

Louvain School of Management

Le BEL ESG Index constitue-t-il un signal fiable de durabilité ? Une analyse de la performance ESG réelle et du discours des entreprises belges.

Auteur : Bouillon Paul
Promotrice(s) : D'Hondt Catherine
Année académique 2025-2026
Travail de fin d'études (TFE) en vue d'obtenir le titre de
Ingénieur de gestion à finalité spécialisée - International Finance
Horaire de jour

Déclaration IA

During the preparation of this master's thesis, the author(s) utilized Chat GPT for the following purpose:

- 1) Reformulation de textes
- 2) Explication de certains concepts méthodologiques
- 3) Traduction d'articles scientifiques

After using Chat GPT, the author(s) diligently reviewed and edited the content produced by the tool. We take full responsibility for the final content presented in this thesis.

By signing this declaration, we affirm that the content of this master's thesis reflects our original work, augmented by the responsible use of AI.

Remerciements

Ce travail est l'aboutissement de quelques mois de travail acharné mais il marque aussi le début d'un nouveau chapitre. Ces années d'études n'auraient pas eu la même saveur sans certaines personnes. Je tiens à remercier mes parents pour leur soutien, ma sœur pour sa relecture ainsi que toutes les personnes avec lesquelles j'ai pu échanger sur le campus durant ces six années. Un grand merci à ma promotrice qui a toujours su répondre dans les temps et a été de bon conseil durant cette année.

Merci à l'UCLouvain pour ces magnifiques années.

Résumé

Ce mémoire examine la cohérence du BEL ESG Index, lancé par Euronext en 2023, comme signal de durabilité fiable pour les investisseurs. Directement inspiré des travaux de Kolling et Busch (2025), notre étude ne s'applique pas aux fonds labellisés FNG mais bien à un indice d'actions. Ce travail teste si l'appartenance au BEL ESG reflète une performance ESG réelle supérieure, une intensité carbone inférieure, une communication ESG plus intense dans les rapports annuels et enfin s'il existe une éventuelle sur-communication relative à la performance. L'échantillon comprend 40 entreprises belges dont 20 membres du BEL ESG et 20 entreprises de contrôle appariées sur le secteur et la taille. Ces 40 entreprises sont observées en panel sur 3 années (2022 à 2024) soit 120 observations entreprise-année. Les données de performances ont été collectées via Bloomberg et l'intensité du discours ESG est mesurée à l'aide d'une analyse textuelle des rapports annuels. Toutes les hypothèses sont testées à l'aide de régressions MCO en panel groupé avec trois spécifications progressives. Le premier modèle est estimé sans variable de contrôle, le deuxième intègre une variable catégorielle d'années (année 2022 comme référence) et le troisième modèle ajoute une variable catégorielle de secteur (secteur financier comme référence) ainsi qu'une variable de contrôle continue (taille).

Les résultats montrent que les niveaux plus élevés d'ESGC pour les entreprises du BEL ESG sont attribués à leur profil structurel plutôt qu'à leur appartenance à l'indice. De plus, la différence d'intensité carbone entre les deux groupes n'est pas statistiquement significative une fois les caractéristiques sectorielles prises en compte. L'intensité du discours ESG est légèrement plus élevée mais significatif qu'à 10%. Et enfin, en contrôlant la performance ESG réelle, les entreprises du BEL ESG ne communiquent pas significativement davantage que leurs pairs.

Ce mémoire contribue à la littérature sur la cohérence des labels ESG et combine une analyse de la performance réelle des entreprises avec leurs discours. Le tout appliqué à un indice d'actions plutôt qu'à des fonds d'investissements.

Table des matières

Table des matières.....	5
1. Introduction	7
2. Revue de littérature.....	8
2.1. Le greenwashing : définitions et formes.....	8
2.1.1. Définitions et formes principales	8
2.1.2. Les sept péchés du greenwashing	9
2.2. Pourquoi les entreprises ont-elles recours au greenwashing ? Les fondements théoriques	10
2.2.1. La théorie de la signalisation : le label comme signal.....	10
2.2.2. La théorie de la légitimité : la pression sociale comme incitant à surenchérir	11
2.2.3. La théorie de l'agence : des conflits d'intérêt qui favorisent le camouflage	11
2.3. Les facteurs déterminants du greenwashing	12
2.3.1. La taille de l'entreprise.....	12
2.3.2. Le secteur d'activité	13
2.3.3. La pression réglementaire : incitant ou frein au greenwashing ?.....	13
2.4. Les conséquences financières et les risques associés au greenwashing	14
2.4.1. Réaction des marchés et risque de krach boursier	14
2.4.2. Le paradoxe de la réputation.....	14
2.5. Labels, indices ESG et mesure du greenwashing.....	14
2.5.1. L'étude de référence : Kolling et Busch (2025)	15
2.5.2. Mesurer l'écart : les approches pour quantifier le greenwashing	16
2.6. Limites de la littérature et positionnement de ce mémoire.....	17
3. Développement des hypothèses	18
3.1. Les entreprises appartenant au BEL ESG Index présentent une performance ESG réelle supérieure aux entreprises de contrôle non incluses dans l'indice	18
3.2. Les entreprises du BEL ESG Index présentent une intensité de discours ESG plus élevée que les entreprises n'appartenant pas à l'indice.....	19
3.3. À performance ESG comparable, les entreprises appartenant au BEL ESG Index présentent une intensité de discours ESG plus élevée que les entreprises comparables non incluses dans l'indice	19
4. Méthodologie	20
4.1. Sources de données.....	21
4.2. Sélection de l'échantillon	22
4.3. Construction des variables	24
4.3.1. Variable d'intérêt : appartenance au BEL ESG Index	25
4.3.2. Variable de discours : ESG Linguistic Score (ESG_LING)	25

4.3.3.	Variable de performance (1) : ESG Combined Score (ESGC).....	26
4.3.4.	Variable de performance (2) : Intensité carbone GHG Scope 1 et 2 (GHG_INT).....	27
4.3.5.	Indice de performance composite (PERF_ESG).....	27
4.4.	Statistiques descriptives.....	29
4.4.1.	Année de référence 2022 (pré-indice)	29
4.4.2.	Période d'analyse (2023-2024)	30
4.5.	Modèles de régression	31
4.5.1.	Spécifications des modèles	31
5.	Résultats.....	32
5.1.	Test de H1a – ESG Combined Score (ESGC).....	32
5.2.	Test de H1b – Intensité carbone GHG_INT.....	34
5.3.	Test de H2 – Intensité du discours ESG (ESG_LING).....	35
5.4.	Test de H3 – Communication excessive (modèle conditionnel).....	36
5.5.	Synthèse des résultats	38
6.	Discussion des résultats	39
7.	Limites de l'étude.....	40
8.	Conclusion	40
	Annexes.....	42
	Annexe A – Codage dynamique BEL_ESG	42
	Annexe B – Script extraction pdf en fichier texte.....	43
	Annexe C – Script analyse textuelle	48
	Annexes D – Script statistiques descriptives	52
	Annexe E – Script régressions OLS.....	55
	Annexe F – Tableaux de régressions détaillés	60
	Bibliographie.....	64

1. Introduction

A l'heure actuelle, la finance durable joue un rôle déterminant dans la transition écologique mondiale. Son objectif est clair : orienter les flux de capitaux vers des entreprises plus responsables. Pour ce faire, la finance durable intègre des critères de durabilité dans les décisions d'investissements ([European Commission, 2023](#)).

Face à la demande croissante des investisseurs, des instruments de mesure de la durabilité ont vu le jour comme les labels et les indices ESG. C'est dans cette lignée qu'Euronext a lancé, en 2023, l'indice BEL ESG. Cet indice a pour ambition de déterminer les entreprises reprises dans le BEL 20 et le BEL MID qui affichent les meilleures pratiques en termes de performances environnementales, sociales et de gouvernances ([Euronext, 2023](#)).

Cependant, le marché est inondé de signaux et la crédibilité de ceux-ci doit être remise en doute. Selon Molnár et al. (2023, cité dans [Ruggeri et al., 2025](#)), les rapports de durabilité sont souvent des auto-déclarations non auditées et cette liberté de divulgation peut inciter certaines entreprises à adopter des stratégies de « Greenwashing ».

En 2025, Kolling et Busch décident de travailler sur le label FNG. A travers leurs travaux, ils démontrent que si les fonds labellisés ont des scores de durabilité supérieurs, les différences sur les scores ESG fournis par les agences sont moins prononcées.

Ce mémoire vise à examiner la cohérence du BEL ESG Index en s'inspirant de la méthodologie de [Kolling et Busch \(2025\)](#) afin de comparer chaque entreprise reprise dans l'index avec une entreprise de contrôle non incluse.

Au-delà des scores ESG généralement utilisés, nous intégrerons une métrique textuelle. Celle-ci viendra mesurer l'intensité du discours ESG dans les rapports annuels des entreprises de l'échantillon. L'objectif est de déterminer si l'appartenance au BEL ESG Index est associée à une performance durable réelle supérieure, à une intensité de discours ESG plus élevée, et si ce discours excède ce que la performance observable justifie — ce qui constituerait un indicateur potentiel de greenwashing

2. Revue de littérature

La finance durable s'est considérablement développée au cours de la dernière décennie, imposant une question fondamentale aux marchés : peut-on faire confiance aux signaux ESG émis par les entreprises ? Afin de répondre à cette question, la revue de littérature qui suit est structurée en quatre grandes étapes qui expliquent pas à pas le raisonnement de ce mémoire. Dans un premier temps, nous définissons le greenwashing ainsi que ces diverses manifestations. Dans un deuxième temps, nous expliquerons trois cadres théoriques qui expliquent pourquoi les entreprises peuvent avoir recours à une communication trompeuse sur leurs performances ESG. Ensuite, nous analyserons les différentes conséquences financières entraînées par le greenwashing et nous finirons par nous interroger sur l'efficacité des labels et indices ESG comme réponse à ce problème.

2.1. Le greenwashing : définitions et formes

En 1986, le terme greenwashing est utilisé pour la première fois par l'écologiste Jay Westervelt. Alors qu'il était à l'hôtel, celui-ci se rend compte que les gérants demandaient aux clients de réutiliser leurs serviettes sous prétexte de protéger l'environnement alors que leur unique objectif était de réduire leurs coûts opérationnels (Pearson, 2010, cité dans [M. Xu et al., 2025](#)). Depuis, le concept est passé d'une simple combine marketing à une préoccupation financière majeure. De nombreux articles scientifiques sur le sujet sont sortis durant ces cinq dernières années avec la montée en puissance de la finance durable et des critères ESG. Aujourd'hui, le greenwashing est perçu comme un risque systémique pouvant entraîner une asymétrie d'information, une méfiance des investisseurs et une instabilité des marchés financiers ([G. Liu et al., 2024](#)).

2.1.1. Définitions et formes principales

Dans la littérature financière, plusieurs définitions existent pour illustrer le concept de greenwashing mais toutes partagent un point commun : un écart délibéré entre ce que l'entreprise communique et ce qu'elle fait réellement. Lyon and Montgomery (2015, cité dans [M. Ruggeri et al., 2025](#)) l'explique comme une communication qui a pour unique but d'embellir la situation d'une organisation. L'objectif est de laisser penser aux parties prenantes que les

performances environnementales de l'organisation sont supérieures à la réalité. [M. Ruggeri et al. \(2025\)](#) l'expliquent par une incohérence « Paroles vs Actes ». Une stratégie où les organisations communiquent à l'aide de discours verts (green talk) uniquement pour satisfaire les parties prenantes, mais sans actions concrètes (green walk). Là où [G. Liu et al. \(2024\)](#) la définissent comme un camouflage de performance : les entreprises divulguent de fausses informations ESG pour cacher des performances durables médiocres.

Dans leurs récents travaux, [M. Xu et al. \(2025\)](#) identifient les deux formes principales de greenwashing. La première, appelée la divulgation sélective (selective disclosure), que l'on observe lorsque l'entreprise communique sur des informations environnementales positives tout en cachant délibérément les impacts négatifs de ses activités. La seconde est le découplage (decoupling) qui survient lorsque l'on observe une différence entre les engagements publics (rapports annuels, publicités) et les pratiques opérationnelles réelles de l'entreprise.

2.1.2. Les sept péchés du greenwashing

L'étude de TerraChoice, effectuée en 2010, est la référence pour identifier et classer les différentes communications trompeuses utilisées par les entreprises. Parmi ces pratiques, on retrouve le « *compromis caché* », qui consiste à définir un produit comme écologique sur base d'un attribut positif tout en omettant de parler de ses attributs négatifs. Une autre pratique fréquente est *l'absence de preuve*, lorsque des affirmations environnementales ne sont pas clairement justifiées. *L'utilisation de termes vagues* comme « écologique » ou encore « naturels » peut également tromper les utilisateurs en raison de leur manque de connaissance. D'autre part, certaines affirmations peuvent être jugées *non pertinentes* lorsqu'elles mettent en avant des caractéristiques déjà imposées par la réglementation. Une autre pratique souvent utilisée par les entreprises se nomme le « *moindre mal* » qui consiste à vendre un produit comme étant plus durable dans une catégorie de produits nuisibles.

