généraux (ex. abonnements, énergie, assurances, entretien, transports...)
- Frais divers (ex. achats divers, commissions, packaging, documentation...)
- Voyages & déplacements (billets, cartes carburant, péages...)
- Frais marketing (publicité, relations presse, goodies, PLV...)
- Sous-traitance (prestations externes, honoraires, formation...)
- Logiciel interne (frais de développement, maintenance)
- Coûts SaaS/Web & redevances

2. Méthodologie d'imputation

- Frais marketing : imputés intégralement à la Direction Marketing.
- Sous-traitance : imputée directement à la direction utilisatrice de la prestation.
- Frais généraux, frais divers, voyages et déplacements : répartis au prorata du nombre d'employés de chaque direction, selon la formule :

$$\text{Coût alloué à } D_i = \frac{\text{Nbr employés } D_i}{\text{Nbr total employés}} \times \text{Coût total}$$

- Coûts SaaS/Web & redevances : considérés comme bruit de modèle ; non répartis dans le cadre de ce périmètre.

8.2 B. Tests de stationnarité des variables log-différenciées

Pour chaque variable, trois tests complémentaires ont été appliqués :

- **ADF** (Dickey-Fuller augmenté) — H_0 : racine unitaire.
- **PP** (Phillips-Perron, statistique $Z-\alpha$) — Si $\text{ratio } |\text{stat}| / |\text{critère } 5 \text{ \%}| < 1,15 \rightarrow \text{stationnarité}$.
- **KPSS** — H_0 : stationnarité autour d'une constante.

Les séries couvrent la période complète de l'étude et sont transformées en log puis différenciées d'ordre 1 afin de préserver leur dimension économique (taux de croissance). Une constante est incluse dans les régressions de tests et le nombre de retards est choisi selon le critère de Schwartz.

| serie | variable | ADF_stat | ADF_p | KPSS_stat | KPSS_p | PP_ratio |
|--------|---------------|----------|-------|-----------|--------|----------|
| niveau | cout_unitaire | -2.520 | 0.365 | 0.135 | 0.100 | 0.791 |
| niveau | ca | -2.167 | 0.507 | 1.469 | 0.010 | 0.071 |
| niveau | articles | -3.190 | 0.098 | 1.469 | 0.010 | 0.125 |
| niveau | RnD | -1.823 | 0.646 | 0.793 | 0.010 | 0.647 |

| | | | | | | |
|--------|-----------|--------|-------|-------|-------|-------|
| niveau | marketing | -2.629 | 0.321 | 0.789 | 0.010 | 0.513 |
| niveau | commerce | -2.917 | 0.205 | 1.383 | 0.010 | 0.113 |
| niveau | Autres | -1.528 | 0.765 | 0.207 | 0.100 | 0.384 |
| niveau | AT | -0.617 | 0.972 | 0.425 | 0.067 | 0.199 |

Pour les huit séries en niveau, les tests concordent globalement :

- **ADF** ne rejette jamais l'hypothèse de racine unitaire (p-value > 5 %).
- **KPSS** rejette la stationnarité pour toutes, sauf peut-être *cout_unitaire* et *Autres* où la statistique reste tout juste au seuil de 10 %.
- **PP** donne des ratios inférieurs à 1,15, mais l'ADF et le KPSS, plus rigoureux ici, pointent vers la non-stationnarité.

8.3 C. Frise temporelle

Table 9: Frise chronologique (valeurs en pourcent)

| Periode | Part du CA total | | | Part dans le budget R&D | |
|---------|------------------|-------|---------|-------------------------|--------|
| | OnPremise | SaaS | Fullweb | Fullweb | Autres |
| 2011 | 100 | — | — | — | 100 |
| - | 100 | — | — | — | 100 |
| 2014 | 98 | 2 | — | — | 100 |
| - | - | - | — | — | 100 |
| 2017 | 92 | 8 | — | — | 100 |
| 2018 | 89 | 11 | — | — | 100 |
| 2019 | 84 | 16 | — | — | 100 |
| 2020 | 78 | 22 | — | 3 | 97 |
| 2021 | 70 | 30 | — | 9 | 91 |
| 2022 | 64 | 35.98 | 0.02 | 15 | 85 |
| 2023 | 58 | 41.66 | 0.34 | 21 | 79 |
| 2024 | 57 | 42 | 1 | 27 | 73 |