Enfin, les cas les plus graves restent la *désinformation directe* ou encore *l'utilisation de faux labels* qui donnent l'illusion qu'une tierce personne a validé l'entreprise ou le produit.

Ce premier chapitre pose une question cruciale : pourquoi certaines entreprises prennent-elles le risque de greenwasher ? La réponse se trouve dans les différents fondements théoriques que nous aborderons dans la partie suivante.

2.2. Pourquoi les entreprises ont-elles recours au greenwashing ? Les fondements théoriques

Dans cette partie, nous utiliserons trois grands cadres théoriques afin d'expliquer pourquoi les entreprises ont parfois intérêt à communiquer de manière trompeuse sur leurs performances ESG. Ces théories se complètent et, ensemble, elles construisent une logique cohérente qui va du constat de l'asymétrie d'information jusqu'à ses conséquences sur la communication des entreprises. Ce sont ces trois théories qui viendront motiver nos hypothèses de recherche.

2.2.1. La théorie de la signalisation : le label comme signal

La théorie de la signalisation part d'un simple constat : il existe une asymétrie d'information entre les dirigeants d'une entreprise et les investisseurs extérieurs. D'un côté, les dirigeants ont accès aux performances réelles de l'entreprise là où, de l'autre côté, les investisseurs n'ont comme information que des simples approximations. Dans le contexte ESG, les investisseurs externes ne peuvent pas directement observer la durabilité réelle des entreprises. Ils se basent donc sur les signaux émis par la firme pour en évaluer la qualité (Connelly et al., 2011, cité dans [M. Xu et al., 2025](#)).

C'est dans ce contexte que Kolling et Busch lancent leur étude sur les labels de durabilité en 2025. Pour les investisseurs qui cherchent à réduire leurs coûts d'information (Gutsche et Zwergel, 2016, cités dans [Kolling et Busch, 2025](#)), les labels de durabilité jouent le rôle de signaux crédibles. Un indice comme le BEL ESG renvoie un signal fort aux investisseurs en suggérant que la qualité de la durabilité de l'entreprise a été validée par un organisme compétent.

Une question émane alors : ce signal est-il fiable ? Imaginons que les entreprises aient accès à un indice ESG uniquement grâce à leur communication et non grâce à leurs actions réelles, le signal perd tout son sens et contribue au problème de greenwashing. C'est la tension que notre première hypothèse cherchera à tester sur le BEL ESG.

H1 : *Les entreprises appartenant au BEL ESG Index présentent une performance ESG réelle supérieure aux entreprises de contrôle non incluses dans l'indice*

2.2.2. La théorie de la légitimité : la pression sociale comme incitant à surenchérir

La théorie de la légitimité nous apporte une autre façon de voir les choses. Selon elle, une entreprise a besoin d'un « contrat social d'opérer ». Les entreprises doivent s'aligner sur les normes et attentes sociales sous peine de ne pas avoir accès aux ressources nécessaires à leur survie. Cet alignement se fait notamment à travers la divulgation d'informations sur leurs pratiques ESG.

Cependant, lorsque leur légitimité est remise en question, à cause de potentielles controverses, certaines entreprises peuvent utiliser la communication comme un outil de « gestion de l'impression » (Marquis et al., 2016, cité dans [G. Liu et al., 2024](#)). Le greenwashing devient alors une stratégie de communication forte capable de créer une image positive aux yeux des investisseurs sans avoir à mettre en place des actions concrètes. Plus la pression sociale et réglementaire est forte, au plus la surenchère discursive l'est aussi.

Dès lors, il est facile de comprendre le comportement des entreprises membres d'un indice ESG : une fois dans l'indice, elles intensifient leur communication pour consolider leur position aux yeux des investisseurs et des agences de notations. Notre deuxième hypothèse viendra vérifier si les entreprises du BEL ESG communiquent davantage sur les thématiques ESG que leurs pairs n'appartenant pas à celui-ci.

H2 : *Les entreprises du BEL ESG Index présentent une intensité de discours ESG plus élevée que les entreprises de contrôle n'appartenant pas à l'indice*

2.2.3. La théorie de l'agence : des conflits d'intérêt qui favorisent le camouflage

La théorie de l'agence nous offre un troisième angle d'analyse. Celle-ci, basée sur les dynamiques internes de l'entreprise, définit le greenwashing comme le résultat d'un conflit d'intérêts entre les dirigeants (agents) et les actionnaires (principaux). En effet, certains dirigeants n'hésitent pas à se servir des rapports de durabilité comme un outil égoïste pour projeter une image positive de l'entreprise en masquant les performances médiocres de celle-ci (Barnea & Rubin, 2010, cité dans [G. Liu et al., 2024](#)). Ce comportement de rétention de l'information aggrave l'asymétrie d'information entre les agents de l'entreprise et les actionnaires.

Une autre facette de la théorie de l'agence est celle du « narcissisme du PDG ». Certains dirigeants se lancent dans des actions de responsabilités sociales dont le but premier est

d'améliorer leur réputation personnelle plutôt que de créer de la valeur à long terme pour les actionnaires ([Oleg V. Petrenko et al., 2015](#)). Le greenwashing devient alors un comportement opportuniste alimenté par des conflits d'intérêts entre dirigeants et actionnaires.

Bien que ces trois théories nous expliquent pourquoi le greenwashing est rationnel pour les entreprises, elles n'expliquent pas dans quels contextes il se manifeste. C'est ce sur quoi nous nous pencherons dans la section suivante, avant d'examiner ses conséquences financières.

2.3. Les facteurs déterminants du greenwashing

Comprendre pourquoi les entreprises ont recours au greenwashing ne suffit pas, il faut également identifier les contextes dans lesquels ce phénomène est le plus susceptible de se produire. Pour se faire, la littérature détermine trois facteurs déterminants : la taille de l'entreprise, son secteur d'activité et enfin la pression réglementaire à laquelle l'entreprise est soumise.

2.3.1. La taille de l'entreprise

La taille de l'entreprise est le premier facteur, celui-ci explique la qualité et l'étendue du reporting ESG d'une entreprise. De fait, les grandes entreprises se doivent de communiquer davantage et de manière plus détaillée car elles sont exposées à la pression des investisseurs et à une visibilité médiatique accrue ([R. Hahn et M. Kühnen, 2013](#)). Cette visibilité accrue peut encourager les entreprises à soigner leur image à travers une communication embellissant leurs performances ESG plutôt que d'investir dans des améliorations réelles mais plus coûteuses.

Cette dynamique peut être transposée à l'échantillon du BEL ESG Index. Les entreprises appartenant à l'indice tendent à avoir une capitalisation boursière plus grande que celle de leurs pairs non-membres, ce que notre analyse empirique contrôlera à l'aide de la variable de capitalisation boursière. Kolling et Busch (2025) utilisent la taille du fonds comme variable de contrôle de leur appariement et reconnaissent que les plus grands fonds ont accès à davantage de ressources pour intégrer des critères de durabilité exigeants.

2.3.2. Le secteur d'activité

Un autre facteur déterminant est le secteur dans lequel l'entreprise exerce son activité. Par exemple, une étude indique que les entreprises manufacturières subissent des pertes de valeurs plus élevées en cas de greenwashing détecté que les entreprises de services. En effet, le secteur manufacturier est un acteur important dans la génération de gaspillage industriel et son impact environnemental est beaucoup plus visible. Il est également soumis à des réglementations bien plus strictes. C'est pourquoi la moindre nouvelle de greenwashing peut entraîner des examens réglementaires plus poussés et la chaîne d'approvisionnement peut être remise en doute ([M. Xu et al., 2025](#)).

Dans notre échantillon belge, les entreprises du BEL ESG exercent dans des secteurs différents, justifiant l'intégration d'une variable de contrôle sectorielle dans nos régressions.

2.3.3. La pression réglementaire : incitant ou frein au greenwashing ?

Etonnamment, des réglementations environnementales strictes peuvent inciter certaines entreprises à adopter des pratiques de greenwashing. Pour éviter des sanctions réglementaires ou par manque de budget, certaines entreprises préfèrent se tourner vers des stratégies de communications trompeuses à défaut de se lancer dans des investissements durables coûteux ([C. Fang et al., 2025](#)).

Face à ce constat, l'Europe a décidé de renforcer le cadre réglementaire en matière de transparence ESG. Les initiatives comme le Corporate Sustainability Reporting Directive (CSRD) ou la Green Claims Directives imposent l'utilisation de méthode scientifique comme l'Analyse du Cycle de Vie (ACV) pour justifier toutes informations environnementales ([M. Ruggeri et al., 2025](#)). L'Analyse du Cycle de Vie est un outil central du reporting de durabilité. Il permet aux entreprises d'évaluer les impacts environnementaux tout au long du cycle de vie de leurs activités et de renforcer la crédibilité des informations ESG qu'ils communiquent ([Stewart et al., 2018](#)).

Ces trois facteurs (taille, secteur et réglementation) définissent le contexte dans lequel le greenwashing se produit. Ils nous en disent également plus sur ses conséquences. Au plus ces facteurs incitent les entreprises à avoir recours au greenwashing, au plus les conséquences financières lors de la détection sont sévères.

2.4. Les conséquences financières et les risques associés au greenwashing

2.4.1. Réaction des marchés et risque de krach boursier

L'exposition d'un cas de greenwashing dans la presse entraîne une réponse immédiate des marchés financiers (Wang et al., 2015, cité dans [G. Liu et al., 2024](#)). Certains investisseurs perçoivent cette nouvelle comme un signe de faiblesse managériale et commencent à douter de la fiabilité des informations publiées l'entreprise. De plus, il est prouvé que les nouvelles négatives ont un impact boursier plus important que les nouvelles positives (Capelle-Blancard et Petit, 2019, cité dans [M. Xu et al., 2025](#)).

Au-delà de cette réponse immédiate, le greenwashing peut également augmenter le risque de chute des cours de la bourse ([G. Liu et al., 2024](#)). Une fois ces pratiques révélées, les entreprises qui avaient décidé de dissimuler leurs informations négatives peuvent être durement pénalisées par les investisseurs ce qui provoque une correction boursière brutale.

2.4.2. Le paradoxe de la réputation

Ce paradoxe vient mettre en lumière une chose assez surprenante. En effet, les entreprises affichant une performance ESG au-dessus de la moyenne subissent des pertes de valeurs boursières plus importantes en cas de découverte d'un cas de greenwashing ([M. Xu et al., 2025](#)). Ces entreprises, à force de fournir des efforts pour être plus durable, ont accumulé une certaine réputation auprès des investisseurs. Plus l'entreprise est réputée, plus les investisseurs lui font confiance. Trahir cette confiance entraîne une grande déception auprès des parties prenantes et génère une chute de valeur boursière amplifiée.

Ce paradoxe est particulièrement pertinent dans le cas du BEL ESG. En effet, les entreprises incluses dans l'index bénéficient d'une image de durabilité renforcée et s'exposent à des conséquences plus sévères si un écart entre le discours et la performance est détecté.

2.5. Labels, indices ESG et mesure du greenwashing

Les indices et labels ESG répondent au problème d'asymétrie d'information identifié par la théorie de la signalisation. Ils jouent un rôle crucial sur les marchés financiers pour 3 raisons. D'une part, ils constituent un signal fort en mettant en avant les entreprises qui répondent à des critères de durabilité. D'autre part, ils ont un impact sur l'allocation du capital, leur objectif est

de diriger les flux de capitaux vers des entreprises qui prônent la durabilité. Enfin, des modifications de la composition de l'indice peuvent entraîner des variations sur le prix des actifs et les volumes échangés ([Barontini et Gioja, 2025](#)). Cependant, l'évidence empirique concernant leur fiabilité peut être plus contrastée que ce qu'il n'y paraît.

2.5.1. L'étude de référence : Kolling et Busch (2025)

L'étude de [Kolling et Busch \(2025\)](#), « Sustainability labels matter », publiée dans Finance Research Letters, est l'article de référence de ce mémoire. Elle se demande si les fonds détenant le FNG-Label, label germanophone de durabilité, affichent une performance de durabilité supérieure à celle des fonds non labellisés. Question que nous allons transposer au BEL ESG Index dans notre étude.

Au niveau de la méthodologie, pour construire leur échantillon, les auteurs adoptent une approche par appariement sur plus de 150 fonds détenant le FNG-Label 2023. L'appariement s'effectue au niveau de la taille du fond, l'âge, la zone géographique et enfin les secteurs dans lesquels ils investissent. Grâce à l'appariement, deux échantillons de contrôle sont construits. Le premier composé de fonds conventionnels et le deuxième de fonds affichant une stratégie durable sans label.

Ensuite, pour tous les fonds, la performance de durabilité est mesurée à l'aide de plusieurs indicateurs. Le premier, le ESG Score (LSEG), mesure la gestion des enjeux ESG d'une entreprise relativement à ses pairs sectoriels. Le deuxième, l'ESG Combined Score (ESGC) complète le premier indicateur en y intégrant un score de controverse. Enfin, le SDG Score mesure la contribution (positive ou négative) d'une entreprise aux Objectifs de Développement Durable. Les émissions de CO2 (Scope 1 et 2) sont aussi analysées en valeur absolue et en intensité relative au chiffre d'affaires.

Mais quels sont les résultats de leur étude ? Premièrement, les fonds FNG présentent des émissions de CO2 totales significativement inférieures pour les trois échantillons. De plus leurs scores SDG sont largement supérieurs. Cependant, ils observent une chose étrange. L'ESG score classiques des fonds FNG n'est pas significativement différent de celui des fonds non labellisés. C'est seulement une fois que l'on intègre les controverses et qu'on travaille donc avec les ESGC qu'on retrouve une différence significative. C'est cette dissociation entre ESG score classique et ESGC qui constitue le cœur de la contribution de leur étude. En effet, les scores ESG classiques sont orientés vers la matérialité financière ([Khan et al., 2016](#)) et ne reflète pas l'impact environnemental réel.

Les auteurs concluent en nous disant que les labels sont capables de réduire les asymétries d'information mais uniquement si on utilise des mesures d'impact réel et pas des scores ESG classiques.

2.5.2. Mesurer l'écart : les approches pour quantifier le greenwashing

En parallèle des travaux réalisés sur les labels, la littérature nous propose différentes approches pour quantifier l'écart entre le discours et la performance durable réelle de l'entreprise. Une première approche, appelée « méthode des scores », identifie le greenwashing comme la différence entre le score de divulgation ESG et le score de performance ESG réelle (Yu et al., 2020, cité dans [G. Liu et al., 2024](#)). Un écart positif indique que l'entreprise communique beaucoup plus que ce qu'elle fait réellement et qu'elle utilise les rapports de durabilité comme un outil de camouflage pour tromper les investisseurs.

Une seconde approche, introduite par Yu et al. en 2023 (cité dans [C. Fang et al., 2025](#)), utilise l'indice de « catering » pour mesurer la déviation entre la promotion verbale d'initiatives vertes et les performances environnementales réelles de l'entreprise. L'indice se construit à partir de trois indicateurs distincts.

Premièrement, l'indice de divulgation ESG (CID) qui mesure l'intensité du discours ESG. Celui-ci est obtenu grâce à une analyse textuelle stricte qui se résume à compter le nombre de répétitions de termes ESG dans les rapports de durabilité de l'entreprise. Une fois la fréquence des mots obtenues, il est nécessaire de la normaliser par rapport au nombre de mots du rapport. Deuxièmement, l'indice d'actions ESG (Action) est calculé pour mesurer la performance réelle de l'entreprise. Cette performance se définit autour de 3 axes, la rigueur, les efforts et les résultats en matière de durabilité. Les actions sont considérées comme des variables binaires et sommées afin d'obtenir un score global standardisé.

Enfin, le degré de greenwashing de l'entreprise est mesuré en standardisant les deux indices précédents et en prenant la différence entre l'indice de divulgation (CID) et l'indice d'actions (Action). Un score élevé reflète un écart entre le discours et la performance réelle de l'entreprise.

C'est dans cette logique de mesure que nous combinerons une analyse textuelle du discours ainsi que des indicateurs de performance réelle. L'ESG linguistic score (ESG_LING) que nous construirons par la suite viendra remplacer le CID de Yu et al. (2023), adapté aux rapports annuels de nos entreprises du marché belge. C'est dans ce contexte que notre troisième

hypothèse prend tout son sens et viendra s'interroger sur les possibles écarts entre discours et performance réelle.

H3 : *A performance ESG comparable, les entreprises appartenant au BEL ESG Index présentent une intensité du discours ESG plus élevée que les entreprises comparables non incluses dans l'indice*

2.6. Limites de la littérature et positionnement de ce mémoire

Malgré la richesse de la littérature sur le greenwashing et les critères ESG, certaines limites peuvent être identifiées qui définissent ce que ce mémoire cherche à combler.

Premièrement, la majorité des études se concentrent sur les scores ESG fournis par les agences de notations classiques dont la crédibilité et la cohérence ne cessent d'être remises en question. En effet, ces scores se basent sur la matérialité financière et non pas sur l'impact environnemental réel des entreprises.

Deuxièmement, la plupart des travaux réalisés étudient séparément la communication ESG et les performances durables réelles. Or, c'est précisément cet écart qui constitue le cœur du problème de greenwashing et sur lequel il faut se pencher.

Troisièmement, il existe peu d'études qui utilisent une analyse textuelle afin de mesurer le discours ESG de manière objective.

C'est dans ce contexte que mon mémoire contribue à la littérature existante en appliquant la démarche de Kolling et Busch (2025) à un indice boursier belge d'actions individuelles, le BEL ESG Index. En combinant une analyse textuelle des discours ESG avec des mesures de performance durable réelle (ESGC, Emissions Scope 1 et 2).

Cette combinaison permet de tester si le BEL ESG Index sélectionne bien des entreprises qui sont plus durables (H1), si ses membres communiquent davantage (H2) et enfin si les entreprises membres ont tendance à surcommuniquer (H3). La section suivante viendra formuler ces trois hypothèses de manière opérationnelle.

3. Développement des hypothèses

Cette section vise à formaliser les hypothèses en termes opérationnels. L'objectif est de déterminer quels indicateurs sont utilisés, comment ils sont calculés et enfin quelle conclusion on pourra tirer s'il existe une différence significative entre les deux groupes.

3.1. Les entreprises appartenant au BEL ESG Index présentent une performance ESG réelle supérieure aux entreprises de contrôle non incluses dans l'indice

Notre première hypothèse testera si l'appartenance au BEL ESG Index est synonyme d'une performance durable réelle supérieure. Il est possible de scinder celle-ci en deux sous-hypothèses. La première sous-hypothèse viendra tester la capacité à gérer les risques ESG là où la deuxième s'attardera sur l'impact réel de l'entreprise sur l'environnement et la société.

Deux variables seront utilisées afin de quantifier la performance ESG réelle des entreprises. La première, l'ESG Combined score (ESGC) combine la gestion ESG ainsi que les potentielles controverses auxquelles l'entreprise a pu faire face. Un ESGC élevé signifie que les agences de cotations n'ont pas identifié de problèmes graves. Ce score est récupéré sur Bloomberg, un fournisseur de données financières.

La deuxième variable est l'intensité carbone des entreprises et elle sera construite à partir de deux indicateurs collectés via Bloomberg, le GHG Scope 1 Intensity per Sales et le GHG Scope 2 Intensity per Sales. Ces deux indicateurs mesurent les émissions directes et indirectes liées à la consommation de l'énergie d'une entreprise. La variable GHG_INT est obtenue en additionnant ces deux indicateurs.

$$\text{GHG_INT} = \text{GHG Scope 1 Intensity per Sales} + \text{GHG Scope 2 Intensity per Sales}$$

Si l'entreprise a une valeur de GHG_INT faible, cela signifie qu'à niveau d'activité comparable, elle a une meilleure performance environnementale.

H1a Opérationnelle : *Les entreprises du BEL ESG Index ont un ESGC supérieur à celui des entreprises comparables n'appartenant pas à l'indice*

H1b Opérationnelle : *Les entreprises appartenant au BEL ESG Index présentent une intensité carbone CO2/CA (GHG_INT) inférieure à celle des entreprises comparables non incluses dans l'indice*

3.2. Les entreprises du BEL ESG Index présentent une intensité de discours ESG plus élevée que les entreprises n'appartenant pas à l'indice

Notre deuxième hypothèse se demande si le fait d'appartenir au BEL ESG incite les entreprises à augmenter l'intensité de leur discours ESG. On vient tester le mécanisme de légitimité. Est-ce qu'une fois qu'elles appartiennent à un indice ESG, les entreprises communiquent davantage pour montrer leur alignement avec les attentes des investisseurs ?

Nous utiliserons la variable ESG Linguistic Score (ESG_LING) que nous construisons à l'aide d'une analyse textuelle des rapports annuels des entreprises. Le ESG_LING correspond au nombre d'occurrences de mots ESG pour 10 000 mots dans le rapport annuel. Mais attention, celui-ci mesure l'intensité du discours et non la qualité de celui-ci.

H2 Opérationnelle : *Les entreprises du BEL ESG ont un ESG_LING significativement plus élevé que les entreprises de contrôle n'y appartenant pas*

3.3. À performance ESG comparable, les entreprises appartenant au BEL ESG Index présentent une intensité de discours ESG plus élevée que les entreprises comparables non incluses dans l'indice

Notre troisième hypothèse est la plus originale de ce mémoire. En effet, elle teste si, avec des niveaux de performance ESG comparable, les entreprises incluses dans le BEL ESG communiquent davantage que leurs pairs.

H3 Opérationnelle : *A performance ESG comparable, les entreprises du BEL ESG présentent un ESG_LING significativement plus élevé que leurs pairs non-membres*

Cette hypothèse est testée à l'aide d'un modèle de régression MCO en panel groupé. Un modèle avec les effets fixes individuels n'a pas été retenu. De fait, avec un échantillon comprenant que 40 entreprises, inclure un effet fixe par entreprise absorberait une trop grande part de la variation et compromettrait l'estimation de nos coefficients.

Le modèle s'écrit :

$$ESG_LING_{i,t} = \alpha + \beta_1 Perf_ESG_{it} + \beta_2 BEL_ESG_{it} + \sum \beta_{3k} Année_k,t \\ + \sum \beta_{4j} Secteur_j,i + \beta_5 Taille_{it} + \epsilon_{it}$$

Où $ESG_LING_{i,t}$ désigne l'intensité du discours de l'entreprise i à l'année t , $Perf_ESG_{it}$ est l'indice composite de la performance ESG réelle (expliqué en détail dans la section 4.3.5) et BEL_ESG_{it} la variable binaire d'appartenance à l'indice.

Si le coefficient β_2 est positif et significatif, cela signifie que les entreprises du BEL ESG utilisent, en moyenne, plus de langage ESG dans leurs rapports que leurs pairs non-membres, une fois la performance ESG prise en compte dans la régression. Attention, ici il ne s'agit pas de faire une comparaison entreprise par entreprise à performance identique comme le permettrait un appariement sur critère. Il s'agit bien d'un contrôle statistique appliqué sur l'ensemble de notre échantillon. Cet excès de communication, non expliqué par la performance observable, serait un indicateur potentiel de greenwashing.

4. Méthodologie

Une description de la méthodologie globale est donnée dans la Figure 1 avec le détail de toutes les étapes dans les sections suivantes.

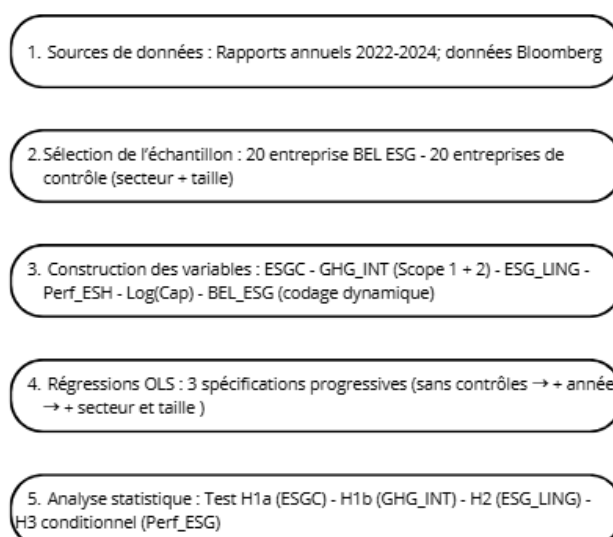


Figure 1. Description de l'approche méthodologique

4.1. Sources de données

Dans le cadre de ce travail, l'utilisation de deux sources de données a été nécessaire.

D'une part, afin d'établir la mesure du discours ESG, les rapports annuels des entreprises comprises dans l'échantillon ont été récoltés. Ces rapports structurent la communication officielle adressée aux différentes parties prenantes. Cette communication doit être analysée en profondeur car elle est précisément le lieu où peuvent apparaître des stratégies de gestion de l'impression ou encore de découplage. Pour toutes les entreprises de notre échantillon, les rapports annuels ont été collectés sur trois années consécutives (2022,2023,2024), ce qui donne à l'échantillon une structure en panel (une observation par couple entreprise-année). Cette période est synonyme de montée en puissance du cadre réglementaire européen ainsi que du reporting.

De fait, avec des événements comme l'Accord de Paris en 2016 ou l'entrée en vigueur de la Taxonomie européenne et de la CSRD, les attentes en matière de stratégie de climat et de transparence s'intensifient.

Pour ce qui est du protocole de collecte, la grande majorité des rapports annuels ont été téléchargés en version anglaise et en format pdf à partir du site companies market cap. Dans le cas où les rapports annuels ne s'y trouvaient pas, le téléchargement s'est fait via le site investisseur des entreprises.