8.4 D. Bootstrap par périodes

Table 10: Bootstrap par périodes

| | Desk | | | Desk + Saas | | | Desk + Saas + Full web | | |
|-------------|--------|---------|--------|-------------|---------|--------|------------------------|---------|--------|
| | Coef. | SE_boot | p-boot | Coef. | SE_boot | p-boot | Coef. | SE_boot | p-boot |
| (Intercept) | 3.555 | 0.000 | 0.244 | 0.243 | 1.692 | 1.056 | 0.000 | 1.454 | 0.026 |
| ca | 0.002 | 0.350 | 0.602 | 0.002 | -0.020 | 0.025 | 0.168 | 0.038 | 0.798 |
| articles | -0.136 | 0.000 | 0.985 | 0.016 | 0.003 | 0.058 | 0.000 | 0.147 | 1.308 |
| RnD | 0.205 | 0.000 | 0.000 | 0.025 | 0.464 | 0.028 | 0.000 | 0.022 | 0.000 |
| marketing | 0.111 | 0.000 | 0.000 | 0.007 | 0.117 | 0.016 | 0.000 | 0.012 | 0.000 |
| commerce | 0.232 | 0.000 | 0.003 | 0.023 | 0.179 | 0.053 | 0.000 | 0.061 | 0.000 |
| Autres | 0.141 | 0.044 | 0.000 | 0.009 | 0.146 | 0.028 | 0.000 | 0.022 | 0.002 |

| | | | | | | | | | |
|----|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| AT | 0.157 | 0.000 | 0.061 | 0.006 | 0.152 | 0.054 | 0.000 | 0.081 | 0.006 |
|----|-------|-------|-------|-------|-------|-------|-------|-------|-------|

8.5 E. Forme flexible : Elastic-Net Translog avec toutes les interactions

Table 11: Coefficients Translog Elastic-Net (variables, carrés et interactions)

| | Desk | Desk_SaaS | Full_Web |
|--------------------|--------|-----------|----------|
| (Intercept) | 3.716 | 3.229 | 2.986 |
| articles | -0.044 | 0.000 | 0.000 |
| RnD | 0.082 | 0.076 | 0.011 |
| commerce | 0.103 | 0.000 | 0.000 |
| Autres | 0.000 | 0.000 | 0.000 |
| AT | 0.040 | 0.000 | 0.000 |
| articles_sq | -0.002 | 0.000 | 0.000 |
| RnD_sq | 0.007 | 0.010 | 0.002 |
| commerce_sq | 0.010 | 0.000 | 0.000 |
| AT_sq | 0.000 | 0.000 | 0.000 |
| RnD_marketing | 0.002 | 0.008 | 0.023 |
| RnD_commerce | 0.010 | 0.035 | 0.071 |
| RnD_Autres | 0.012 | 0.000 | 0.029 |
| RnD_AT | 0.009 | 0.006 | 0.044 |
| marketing_commerce | 0.010 | 0.009 | 0.000 |
| marketing_Autres | 0.013 | 0.000 | 0.001 |
| marketing_AT | 0.005 | 0.008 | 0.000 |
| commerce_Autres | 0.010 | 0.000 | 0.000 |
| commerce_AT | 0.009 | 0.029 | 0.000 |
| Autres_AT | 0.005 | 0.021 | 0.000 |
| %Var | 99.740 | 98.870 | 98.350 |
| RMSE | 0.013 | 0.016 | 0.013 |

8.6 F. Bootstrap translog

Table 12: Translog – Phase Desk : écarts-types et p-values bootstrap

| | Coefficient | SE_boot | p_boot |
|-------------|-------------|---------|--------|
| (Intercept) | 3.7731 | 0.2565 | 0.000 |
| ca | 0.0000 | 0.0000 | 2.000 |
| articles | -0.0403 | 0.0142 | 0.000 |
| RnD | 0.0706 | 0.0116 | 0.000 |
| marketing | 0.0000 | 0.0031 | 1.954 |
| commerce | 0.0913 | 0.0106 | 0.000 |

| | | | |
|--------------------|---------|--------|-------|
| Autres | 0.0000 | 0.0172 | 1.004 |
| AT | 0.0342 | 0.0123 | 0.000 |
| ca_sq | 0.0000 | 0.0000 | 2.000 |
| articles_sq | -0.0022 | 0.0007 | 0.000 |
| RnD_sq | 0.0081 | 0.0015 | 0.000 |
| marketing_sq | 0.0000 | 0.0000 | 2.000 |
| commerce_sq | 0.0103 | 0.0011 | 0.000 |
| Autres_sq | 0.0000 | 0.0014 | 1.486 |
| AT_sq | 0.0007 | 0.0032 | 0.394 |
| RnD_marketing | 0.0028 | 0.0028 | 0.078 |
| RnD_commerce | 0.0103 | 0.0011 | 0.000 |
| RnD_Autres | 0.0120 | 0.0018 | 0.000 |
| RnD_AT | 0.0094 | 0.0015 | 0.000 |
| marketing_commerce | 0.0089 | 0.0023 | 0.002 |
| marketing_Autres | 0.0123 | 0.0042 | 0.032 |
| marketing_AT | 0.0053 | 0.0036 | 0.314 |
| commerce_Autres | 0.0099 | 0.0015 | 0.000 |
| commerce_AT | 0.0094 | 0.0017 | 0.000 |
| Autres_AT | 0.0040 | 0.0025 | 0.010 |