D'autres part, afin de mesurer la dimension ESG des entreprises, ce mémoire mobilise des données issues de Bloomberg. Les indicateurs ont été collectés pour chaque entreprise, sur les mêmes trois années que les rapports annuels à savoir 2022, 2023 et 2024 afin de garantir une cohérence temporelle entre les deux sources de données. Ces données permettent de comparer de manière objective les différentes entreprises. Bloomberg contient une multitude d'indicateurs mais nous nous intéresserons qu'à trois d'entre eux. Le premier, le ESG Combined Score, est un indicateur qui mesure la performance ESG globale d'une entreprise. Il tient compte des critères environnementaux, sociaux et de gouvernance et nous est donc utile pour positionner une entreprise par rapport aux autres. Les deux indicateurs suivants vont de pairs, il s'agit des émissions de CO2 (Scope 1 et Scope 2). Les émissions Scope 1 mesurent les émissions directes, alors que les émissions Scope 2 concernent les émissions indirectes liées à la consommation d'énergie. Dans le cadre de ce mémoire, nous avons choisi de ne pas inclure le Scope 3 dans notre analyse car certaines entreprises reportent partiellement, voire pas du tout sur cet indicateur.

4.2. Sélection de l'échantillon

Le défi central dans l'échantillonnage était d'éviter un biais de sélection. En effet, les entreprises du BEL ESG sont sélectionnées car celles-ci satisfont des critères ESG. L'objectif du matching est de construire un groupe d'entreprises de contrôles afin de réussir à isoler l'effet de l'appartenance à l'indice BEL ESG sur les différentes mesures étudiées. Cet appariement s'est appuyé sur la composition de l'indice BEL ESG de décembre 2025, la plus récente au moment de construire notre échantillon.

Le matching s'est déroulé en deux étapes, l'appariement exact sur secteur et par la suite l'appariement sur la taille. Pour chaque entreprise appartenant au BEL ESG, une entreprise du BEL Mid, du BEL 20 ou cas échéant du BEL ALL opérant dans le même secteur a été sélectionnée. Les informations concernant le secteur d'activité des entreprises (par exemple Industriel, Immobilier, Financier, Santé) de l'échantillon ont été trouvées sur le site d'[Euronext](#). Une fois la première phase de l'échantillonnage effectuée, un appariement sur la taille a été réalisé à l'aide de la capitalisation boursière. La méthode employée a été celle du « plus proche voisin ». L'objectif est d'associer à chaque entreprise du BEL ESG une entreprise du groupe de contrôle présentant une capitalisation boursière la plus proche possible au sein du même secteur.

BEL ESG Entreprise	Index	Capitalisation Boursière	log(capitalisation)	Secteur	Entreprise Contrôle	Index	Capitalisation Boursière	log(capitalisation)
ACKERMANS	BEL 20	9,291	0,96806246	Industriel	DEME GROUP	BEL MID	4,86	0,686636269
AEDIFICA	BEL 20	3,443	0,536937023	Immobilier	RETAIL ESTATES	BEL MID	1,008	0,003460532
AGEAS	BEL 20	12,205	1,086537784	Financier	BEKAERT	BEL MID	2,016	0,304490528
AZELIS GROUP	BEL 20	1,886	0,275541688	Industriel	CMB.TECH	BEL MID	3,817	0,58172216
COFINIMMO	BEL 20	3,187	0,503382063	Immobilier	EUROCOMMERCIAL PROPERTIES	BEL ALL	1,531	0,184975191
DIETEREN GROUP	BEL 20	9,965	0,998477303	Consommation discrétionnaire	KINEPOLIS GROUP	BEL MID	0,724	-0,140261434
ELIA GROUP	BEL 20	14,561	1,163191202	Services publics	FLUXYS	BEL ALL	0,248	-0,605548319
GBL	BEL 20	11,082	1,044618146	Financier	EURONEXT	BEL ALL	14,458	1,16010822
KBC	BEL 20	46,695	1,66927038	Financier	ING GROEP N.V.	BEL ALL	66,613	1,823558993
MONTEA	BEL 20	1,751	0,243286146	Immobilier	WERELDHAVE BELGIUM	BEL ALL	0,599	-0,222573178
SOFINA	BEL 20	8,895	0,949145952	Financier	CIE BOIS SAUVAGE	BEL ALL	0,505	-0,296708622
UCB	BEL 20	50,922	1,706905453	Santé	ARGENX SE	BEL 20	40,076	1,602884368
UMICORE	BEL 20	4,438	0,647187298	Matériaux de base	SOLVAY	BEL 20	2,829	0,451632947
WDP	BEL 20	5,85	0,767155866	Immobilier	SHURGARD	BEL MID	2,777	0,44357588
APERAM	BEL MID	2,96	0,471291711	Matériaux de base	TESSENDERLO	BEL MID	1,24	0,093421685
BARCO	BEL MID	0,994	-0,002613616	Technologie	MELEXIS	BEL 20	2,202	0,342817315
FAGRON	BEL MID	1,647	0,216693599	Santé	GALAPAGOS	BEL MID	1,951	0,290257269
GIMV	BEL MID	1,703	0,231214648	Financier	BANQUE NATIONALE BELGE	BEL ALL	0,144	-0,841637508
PROXIMUS	BEL MID	2,593	0,413802517	Télécommunication	EVS BROADC.EQUPM.	BEL MID	0,494	-0,306273051
VGP	BEL MID	2,806	0,448087667	Immobilier	XIOR	BEL MID	1,31	0,117271296

Tableau 1. Echantillon

Pour évaluer la qualité de l'appariement entre les entreprises du BEL ESG et le groupe de contrôle, nous avons réalisé un test de balance. L'objectif est de vérifier que les deux groupes présentent des caractéristiques similaires afin d'assurer la validité des comparaisons empiriques.

Le contrôle repose, en premier lieu, sur l'analyse des statistiques descriptives comme les moyennes et les médianes de la capitalisation boursière. Pour limiter l'influence des valeurs extrêmes, nous avons appliqué une transformation logarithmique aux données de capitalisation. Ensuite, un test de différence de moyennes a été effectué pour identifier de potentielles différences statistiques entre les deux groupes. Ce type de test étant sensible à la taille de l'échantillon, nous sommes partis sur une mesure standardisée indépendante des unités, à savoir le Standardized Mean Difference (SMD).

$$d = \frac{(\bar{x}_{\text{treatment}} - \bar{x}_{\text{control}})}{\sqrt{\frac{s_{\text{treatment}}^2 + s_{\text{control}}^2}{2}}}$$

Figure 2. Formule du SMD

Le SMD est défini comme la différence de moyennes entre les deux groupes, normalisée par leur écart-type combiné. Cet indicateur de similarité entre deux groupes est largement utilisé dans la littérature empirique afin d'évaluer la qualité d'un appariement. Selon Austin ([2009](#)), le SMD permet d'évaluer l'équilibre entre deux groupes indépendamment de la taille de l'échantillon. Des valeurs proches de zéro indique une bonne balance entre les deux groupes, alors que plus les valeurs sont grandes, plus c'est un signal de déséquilibre.

Moyennes				Médianes				Ecart type	
BEL ESG		hors BEL ESG		BEL ESG		hors BEL ESG		BEL ESG	hors BEL ESG
Capitalisation	log(capitalisation)	Capitalisation	log(capitalisation)	Capitalisation	log(capitalisation)	Capitalisation	log(capitalisation)	0,482083618	0,686763285
10	0,731057243	8	0,292449434	4	0,647187298	2	0,290257269		
SMD	0,74708								

Tableau 2. Test de balance

Le test de balance met en évidence un certain déséquilibre entre les entreprises du BEL ESG Index et les entreprises du groupe de contrôle. En effet, l'indice présente, en moyenne et en médiane, une capitalisation boursière plus élevée que les entreprises de contrôles.

Toujours selon Austin (2009), un SMD de 0,74 reflète un gros déséquilibre entre les deux groupes suggérant que l'appariement n'a pas su atteindre un équilibre satisfaisant pour la capitalisation boursière.

Cependant, ce résultat n'est pas anodin. Il montre un constat du marché Belge, les entreprises de plus grande taille sont davantage représentées dans les indices ESG et cela peut être expliqué par des plus grandes ressources ou une meilleure visibilité. Par conséquent, bien que les caractéristiques sectorielles aient été respectées, on observe des différences au niveau de la taille des entreprises. Pour tenir compte de ce déséquilibre, les analyses empiriques futures intégreront des variables de contrôles, notamment la capitalisation boursière des entreprises.

4.3. Construction des variables

Dans cette section, nous présentons les différentes variables utilisées dans l'analyse empirique. Pour offrir une vision plus claire et synthétique, le tableau suivant résume les différentes variables, leur définition ainsi que leur source.

Ensuite, chaque variable sera détaillée afin de préciser leur construction et leur interprétation.

Variable	Type	Définition	Ordre de grandeur	Source
BEL_ESG	Variable d'intérêt (H1a, H1b, H2, H3)	Variable binaire qui prend la valeur 1 si l'entreprise appartient au BEL ESG Index pour l'année t, 0 sinon	0 ou 1	Euronext
ESG_LING	Variable dépendante (H2, H3)	Nombre d'occurrences de mots ESG pour 10 000 mots dans le rapport annuel	150 - 650 (dans l'échantillon)	Rapports annuels
ESGC	Variable dépendante (H1a)	Score combinant la gestion des risques ESG et les controverses. Pénalise les entreprises impliquées dans des comportements problématiques	0-100	Bloomberg
GHG Scope 1 Intensity per Sales	Variable dépendante brute (H1b)	Emissions directes de gaz à effet de serre générées par les activités de l'entreprise, rapportées au chiffre d'affaires	En tonnes CO2e/millions d'euros de chiffre d'affaires	Bloomberg
GHG Scope 2 Intensity per Sales	Variable dépendante brute (H1b)	Emissions indirectes liées à la consommation d'énergie achetée, rapportées au chiffre d'affaires	En tonnes CO2e/millions d'euros de chiffre d'affaires	Bloomberg
GHG_INT	Variable dépendante construite (H1b)	Somme des intensités carbone Scope 1 et Scope 2 rapportées au chiffre d'affaires. Une valeur faible indique une meilleure performance environnementale	En tonnes CO2e/millions d'euros de chiffre d'affaires	Construit à partir de Bloomberg
PERF_ESG	Variable de contrôle construite (H3)	Indice de performance réelle combinant l'ESGC et l'intensité carbone	Score standardisé, centré autour de 0	Construit à partir de Bloomberg
Taille (log)	Variable de contrôle	Logarithme de la capitalisation boursière pour contrôler l'effet taille	En log de milliards	Euronext
Secteur	Variable de contrôle	Variable binaire représentant le secteur d'activité	0 ou 1	Companies Market Cap
Année	Variable de contrôle	Variable binaire capturant les chocs temporels communs (2022 est la référence)	0 ou 1	/

Tableau 3. Tableau synthétique des variables

4.3.1. Variable d'intérêt : appartenance au BEL ESG Index

La variable principale de notre analyse est une variable binaire notée BEL_ESG_{it} , qui prend comme valeur 1 si l'entreprise i appartient au BEL ESG Index pour l'année t , 0 sinon. Cette variable est construite à partir des compositions de l'indice publiées par Euronext depuis son lancement en février 2023. L'indice n'existant pas avant cette date, la variable prend la valeur 0 pour l'ensemble des entreprises de l'échantillon en 2022, année de référence.

Notre étude adopte une approche dynamique pour les années 2023 et 2024 afin de tenir compte de potentiels changements dans la composition de l'indice au fil du temps. En effet, le BEL ESG Index fait l'objet de révisions trimestrielles basées sur les profils ESG des entreprises belges entraînant des entrées et sorties dans l'indice.

Dans un premier temps, l'appartenance à l'indice est observée sur une base trimestrielle. Dans un second temps, afin d'assurer une certaine cohérence temporelle avec les autres variables (notamment la variable linguistique issue des rapports annuels), la variable BEL_ESG_{it} est construite à partir de la règle suivante : une entreprise est considérée comme appartenant à l'indice pour une année donnée si celle-ci est incluse dans au moins deux périodes sur les quatre au cours de l'année. Sinon, la variable prend la valeur 0 ([voir Annexe A](#))

Au niveau de l'interprétation, la variable BEL_ESG_{it} peut être considérée comme un indicateur de labellisation ESG. En effet, les entreprises de l'indice sont sélectionnées sur base de critères ESG spécifique et l'appartenance à celui-ci représente un signal fort pour les investisseurs.

4.3.2. Variable de discours : ESG Linguistic Score (ESG_LING)

Afin de mesurer l'intensité du discours ESG des entreprises, cette étude mobilise une analyse textuelle basée sur la méthode « bag-of-words » ainsi que l'utilisation d'un dictionnaire de termes ESG. La méthode, largement utilisée en finance, consiste à extraire des informations quantitatives de documents textuels comme les rapports annuels ou les publications réglementaires.

La méthode repose sur l'idée que certaines caractéristiques du discours telles que le ton, la lisibilité ou l'importance accordée à certaines thématiques peuvent être capturer à l'aide d'un comptage standardisé de mots spécifiques. Dans le cadre de notre étude, l'analyse textuelle viendra mesurer l'intensité du discours lié aux enjeux environnementaux, sociaux et de gouvernance. Celle-ci sera effectuée via RStudio et le code relatif à l'analyse est disponible en [annexe C](#).

La variable ESG_LING_{it} est construite pour chaque entreprise i et chaque année t sur base des rapports annuels. Elle est calculée comme suit : $ESG_LING_{it} = \frac{C_{i,t}}{W_{i,t}} \times 10000$ où $C_{i,t}$ représente le nombre total d'occurrences de mots appartenant au dictionnaire ESG et $W_{i,t}$ correspond au nombre total de mots dans le rapport annuel. Afin d'obtenir une mesure standardisée et facilement interprétable, on multiplie le résultat de la division par 10000. Par exemple, un score de 350 mots signifie que 3,5% des mots du rapport appartiennent au lexique ESG. La normalisation est obligatoire afin de rendre les scores comparables entre entreprises, indépendamment de la longueur des rapports annuels.

Le choix du dictionnaire ESG est une étape cruciale dans notre méthodologie car il détermine la pertinence de notre score linguistique. C'est pourquoi nous avons décidé de reprendre le dictionnaire publié par [Baier et al. \(2020\)](#) qu'ils ont utilisé dans leur étude s'intitulant « *Environmental, Social and Governance Reporting: A Textual Analysis* ». Outre la version anglaise de ce dictionnaire, une version française a été traduite manuellement afin de répondre à un problème rencontré lors de la collecte des rapports annuels ([voir lien dictionnaire Baier](#)). En effet, trois rapports de notre échantillon (rapports annuels de CIE BOIS SAUVAGE) sont rédigés en français et nécessite l'utilisation de la version française du dictionnaire.

Il est important de noter que la variable ESG_LING mesure l'intensité du discours ESG et non sa qualité. Une entreprise qui utilise fréquemment des termes repris dans le dictionnaire ESG obtiendra un score ESG élevé, que ce soit dans des formulations vagues ou dans des engagements concrets.

Cette limite, reconnue dans la littérature, sera détaillée dans la section dédiée.

4.3.3. Variable de performance (1) : ESG Combined Score (ESGC)

Le ESG Combined Score (ESGC) est la première mesure de performance ESG réelle et est collectée via Bloomberg. Cet indicateur est composé de deux scores : le score ESG classique qui évalue la gestion des risques ESG par rapport aux pairs du même secteur et un score de controverse qui capte les potentiels comportements problématiques dans lesquels l'entreprise a pu être impliqué.

Au niveau de la construction, l'ESGC correspond à la moyenne entre le score ESG brut et le score de controverse. Cette formulation empêche qu'une entreprise puisse compenser ses comportements problématiques par un bon score ESG général.

L'ESGC oscille entre 0 et 100 et un score élevé indique une bonne gestion des risques ESG ainsi que l'absence de comportements problématiques.

Cette variable sera mobilisée pour tester l'hypothèse H1a.

4.3.4. Variable de performance (2) : Intensité carbone GHG Scope 1 et 2 (GHG_INT)

La deuxième variable de performance réelle mesure l'intensité carbone. Elle est construite à l'aide des émissions de gaz à effet de serre de Scope 1 et Scope 2 qui sont collectées via Bloomberg. Pour rappel, les émissions de Scope 1 englobent les émissions directes produites par les activités de l'entreprise (combustion de carburant, procédés de fabrication) alors que les émissions de Scope 2 concernent les émissions indirectes liées à la consommation d'énergie achetée (électricité, vapeur). Le Scope 3, qui recouvre les émissions de toute la chaîne de valeur, a été exclu de l'analyse à cause de son taux de couverture insuffisant dans notre échantillon. Pour rendre les émissions comparables entre les différentes entreprises de notre échantillon, nous avons décidé de standardiser celles-ci par le chiffre d'affaires de l'entreprise. Nous pouvons donc définir la variable GHG_INT comme suit :

$$GHG_INT_{i,t} = (Scope\ 1_{i,t} + Scope\ 2_{i,t}) / Chiffre\ d'affaires_{i,t}$$

Cette mesure finale, exprimée en tonnes de CO2 équivalent par million d'euros de chiffre d'affaires, quantifie un aspect de l'impact environnemental réel de l'entreprise indépendamment de sa taille. Contrairement à l'ESGC qui relève de la question de la matérialité financière, le GHG_INT traite la matérialité environnementale. Mobiliser ces deux perspectives simultanément permet d'obtenir une image plus complète de la performance durable réelle de l'entreprise. C'est ce que nous ferons dans la section suivante à l'aide de la variable PERF_ESG. Dans les régressions qui vont suivre, une valeur plus élevée de GHG_INT reflète une performance environnementale plus faible.

Nous utiliserons cette variable pour tester l'hypothèse H1b.

4.3.5. Indice de performance composite (PERF_ESG)

Afin de tester notre troisième et dernière hypothèse, qui nécessite de contrôler la performance réelle sur ses deux dimensions, l'indice composite PERF_ESG a été construit. Celui-ci combine l'ESGC et l'intensité carbone GHG_INT pour créer un score unique de performance. Une valeur plus élevée de l'indice correspond à une meilleure performance réelle de l'entreprise. L'indice est défini comme suit :

$$PERF_ESG_{i,t} = \left[z(ESGC_{i,t}) + z\left(-\ln(1 + GHG_{INT_{i,t}})\right) \right] / 2$$

La construction de l'indice repose sur trois étapes successives. La première étape est de standardiser par z-score les deux variantes. En effet, l'ESGC varie entre 0 et 100 alors que GHG_INT s'exprime en tonnes de CO2 par million d'euros de chiffre d'affaires. Si l'on effectue une simple addition, la variable GHG_INT dominerait le score composite par son ordre de grandeur plus élevé. La standardisation par z-score ramène chaque variable à une moyenne de 0 et un écart-type de 1 rendant le score plus pertinent.

La deuxième étape consiste à appliquer une transformation logarithmique à la variable GHG_INT. En effet, on observe que certaines entreprises ont des valeurs d'émissions carbone extrêmes, notamment dans le secteur de l'industrie ou des matériaux de base. Ces valeurs risqueraient d'exercer une influence disproportionnée sur les estimations. En utilisant le logarithme de la variable GHG_INT, on comprime ces valeurs extrêmes, condition nécessaire pour les régressions dont nous parlerons dans la section suivante. Le terme +1 est également ajouté afin d'éviter le cas mathématiquement incorrect où GHG_INT serait égal à 0. Ce cas n'est pas présent dans notre échantillon puisque la valeur minimale de GHG_INT est de 0,03 tonnes de CO2e par million d'euros de chiffre d'affaires et une entreprise affichant une intensité carbone nulle est rare mais pas impossible. Il convient donc d'ajuster notre variable en y ajoutant le terme +1.

Enfin, l'inversion du signe sur GHG_INT est la troisième et dernière étape. Pour la variable ESGC, une valeur élevée reflète une bonne performance ESG alors que pour GHG_INT, une valeur élevée traduit une mauvaise performance environnementale. Sans correction, les deux variables tireraient l'indice dans des directions opposées. En appliquant un signe négatif, on s'assure qu'une valeur élevée de l'indice correspond systématiquement à une meilleure performance. Ensuite, on réalise la moyenne des deux z-scores traduisant l'hypothèse que la matérialité financière et la matérialité environnementale participent de manière équivalente à la performance durable réelle de l'entreprise.

Illustration numérique : BEKAERT en 2022 (ESGC = 57,29 et GHG_INT = 307,24)

Etape	Calcul	Résultat
1. Z-score de l'ESGC Moyenne = 50,45 Ecart-type = 10,46	$(57,28 - 50,45) / 10,46$	0,65
2. Log puis Z-score de GHG_INT Moyenne du log = 2,89 Ecart-type du log = 1,98	$\ln(307,24 + 1) = 5,73$ $(5,73 - 2,89) / 1,98$	1,44
3. Inversion du signe	$-(1,44)$	-1,44
Résultat : Perf_ESG	$(0,65 + (-1,44)) / 2$	-0,39

4.4. Statistiques descriptives

Avant de procéder aux différents tests économétriques, il est nécessaire de faire le point sur les statistiques descriptives de notre échantillon. Directement inspirée de [Kolling et Busch \(2025\)](#), cette étape donne une première idée sur les différences qui peuvent exister entre les deux groupes. Les statistiques sont présentées en séparant la période pré-indice (2022) et la période durant laquelle la variable BEL_ESG est définie et varie (2023-2024).

4.4.1. Année de référence 2022 (pré-indice)

L'année 2022 est l'année de référence pré-indice. En effet, Euronext a lancé le BEL ESG qu'en février 2023 obligeant la variable BEL_ESG à prendre la valeur 0 pour toutes les entreprises de l'échantillon en 2022. Les statistiques descriptives de la période pré-indice permettent de comparer le profil des entreprises qui seront sélectionnées dans l'indice en 2023 à celui des futures entreprises de contrôles.

Le tableau ci-dessous résume les statistiques pour l'année 2022, scindées en deux groupes (BEL ESG et Contrôle, basée sur la composition de l'indice en 2023)

Variable	Groupe	Observations	Moyenne	Médiane	Ecart-type	Min	Max
ESG_LING	BEL ESG	17	342,4	336,8	64,8	254	464
	Contrôle	23	309,8	309,5	70,3	212,6	490,2
ESGC	BEL ESG	17	53	53,7	8,9	35,7	67,2
	Contrôle	23	47	47,5	11,6	21,6	70,9
GHG_INT	BEL ESG	17	104	8,4	269,1	0,04	1077,830
	Contrôle	23	108,8	16,1	548,1	0,03	2651,280
Perf_ESG	BEL ESG	17	0,252	0,418	0,458	-0,722	0,915
	Contrôle	23	-0,265	-0,261	0,703	-1,557	0,995

En 2022, avant même le lancement du BEL ESG Index, les entreprises qui seront sélectionnées dans l'indice l'année suivante affichent déjà un meilleur profil ESG que les entreprises du groupe de contrôle. En effet, elles affichent un ESGC moyen de 53 contre 47 pour le groupe de contrôle et surtout un Perf_ESG moyen de 0,25 contre -0,27. La distribution n'est donc pas parfaitement homogène entre les deux groupes, même avant le lancement de l'indice. Il faut aussi noter la forte dispersion de la variable GHG_INT (SD=448) indiquant une différence

importante d'intensité carbone entre les différents secteurs. Ce constat justifie l'inclusion d'effets fixes secteur dans nos futures régressions.

4.4.2. Période d'analyse (2023-2024)

Sur notre période d'analyse, l'échantillon est composé de 80 observations dont 37 avec BEL_ESG = 1 et 43 avec BEL_ESG = 0.

Le tableau ci-dessous résume les statistiques descriptives pour chaque groupe.

Variable	Groupe	Observations	Moyenne	Médiane	SD	Min	Max
ESG_LING	Hors BEL ESG	43	343,904	347,857	90,935	191,152	580,775
	BEL ESG	37	377,133	365,398	103,344	159,076	659,790
ESGC	Hors BEL ESG	43	48,445	47,8	10,956	21,46	69,22
	BEL ESG	37	53,770	53,44	8,760	33,39	68,915
GHG_INT	Hors BEL ESG	43	183,262	16,91	428,884	0,23	1949,58
	BEL ESG	37	68,333	9,86	209,789	0,03	1228,71
Perf_ESG	Hors BEL ESG	43	-0,2	-0,165	0,641	-1,444	0,829
	BEL ESG	37	0,281	0,475	0,537	-0,985	1,087

Premièrement, les entreprises appartenant au BEL ESG Index présentent un ESG Linguistic Score moyen de 377 occurrences pour 10 000 mots contre 344 pour les entreprises du groupe de contrôle. Cet écart de 33 mots suggère que les entreprises du BEL ESG Index consacrent une partie plus importante de leur rapport annuel aux thématiques ESG. Cependant, la dispersion est élevée dans les deux groupes (écart-type de 103 pour le groupe BEL ESG et 91 pour le groupe de contrôle), indiquant des pratiques de communication différente au sein de chaque groupe. Cette première observation est tout à fait cohérente avec la prédiction de H2 sans pour autant être significative à ce stade.

Deuxièmement, l'ESG Combined Score moyen des entreprises incluses dans le BEL ESG s'élève à 53,8 points contre 48,4 pour le groupe de contrôle soit une différence de 5,3 points. Cet écart est totalement cohérent avec les résultats obtenus par Kolling et Busch, qui observent un écart situé entre +3,78 et +4,75 points d'ESGC pour les fonds labellisés. Le fait de se retrouver dans le même intervalle sur notre échantillon d'entreprises belges représente un signal en faveur de l'hypothèse H1a.

Troisièmement, on observe une intensité carbone moyenne pour les entreprises du BEL ESG de 68 tonnes de CO₂ par million d'euros de chiffre d'affaires contre 183 tonnes pour le groupe de contrôle. Cette différence de -115 va dans le sens de notre hypothèse H1b. Cependant, il est important d'interpréter cette différence avec prudence. En effet, la dispersion est très élevée dans les deux groupes, reflétant la présence d'entreprises très émettrices notamment dans le secteur industriel. Cette hétérogénéité nécessite l'introduction d'effets fixes secteur dans les régressions pour isoler l'effet de l'appartenance au BEL ESG de l'effet lié au secteur d'activité. Enfin, l'indice de performance Perf_ESG présente la performance la plus marquée entre nos deux groupes. Enfin, l'indice composite Perf_ESG confirme ce que montrent déjà l'ESGC et l'intensité carbone pris séparément : les entreprises de l'indice se situent en moyenne au-dessus de la moyenne de l'échantillon (+0,28), tandis que les entreprises non incluses se situent en dessous (-0,20). Cet écart s'explique par le fait que Perf_ESG capte simultanément les deux dimensions de la performance durable.