Table 13: Translog – Phase Desk + SaaS : écarts-types et p-values bootstrap

| | Coefficient | SE_boot | p_boot |
|--------------------|-------------|---------|--------|
| (Intercept) | 3.0992 | 0.7496 | 0.000 |
| ca | 0.0000 | 0.0243 | 1.974 |
| articles | 0.0000 | 0.0097 | 2.000 |
| RnD | 0.0903 | 0.0329 | 0.032 |
| marketing | 0.0000 | 0.0002 | 1.982 |
| commerce | 0.0000 | 0.0082 | 1.768 |
| Autres | 0.0000 | 0.0000 | 2.000 |
| AT | 0.0037 | 0.0164 | 1.394 |
| ca_sq | 0.0000 | 0.0007 | 1.990 |
| articles_sq | 0.0000 | 0.0004 | 2.000 |
| RnD_sq | 0.0098 | 0.0047 | 0.038 |
| marketing_sq | 0.0000 | 0.0002 | 1.960 |
| commerce_sq | 0.0000 | 0.0012 | 1.792 |
| Autres_sq | 0.0000 | 0.0000 | 2.000 |
| AT_sq | 0.0000 | 0.0023 | 1.544 |
| RnD_marketing | 0.0081 | 0.0046 | 0.102 |
| RnD_commerce | 0.0361 | 0.0076 | 0.000 |
| RnD_Autres | 0.0000 | 0.0026 | 1.644 |
| RnD_AT | 0.0063 | 0.0093 | 0.298 |
| marketing_commerce | 0.0123 | 0.0038 | 0.014 |
| marketing_Autres | 0.0000 | 0.0008 | 1.906 |
| marketing_AT | 0.0062 | 0.0036 | 0.042 |

| | | | |
|-----------------|--------|--------|-------|
| commerce_Autres | 0.0000 | 0.0011 | 1.874 |
| commerce_AT | 0.0312 | 0.0075 | 0.000 |
| Autres_AT | 0.0228 | 0.0069 | 0.032 |

Table 14: Translog – Phase Desk + SaaS : écarts-types et p-values bootstrap

| | Coefficient | SE_boot | p_boot |
|--------------------|-------------|---------|--------|
| (Intercept) | 2.9859 | 0.3411 | 0.000 |
| ca | 0.0000 | 0.0075 | 2.000 |
| articles | 0.0000 | 0.0120 | 2.000 |
| RnD | 0.0115 | 0.0317 | 0.184 |
| marketing | 0.0000 | 0.0021 | 1.802 |
| commerce | 0.0000 | 0.0113 | 1.700 |
| Autres | 0.0000 | 0.0001 | 1.996 |
| AT | 0.0000 | 0.0049 | 1.932 |
| ca_sq | 0.0000 | 0.0002 | 2.000 |
| articles_sq | 0.0000 | 0.0005 | 2.000 |
| RnD_sq | 0.0015 | 0.0045 | 0.130 |
| marketing_sq | 0.0000 | 0.0007 | 1.766 |
| commerce_sq | 0.0000 | 0.0024 | 1.668 |
| Autres_sq | 0.0000 | 0.0001 | 1.992 |
| AT_sq | 0.0000 | 0.0009 | 1.918 |
| RnD_marketing | 0.0229 | 0.0048 | 0.000 |
| RnD_commerce | 0.0707 | 0.0106 | 0.000 |
| RnD_Autres | 0.0286 | 0.0071 | 0.012 |
| RnD_AT | 0.0441 | 0.0079 | 0.000 |
| marketing_commerce | 0.0000 | 0.0022 | 1.244 |
| marketing_Autres | 0.0014 | 0.0047 | 0.536 |
| marketing_AT | 0.0000 | 0.0012 | 1.626 |
| commerce_Autres | 0.0000 | 0.0037 | 1.444 |
| commerce_AT | 0.0000 | 0.0018 | 1.820 |
| Autres_AT | 0.0000 | 0.0023 | 1.504 |