Ces observations ne permettent pas, à elles seules, de conclure à des différences statistiquement significatives entre les deux groupes. Sans contrôle sur le secteur d'activité et l'année, les écarts observés peuvent s'expliquer par des différences structurelles entre les groupes. Les régressions présentées dans la section 5 viendront apporter une réponse plus rigoureuse.

4.5. Modèles de régression

Nos quatre hypothèses sont testées à l'aide de régressions linéaires MCO en panel groupé (Pooled Ordinary Least Squares), estimées sur un panel de 40 entreprises observées sur trois années (2022, 2023 et 2024). Le choix des MCO est justifié par la nature des variables dépendantes (ESGC, GHG_INT et ESG_LING).

4.5.1. Spécifications des modèles

Pour chaque hypothèse, trois spécifications progressives sont estimées, cette démarche est directement inspirée de l'article de référence de Kolling et Busch. Le premier modèle est estimé sans variable de contrôle, le deuxième intègre une variable catégorielle d'années (année 2022 comme référence) afin de neutraliser les chocs temporels communs aux entreprises comme la montée en puissance de la CSRD. Enfin, le troisième modèle ajoute une variable catégorielle de secteur (secteur financier comme référence) pour neutraliser les différences d'intensité ESG

et carbone entre secteurs ainsi qu'une variable de contrôle continue (taille) correspondant au logarithme de la capitalisation boursière de l'entreprise i à l'année t .

Les quatre modèles finaux estimés sont les suivants :

H1a :
$$ESGC_{it} = \alpha + \beta_1 BEL_ESG_{it} + \sum \beta_{2k} Année_k, t + \sum \beta_{3j} Secteur_j, i + \beta_4 Taille_{it} + \epsilon_{it}$$

Le coefficient β_1 teste si l'appartenance au BEL ESG est associé à un ESGC significativement plus élevé.

H1b :
$$\ln(1 + GHG_{INT_{it}}) = \alpha + \beta_1 BEL_ESG_{it} + \sum \beta_{2k} Année_k, t + \sum \beta_{3j} Secteur_j, i + \beta_4 Taille_{it} + \epsilon_{it}$$

Un coefficient β_1 négatif et significatif confirmerait H1b.

H2 :
$$ESG_LING_{i,t} = \alpha + \beta_1 BEL_ESG_{it} + \sum \beta_{3k} Année_k, t + \sum \beta_{3j} Secteur_j, i + \beta_4 Taille_{it} + \epsilon_{it}$$

Le coefficient β_1 teste si les entreprises appartenant au BEL ESG utilisent davantage de mots ESG dans leurs rapports que leurs pairs.

H3 :
$$ESG_LING_{i,t} = \alpha + \beta_1 Perf_ESG_{it} + \beta_2 BEL_ESG_{it} + \sum \beta_{3k} Année_k, t + \sum \beta_{4j} Secteur_j, i + \beta_5 Taille_{it} + \epsilon_{it}$$

La variable $Perf_ESG_{it}$ vient neutraliser l'effet de la performance réelle afin que β_2 capte exclusivement l'excès de communication qui n'est pas expliqué par la performance.

5. Résultats

5.1. Test de H1a – ESG Combined Score (ESGC)

Pour rappel, l'hypothèse H1a vient tester si les entreprises appartenant au BEL ESG présentent un ESGC significativement supérieur au groupe de contrôle. Le tableau ci-dessous reprend les résultats des trois spécifications.

	Variable dépendante : ESGC		
	(1) Sans contrôles	(2) + Année	(3) + Secteur et Taille
BEL_ESG	4.798**	5.245**	2.822
ln_cap			9.489***
year_f2023		-1.588	-0.664
year_f2024		-0.525	1.309
secteur_fConsommation			1.992
secteur_fimmobilier			11.211***
secteur_findustriel			-2.210
secteur_fMateriaux_de_base			18.435***
secteur_fSante			-0.047
secteur_fServices_Publics			5.096*
secteur_fTechnologie			12.223***
secteur_fTelecommunication			12.611***
Observations	120	120	120
R2	0.045	0.049	0.510
Adjusted_R2	0.037	0.024	0.455

Le modèle sans variables de contrôle (1) présente un coefficient positif et significatif au seuil de 5% ($\beta = 4,798$; $p = 0,01238$). Les entreprises du BEL ESG affichent un ESGC supérieur d'environ 4,8 points à celui des entreprises du groupe de contrôle. L'écart est directement comparable à celui identifié par Kolling et Busch (entre +3,78 et +4,75 points).

Le deuxième modèle (2), avec effets fixes d'années, confirme la significativité ($\beta = 5,245$; $p = 0,02229$). Le coefficient associé à notre variable BEL ESG reste positif et significatif indiquant que les différences observées ne s'expliquent pas par un effet de période. Cependant, lorsque l'on intègre à notre dernier modèle (3) la taille et les effets fixes secteurs, le coefficient chute à 2,822 et n'est plus significatif ($p = 0,145$). Cette détérioration du coefficient suggère qu'une part importante de l'écart observé dans nos deux précédents modèles est expliquée par des caractéristiques structurelles des entreprises plutôt que par leur appartenance à l'indice BEL ESG.

Parmi les variables de contrôle, la taille de l'entreprise a un impact positif et significatif ($\beta = 9,489$; $p < 0,01$), suggérant que les entreprises de plus grande taille ont en moyenne des ESG Combined Score plus élevés. Plusieurs secteurs ont également des coefficients positifs significatifs comme les matériaux de base ($\beta = 18,435$; $p < 0,01$), les télécommunications ($\beta = 12,611$; $p < 0,01$), la technologie ($\beta = 12,223$; $p < 0,01$) ou encore l'immobilier ($\beta = 11,211$;

$p < 0,01$). L'ajout de ces variables permet d'améliorer considérablement le pouvoir explicatif du modèle. En effet, le R^2 passe de 4,5% dans le premier modèle à 51% dans le modèle principal.

L'hypothèse H1a n'est donc pas confirmée. Les niveaux plus élevés d'ESGC pour les entreprises du BEL ESG sont attribués à leur profil structurel plutôt qu'à leur appartenance à l'indice.

5.2. Test de H1b – Intensité carbone GHG_INT

L'hypothèse H1b teste si les entreprises appartenant au BEL ESG ont une intensité carbone plus faible que celles non incluses dans l'indice. Le tableau ci-dessous reprend les résultats des trois spécifications.

	Variable dépendante : $\ln(\text{GHG_INT})$		
	(1) Sans contrôles	(2) + Année	(3) + Secteur et Taille
BEL_ESG	-0.702*	-0.906**	-0.353
$\ln_cap$			-1.081***
year_f2023		0.349	0.126
year_f2024		0.468	0.121
secteur_fConsommation			0.464
secteur_fImmobilier			0.653
secteur_fIndustriel			2.348**
secteur_fMateriaux_de_base			3.433***
secteur_fSante			0.580
secteur_fServices_Publics			3.940***
secteur_fTechnologie			0.436
secteur_fTelecommunication			0.189
Observations	120	120	120
R^2	0.027	0.035	0.494
Adjusted_R2	0.019	0.010	0.437

Le modèle (1), sans variables de contrôle, présente un coefficient négatif significatif au seuil de 10% ($\beta = -0,702$; $p = 0,061$). Ce coefficient indique que les entreprises appartenant au BEL ESG ont une intensité carbone plus faible que celles n'appartenant pas à l'indice.

Le modèle (2), qui contrôle les différences entre années, confirme la relation du premier modèle et rend le coefficient significatif au seuil de 5% ($\beta = -0,906$; $p = 0,042$).

En revanche, on observe dans le modèle principal (3), qui contrôle les effets fixes d'années, de secteurs et la taille des entreprises, un coefficient associé à BEL_ESG qui diminue et perd sa significativité statistique ($\beta = -0,353$; $p = 0,333$).

L'hypothèse H1b n'est donc pas confirmée.

Ce résultat s'explique par l'introduction des variables de contrôle. On voit que certains secteurs sont beaucoup plus émetteurs que d'autres comme le secteur des matériaux de base ($\beta = 3,433$; $p < 0,01$), des services publics ($\beta = 3,940$; $p < 0,01$) ou encore de l'industrie ($\beta = 2,348$; $p < 0,05$). Ces secteurs présentent une intensité carbone significativement plus élevée que le secteur de référence (financier).

De plus, la taille de l'entreprise présente un coefficient négatif et fortement significatif ($\beta = -1,081$; $p < 0,01$) ce qui indique que les entreprises de grandes tailles ont, en moyenne, une intensité carbone plus faible. L'ajout des variables de contrôle améliore le pouvoir explicatif du modèle en passant d'un R2 de 2,7% dans le modèle sans contrôle à 49,4% dans le dernier modèle.

Nos résultats sont en adéquation avec ceux de Kolling et Busch qui ont observé que l'intensité carbone entre entreprises labellisées et non labellisées s'atténuait une fois que les caractéristiques structurelles des entreprises étaient prises en compte.

5.3. Test de H2 – Intensité du discours ESG (ESG_LING)

L'hypothèse H2 teste si les entreprises du BEL ESG utilisent davantage de termes appartenant au lexique ESG dans leurs rapports annuels que le groupe de contrôle.

	Variable dépendante : ESG_LING		
	(1) Sans contrôles	(2) + Année	(3) + Secteur et Taille
BEL_ESG	42.989**	29.832	36.497*
ln_cap			15.730
year_f2023		-0.705	-3.712
year_f2024		44.348**	42.047**
secteur_fConsommation			45.405
secteur_fimmobilier			23.916
secteur_findustriel			79.472***
secteur_fMateriaux_de_base			124.412***
secteur_fSante			112.994***
secteur_fServices_Publics			79.085**
secteur_fTechnologie			159.127***
secteur_fTelecommunication			50.133
Observations	120	120	120
R2	0.049	0.099	0.413
Adjusted_R2	0.040	0.075	0.347

Le modèle (1) sans contrôles livre un coefficient significatif au seuil de 5%

($\beta = 42,99$; $p = 0,029$). Ce coefficient signifie que les entreprises incluses dans le BEL ESG utilisent, en moyenne, 43 termes ESG supplémentaires pour 10 000 mots dans leur rapport annuel par rapport aux entreprises du groupe de contrôle.

Dans le modèle (2), intégrant les effets fixes d'années, on obtient un coefficient associé à BEL_ESG de 29,83 qui n'est plus statistiquement significatif ($p = 0,174$). Cette diminution suggère qu'une partie de la différence que nous avons observé dans le premier modèle est liée à des changements communs apparus au cours de la période étudiée. Si on regarde le coefficient de l'effet fixe de l'année 2024, on voit qu'il est positif et significatif au seuil de 5% ($\beta = 44,35$; $p = 0,040$) indiquant qu'il y a eu une augmentation moyenne d'utilisation du vocabulaire ESG durant cette année. Cette augmentation peut être expliquée par la hausse des exigences de la CSRD.

Enfin, le dernier modèle (3) qui contrôle les effets fixes d'années, de secteur et la taille des entreprises nous livre un coefficient associé à la variable BEL_ESG de 36,5. Ce coefficient est positif mais significatif qu'à un seuil de 10% ($p = 0,082$).

L'hypothèse H2 est donc partiellement confirmée. A taille, secteur et année comparables, les entreprises du BEL ESG utilisent en moyenne davantage de vocabulaire ESG dans leurs rapports. Cependant, cet effet reste statistiquement limité. Ce résultat fait directement lien avec la théorie de la légitimité précédemment identifiée dans la revue de littérature selon laquelle les entreprises détentrices d'un label améliorent leur communication ESG afin de répondre aux attentes de leurs différentes parties prenantes.

5.4. Test de H3 – Communication excessive (modèle conditionnel)

Notre troisième et dernière hypothèse vise à déterminer si, en contrôlant statistiquement la performance ESG réelle, les entreprises du BEL ESG communiquent davantage sur les enjeux ESG que les entreprises non incluses dans l'indice.

	Variable dépendante : ESG_LING		
	(1) Sans contrôles	(2) + Année	(3) + Secteur et Taille
Perf_ESG	14.761	15.967	36.250**
BEL_ESG	36.983*	22.171	28.369
ln_cap			-10.614
year_f2023		1.916	-1.406
year_f2024		46.638**	40.885**
secteur_fConsommation			46.204
secteur_fImmobilier			10.471
secteur_fIndustriel			104.812***
secteur_fMateriaux_de_base			123.908***
secteur_fSante			118.388***
secteur_fServices_Publics			106.345**
secteur_fTechnologie			141.934***
secteur_fTelecommunication			30.006
Observations	120	120	120
R2	0.059	0.110	0.442
Adjusted_R2	0.043	0.079	0.373

Le premier modèle (1), constitué uniquement de la performance ESG et de l'appartenance au BEL ESG, ne montre aucune relation statistiquement significative entre la performance ESG et l'intensité du discours ESG ($\beta = 14,76$; $p = 0,212$). Le fait d'introduire les effets fixes d'années dans le modèle (2) ne change pas cette conclusion ($\beta = 15,97$; $p = 0,187$).

Cependant, dans notre modèle principal (3), le coefficient de Perf_ESG devient positif et significatif au seuil de 5% ($\beta = 36,25$; $p = 0,032$). Ce résultat montre qu'en contrôlant statistiquement les autres caractéristiques du modèle, les entreprises qui affichent une meilleure performance ESG tendent à utiliser plus de vocabulaire ESG dans leurs rapports annuels.

Pour ce qui est du coefficient de la variable BEL_ESG, celui-ci reste positif mais n'est pas statistiquement significatif ($\beta = 28,37$; $p = 0,165$).

L'hypothèse H3 n'est donc pas confirmée. En contrôlant la performance ESG, la taille, le secteur et l'année, les entreprises incluses dans le BEL ESG n'ont pas une communication significativement plus élevée que les entreprises du groupe de contrôle.

Plusieurs variables de contrôle sont significatives dans notre modèle complet. L'année 2024 est synonyme d'augmentation générale du vocabulaire ESG ($\beta = 40,89$; $p = 0,049$). Certains secteurs sortent un peu du lot et présentent une intensité de communication significativement plus élevée que le secteur de référence. Parmi ces secteurs on retrouve le secteur technologique ($\beta = 141,93$; $p < 0,01$), les matériaux de base ($\beta = 123,91$; $p < 0,01$), la santé ($\beta = 118,39$; p

< 0,01), les services publics ($\beta = 106,35$; $p < 0,01$) ou encore l'industrie ($\beta = 104,81$; $p < 0,01$),

Enfin, l'ajout des variables de contrôle améliore le pouvoir explicatif du modèle puisque le R2 passe de 5% dans le premier modèle à 44% dans le modèle (3).

5.5. Synthèse des résultats

	Tableau Synthétique			
	H1a : ESGC	H1b : $\ln(\text{GHG_INT})$	H2 : ESG_LING	H3 : ESG_LING Cond.
Perf_ESG				36.250**
BEL_ESG	2.822	-0.353	36.497*	28.369
ln_cap	9.489***	-1.081***	15.730	-10.614
year_f2023	-0.664	0.126	-3.712	-1.406
year_f2024	1.309	0.121	42.047**	40.885**
secteur_fConsommation	1.992	0.464	45.405	46.204
secteur_fimmobilier	11.211***	0.653	23.916	10.471
secteur_findustriel	-2.210	2.348**	79.472***	104.812***
secteur_fMateriaux_de_base	18.435***	3.433***	124.412***	123.908***
secteur_fSante	-0.047	0.580	112.994***	118.388***
secteur_fServices_Publics	5.096*	3.940***	79.085**	106.345**
secteur_fTechnologie	12.223***	0.436	159.127***	141.934***
secteur_fTelecommunication	12.611***	0.189	50.133	30.006
Observations	120	120	120	120
R2	0.510	0.494	0.413	0.442
Adjusted_R2	0.455	0.437	0.347	0.373

Ce tableau résume les différents résultats obtenus pour chaque hypothèse. La première chose frappante est qu'aucune de nos quatre hypothèses n'est confirmée au seuil de 5% dans nos spécifications principales. Notre deuxième hypothèse est la seule qui est partiellement confirmée au seuil de 10%. Nos modèles sans contrôles nous livrent des résultats significatifs, mais cette significativité s'atténue une fois l'introduction des effets fixes secteur et de la taille. Ces résultats ne signifient pas qu'ils n'existent aucun effet entre nos variables étudiées mais la puissance statistique limitée (120 observations pour 37 avec BEL_ESG =1) peut expliquer ce manque de significativité.

6. Discussion des résultats

Plusieurs enseignements peuvent être tirés de ce mémoire. Dans nos modèles les plus simples sans variables de contrôle (1), on observe des différences significatives entre les entreprises du BEL ESG et celles non incluses dans l'indice pour H1a, H1b et H2. Cependant, l'introduction progressive des effets temporels, des effets fixes sectoriels ou encore de la taille des entreprises entraîne une diminution de ces effets voir même à leur disparition dans nos modèles principaux (3). Ce constat est un enseignement en lui-même : les différences apparues dans nos premiers modèles (1) entre les entreprises du BEL ESG et leurs pairs ne sont pas associées à l'appartenance à l'indice mais bien à des caractéristiques structurelles comme la taille ou le secteur de l'entreprise. La simple appartenance à l'indice n'est pas capable d'identifier une performance ESG supérieure, indépendamment des caractéristiques structurelles des entreprises qui y sont incluses.

Les coefficients positifs et significatifs de l'année 2024 observés dans notre deuxième et troisième hypothèse sont également instructifs. On observe une augmentation flagrante du nombre de termes ESG utilisés dans les rapports des entreprises de l'échantillon de 2024. Cette augmentation est à mettre en perspective avec la montée en puissance du cadre réglementaire européen et des exigences associées à la CSRD.

Notre étude a des implications directes pour les investisseurs. Les résultats obtenus invitent à la prudence concernant la valeur informationnelle de l'indice BEL ESG en tant que signal de durabilité. En effet, comme l'appartenance à l'indice ne garantit pas une performance ESG significativement plus élevée une fois les caractéristiques structurelles prises en comptes, le caractère durable de l'entreprise mérite d'être étudiée sur une plus longue période.

Pour Euronext, nos résultats remettent en doute la méthodologie de sélection de l'indice et sa capacité à identifier les entreprises performantes sur les critères ESG indépendamment des caractéristiques structurelles des entreprises.

Enfin, pour les décideurs politiques et les différentes parties prenantes qui s'intéressent à la transparence ESG, l'absence de confirmation de H3 nous livre un élément de réponse. Dans notre échantillon, une fois la performance ESG, la taille, le secteur et l'année d'observation contrôlés, les entreprises incluses dans l'indice ne présentent pas de surplus de communication ESG statistiquement significatif. Le fait d'appartenir à l'indice n'est donc pas synonyme de sur-communication systématique. Cependant, les résultats ne permettent pas, à eux-seuls, d'exclure toute forme de greenwashing puisque la mesure utilisée se concentre sur l'intensité

du discours ESG et non sur la qualité et la véracité des informations communiquées par les entreprises.

7. Limites de l'étude

Quatre limites doivent tout de même être soulignées.

Premièrement, notre échantillon n'est constitué que de 120 observations dont 37 correspondent à des entreprises appartenant au BEL ESG. La taille restreinte de celui-ci limite la précision de nos estimations notamment une fois l'introduction de plusieurs variables de contrôle et d'effets fixes.

Deuxièmement, la période d'observation s'étend de 2022 à 2024. Le BEL ESG ayant été lancé par Euronext en février 2023, seules deux années post lancement peuvent être observées.

Troisièmement, la métrique textuelle utilisée pour mesurer la communication ESG (ESG_LING) capture l'intensité du discours ESG et non la véracité, la pertinence ou la qualité des informations communiquées dans les rapports annuels.

Enfin, nous avons utilisé un fournisseur unique de données ESG. Les scores ESG peuvent varier en fonction des méthodologies et des critères appliqués par les différents fournisseurs. Les résultats obtenus dans notre étude sont à interprétés en prenant en compte la méthodologie appliquée par Bloomberg.

8. Conclusion

Ce mémoire avait pour but d'évaluer la cohérence du BEL ESG Index en tant que signal de durabilité à travers quatre hypothèses : la performance ESG supérieure des entreprises appartenant à l'indice (H1a et H1b), une intensité du vocabulaire ESG également plus élevée (H2) et enfin l'existence d'une sur-communication relative à leur performance (H3). La méthodologie utilisée s'inspire directement des travaux de Kolling et Busch en intégrant une analyse textuelle des rapports annuels.

Les résultats obtenus sont nuancés. En effet, aucune de nos quatre hypothèses n'est confirmée au seuil de 5% dans le modèle principal. L'hypothèse H2 est partiellement confirmée au seuil de 10% indiquant une tendance des entreprises appartenant au BEL ESG à communiquer davantage sur les thématiques ESG que leurs pairs. Les spécifications sans contrôle sont significatives mais cette significativité s'estompe une fois l'introduction des effets fixes secteur et de la taille. Les différences observées entre les entreprises incluses dans l'indice et celles non

incluses s'expliquent en grande partie par leurs caractéristiques structurelle (secteur, taille) plutôt que par un effet propre de l'appartenance à l'indice.

L'absence de confirmation de H3 est un enseignement en lui-même. Dans notre échantillon, les entreprises du BEL ESG ne pratiquent pas de communication excessive systématique compte tenu de leur performance observable.

Nos résultats doivent être interprétés avec des pincettes à cause de la taille restreinte de notre échantillon. L'absence de significativité statistique n'est pas une preuve de l'absence d'effet et Kolling et Busch disposaient de 150 fonds, leur donnant une puissance statistique nettement supérieure à la nôtre.

Plusieurs pistes de recherche future sont possibles. La première serait d'enrichir l'historique du BEL ESG au fil des années afin d'étoffer l'échantillon. Ensuite, essayer de capturer la qualité et non seulement l'intensité du discours à l'aide des approches NLP (Traitement Automatique du Langage Naturel) plus avancées. Enfin, utiliser un fournisseur comme Sustainalytics afin de croiser les scores ESG avec ceux de Bloomberg. Ce choix est d'autant plus pertinent que la composition du BEL ESG Index repose sur les scores proposés par Sustainalytics.

Ce mémoire contribue à la littérature sur la cohérence des labels ESG en appliquant une double approche : performance et discours à un indice d'actions là où la littérature s'était attardée sur les fonds d'investissement.

Annexes

Annexe A – Codage dynamique BEL_ESG

Lancement	mars-23	juin-23	sept-23	déc-23	mars-24	juin-24	sept-24	déc-24	mars-25	juin-25	sept-25	déc-25
AB Inbev	AB Inbev	AB Inbev	AB Inbev	AB Inbev	AB Inbev	AB Inbev	AB Inbev	AB Inbev	AB Inbev	AB Inbev	AB Inbev	AB Inbev
Ackerman s & van Haaren	Ackerman s & van Haaren	Ackerman s & van Haaren	Ackerman s & van Haaren	Ackerman s & van Haaren	Ackerman s & van Haaren	Ackerman s & van Haaren	Ackerman s & van Haaren	Ackerman s & van Haaren	Ackerman s & van Haaren	Ackerman s & van Haaren	Ackerman s & van Haaren	Ackerman s & van Haaren
Aedifica	Aedifica	Aedifica	Aedifica	Aedifica	Aedifica	Aedifica	Aedifica	Aedifica	Aedifica	Aedifica	Aedifica	Aedifica
Aperam	Aperam	Aperam	Aperam	Aperam	Aperam	Aperam	Aperam	Aperam	Aperam	Aperam	Aperam	Aperam
Barco	Barco	Barco	Barco	Barco	Barco	Barco	Barco	Barco	Barco	Barco	Barco	Barco
Bekaert	Bekaert	Bekaert	Bekaert	Bekaert	Bekaert	Bekaert	Bekaert	Bekaert	Bekaert	Bekaert	Bekaert	Bekaert
Bpost	Bpost	Bpost	Bpost	Bpost	Bpost	Bpost	Bpost	Bpost	Bpost	Bpost	Bpost	Bpost
Cofinimmo	Cofinimmo	Cofinimmo	Cofinimmo	Cofinimmo	Cofinimmo	Cofinimmo	Cofinimmo	Cofinimmo	Cofinimmo	Cofinimmo	Cofinimmo	Cofinimmo
Elia Group	Elia Group	Elia Group	Elia Group	Elia Group	Elia Group	Elia Group	Elia Group	Elia Group	Elia Group	Elia Group	Elia Group	Elia Group
Fagron	Fagron	Fagron	Fagron	Fagron	Fagron	Fagron	Fagron	Fagron	Fagron	Fagron	Fagron	Fagron
GBL	GBL	GBL	GBL	GBL	GBL	GBL	GBL	GBL	GBL	GBL	GBL	GBL
KBC	KBC	KBC	KBC	KBC	KBC	KBC	KBC	KBC	KBC	KBC	KBC	KBC
Lotus	Lotus	Lotus	Lotus	Lotus	Lotus	Lotus	Lotus	Lotus	Lotus	Lotus	Lotus	Lotus
Bakeries	Bakeries	Bakeries	Bakeries	Bakeries	Bakeries	Bakeries	Bakeries	Bakeries	Bakeries	Bakeries	Bakeries	Bakeries
Proximus	Proximus	Proximus	Proximus	Proximus	Proximus	Proximus	Proximus	Proximus	Proximus	Proximus	Proximus	Proximus
Shurgard	Shurgard	Shurgard	Shurgard	Shurgard	Shurgard	Shurgard	Shurgard	Shurgard	Shurgard	Shurgard	Shurgard	Shurgard
Sofina	Sofina	Sofina	Sofina	Sofina	Sofina	Sofina	Sofina	Sofina	Sofina	Sofina	Sofina	Sofina
Solvay	Solvay	Solvay	Solvay	Solvay	Solvay	Solvay	Solvay	Solvay	Solvay	Solvay	Solvay	Solvay
Telenet	Telenet	Telenet	Telenet	Telenet	Telenet	Telenet	Telenet	Telenet	Telenet	Telenet	Telenet	Telenet
UCB	UCB	UCB	UCB	UCB	UCB	UCB	UCB	UCB	UCB	UCB	UCB	UCB
WDP	WDP	WDP	WDP	WDP	WDP	WDP	WDP	WDP	WDP	WDP	WDP	WDP
	Gim V	Gim V	Gim V	Gim V	Gim V	Gim V	Gim V	Gim V	Gim V	Gim V	Gim V	Gim V
	Ontex	Ontex	Ontex	Ontex	Ontex	Ontex	Ontex	Ontex	Ontex	Ontex	Ontex	Ontex
	VGP	VGP	VGP	VGP	VGP	VGP	VGP	VGP	VGP	VGP	VGP	VGP
			Montea	Montea	Montea	Montea	Montea	Montea	Montea	Montea	Montea	Montea
				Azelis	Azelis	Azelis	Azelis	Azelis	Azelis	Azelis	Azelis	Azelis
				Intervest	Intervest	Intervest	Intervest	Intervest	Intervest	Intervest	Intervest	Intervest
				Off-ware	Off-ware	Off-ware	Off-ware	Off-ware	Off-ware	Off-ware	Off-ware	Off-ware
				XIOR	XIOR	XIOR	XIOR	XIOR	XIOR	XIOR	XIOR	XIOR
					Ageas	Ageas	Ageas	Ageas	Ageas	Ageas	Ageas	Ageas
					EVS	EVS	EVS	EVS	EVS	EVS	EVS	EVS
					Broadcast	Broadcast	Broadcast	Broadcast	Broadcast	Broadcast	Broadcast	Broadcast
	Appartient à l'indice					Colruyt	Colruyt	Colruyt	Colruyt	Colruyt	Colruyt	Colruyt
	Sorti de l'indice					D'ieteren Group	D'ieteren Group	D'ieteren Group	D'ieteren Group	D'ieteren Group	D'ieteren Group	D'ieteren Group
	Intègre l'indice							Recticel	Recticel	Recticel	Recticel	Recticel
								Umicore	Umicore	Umicore	Umicore	Umicore

2023	2024	2025
AB Inbev	AB Inbev	AB Inbev
Ackermans & van Haaren	Ackermans & van Haaren	Ackermans & van Haaren
Aedifica	Aedifica	Aedifica
Aperam	Aperam	Aperam
Barco	Barco	Barco
Bekaert	Bekaert	Bekaert
Bpost	Bpost	Bpost
Cofinimmo	Cofinimmo	Cofinimmo
Elia Group	Elia Group	Elia Group
Fagron	Fagron	Fagron
GBL	GBL	GBL
KBC Group	KBC Group	KBC Group
Proximus	Proximus	Proximus
Shurgard	Shurgard	Shurgard
Sofina	Sofina	Sofina
Solvay	Solvay	Solvay
Telenet	Telenet	Telenet
UCB	UCB	UCB
WDP	WDP	WDP
Gim V	Gim V	Gim V
Ontex Group	Ontex Group	Ontex Group
VGP	VGP	VGP
	Azelis Group	Azelis Group
	XIOR	XIOR
	Ageas	Ageas
	Montea	Montea
		D'ieteren Group
		Umicore

Annexe B – Script extraction pdf en fichier texte

```
import os
import re
import csv
import unicodedata
from pathlib import Path
```

```
try:
    import pdfplumber
except ImportError:
    raise ImportError("Installe pdfplumber : pip install pdfplumber")
```

```
INPUT_DIR = r"C:\Users\bouil\OneDrive\Documents\Master 2\Mémoire\Rapports annuels
hors BEL ESG" # Dossier contenant tes PDF
OUTPUT_DIR = r"C:\Users\bouil\OneDrive\Documents\Master 2\Mémoire\Rapports txt
Hors BEL ESG" # Dossier où seront écrits les TXT
LOG_FILE = r"./extraction_log.csv"
```

```
# Seuil d'alerte : si un fichier TXT a moins de MIN_WORDS mots,
# il est marqué comme suspect dans le log
MIN_WORDS = 500
```

```
#
```

```
# NETTOYAGE DU TEXTE
```

```
#
```

```
# Lignes à supprimer si elles correspondent exactement à ces patterns.
```

```
# Supprime les en-têtes, pieds de page, numéros de page, et ToC
```

```
# qui gonflent artificiellement le dénominateur de ESG_LING.
```

```
NOISE_PATTERNS = [
```

```
    r"^\s*\d+\s*$",          # Numéro de page seul
    r"^\s*page\s+\d+\s*(of|sur|/)\s*\d+", # "Page 3 of 120"
    r"^\s*table\s+of\s+contents?\s*$",    # Table of contents
    r"^\s*table\s+des\s+mati.res\s*$",    # Table des matières
    r"^\s*inhoudstafel\s*$",              # Inhoudstafel (NL)
    r"^\s*annual\s+report\s+\d{4}\s*$",   # "Annual Report 2023"
    r"^\s*rapport\s+annuel\s+\d{4}\s*$",  # "Rapport annuel 2023"
    r"^\s*jaarverslag\s+\d{4}\s*$",      # "Jaarverslag 2023" (NL)
    r"^\s*\.{3,}\s*\d+\s*$",             # "..... 45" (ToC)
    r"^\s*[\—]{3,}\s*$",                 # Lignes de séparation
    r"^\s*copyright\s*[\©®]",            # Copyright
    r"^\s*confidential",                  # Mentions confidentielles
```

```
]
```

```
NOISE_RE = [re.compile(p, re.IGNORECASE) for p in NOISE_PATTERNS]
```

```
def is_noise_line(line: str) -> bool:
```

```
    """Retourne True si la ligne est du bruit à supprimer."""
```

```
    for pattern in NOISE_RE:
```

```
        if pattern.match(line):
```

```
            return True
```

```
    return False
```

```
def normalize_text(text: str) -> str:
```

```
    """
```

```
    Nettoyage du texte extrait :
```

```
    1. Normalise les caractères Unicode (é → e si problème d'encodage)
```

```
    2. Supprime les lignes bruit (en-têtes, pieds de page, numéros)
```

```
    3. Réduit les espaces et lignes vides multiples
```

```
    4. Met tout en minuscules (nécessaire pour le comptage dictionnaire)
```

```
    """
```

```
    # 1. Normalise unicode (NFC = forme canonique composée)
```

```
    text = unicodedata.normalize("NFC", text)
```

```
    # 2. Traite ligne par ligne
```

```
    lines = text.splitlines()
```

```
    clean_lines = []
```

```
    for line in lines:
```

```

        stripped = line.strip()
        if not stripped:
            continue
        if is_noise_line(stripped):
            continue
        # Supprime les tokens qui ne sont que des symboles/chiffres isolés
        if re.match(r"^\[d\W]+\$", stripped) and len(stripped) < 5:
            continue
        clean_lines.append(stripped)

# 3. Réduit les sauts de ligne multiples (garde max 2 lignes vides)
joined = "\n".join(clean_lines)
joined = re.sub(r"\n{3,}", "\n\n", joined)

# 4. Minuscules
joined = joined.lower()

# 5. Supprime les caractères non-texte (garde lettres, chiffres, ponctuation)
joined = re.sub(r"^[^w\s\-\.,;:!?\"()/'&@%]", " ", joined)

# 6. Normalise les espaces multiples
joined = re.sub(r"[ \t]{2,}", " ", joined)

return joined.strip()

#


---




---


# EXTRACTION PDF
#


---




---



def extract_pdf_to_text(pdf_path: Path) -> tuple[str, dict]:
    """
    Extrait le texte d'un PDF avec pdfplumber.

    Retourne :
        text : texte nettoyé (str)
        stats : dict avec nb_pages, nb_words, nb_words_raw, warnings
    """
    stats = {
        "nb_pages": 0,
        "nb_pages_extracted": 0,
        "nb_words_raw": 0,
        "nb_words_clean": 0,
        "warnings": [],
    }

    raw_pages = []

```

```

try:
    with pdfplumber.open(pdf_path) as pdf:
        stats["nb_pages"] = len(pdf.pages)

        for i, page in enumerate(pdf.pages):
            try:
                # Extrait texte de la page
                page_text = page.extract_text(
                    x_tolerance=2,
                    y_tolerance=2,
                    layout=True,      # préserve mise en page
                    x_density=7.25,
                    y_density=13,
                )
                if page_text:
                    raw_pages.append(page_text)
                    stats["nb_pages_extracted"] += 1
            except Exception as e:
                stats["warnings"].append(f"Page {i+1} erreur: {str(e)[:60]}")

except Exception as e:
    stats["warnings"].append(f"Erreur ouverture PDF: {str(e)[:100]}")
    return "", stats

raw_text = "\n".join(raw_pages)
stats["nb_words_raw"] = len(raw_text.split())

# Nettoyage
clean_text = normalize_text(raw_text)
stats["nb_words_clean"] = len(clean_text.split())

return clean_text, stats

```

#

PROGRAMME PRINCIPAL

#

```

def main():
    input_dir = Path(INPUT_DIR)
    output_dir = Path(OUTPUT_DIR)
    output_dir.mkdir(parents=True, exist_ok=True)

    # Recherche tous les PDFs dans le dossier d'entrée
    pdf_files = sorted(input_dir.glob("*.pdf"))

```

```

if not pdf_files:
    print(f"\n⚠️ Aucun PDF trouvé dans : {input_dir.resolve()}")
    print(" Vérifie que tes fichiers sont dans ce dossier.")
    return

print(f"\n{'='*60}")
print(f" Extraction de {len(pdf_files)} fichiers PDF")
print(f" Source : {input_dir.resolve()}")
print(f" Sortie : {output_dir.resolve()}")
print(f"{'='*60}\n")

log_rows = []

for pdf_path in pdf_files:
    stem = pdf_path.stem # ex: "ACKB_2022"
    out_path = output_dir / f"{stem}.txt"

    print(f" 📄 {pdf_path.name:<30}", end=" ", flush=True)

    text, stats = extract_pdf_to_text(pdf_path)

    # Détermine le statut
    if not text:
        status = "ECHEC"
    elif stats["nb_words_clean"] < MIN_WORDS:
        status = "SUSPECT (peu de mots)"
    elif stats["warnings"]:
        status = "OK (avec avertissements)"
    else:
        status = "OK"

    # Écrit le fichier TXT
    if text:
        with open(out_path, "w", encoding="utf-8") as f:
            f.write(text)

    # Log
    log_rows.append({
        "fichier": pdf_path.name,
        "statut": status,
        "pages_total": stats["nb_pages"],
        "pages_extraites": stats["nb_pages_extracted"],
        "mots_bruts": stats["nb_words_raw"],
        "mots_nettoyés": stats["nb_words_clean"],
        "avertissements": " | ".join(stats["warnings"]) if stats["warnings"] else "",
    })

    # Affichage console
    icon = "✅" if status == "OK" else ("⚠️" if "avert" in status.lower() else "❌")
    print(f" {icon} {status:<30} | {stats['nb_words_clean']:>7} mots")

```

```

# Écrit le log CSV
with open(LOG_FILE, "w", newline="", encoding="utf-8") as f:
    writer = csv.DictWriter(f, fieldnames=log_rows[0].keys())
    writer.writeheader()
    writer.writerows(log_rows)

# Résumé final
ok_count    = sum(1 for r in log_rows if r["statut"].startswith("OK"))
suspect_count = sum(1 for r in log_rows if "SUSPECT" in r["statut"])
fail_count   = sum(1 for r in log_rows if r["statut"] == "ECHEC")

print(f"\n{' '*60}")
print(f"  ✓ OK          : {ok_count}")
print(f"  ⚠ Suspects     : {suspect_count}")
print(f"  ✗ Échecs       : {fail_count}")
print(f"  📄 Log détaillé : {LOG_FILE}")
print(f"{' '*60}\n")

if suspect_count or fail_count:
    print(" Fichiers à vérifier manuellement :")
    for r in log_rows:
        if r["statut"] != "OK":
            print(f"   → {r['fichier']} : {r['statut']}")
            if r["avertissements"]:
                print(f"       {r['avertissements'][:100]}")
    print()

if __name__ == "__main__":
    main()

```

Annexe C – Script analyse textuelle

```

install.packages(c("readxl", "stringr", "dplyr", "purrr", "writexl", "stringi"))

library(readxl)
library(stringr)
library(dplyr)
library(purrr)
library(writexl)
library(stringi)

# =====
# 1) CHEMINS
# =====

dict_en_path <- "C:/Users/bouil/OneDrive/Documents/Master
2/Mémoire/ESG_dictionary_EN.xlsx"

```

```

dict_fr_path <- "C:/Users/bouil/OneDrive/Documents/Master
2/Mémoire/ESG_dictionary_FR.xlsx"

folder_non_esg <- "C:/Users/bouil/OneDrive/Documents/Master 2/Mémoire/Rapports txt
Hors BEL ESG"
folder_bel_esg <- "C:/Users/bouil/OneDrive/Documents/Master 2/Mémoire/Rapports txt BEL
ESG"

output_path <- "C:/Users/bouil/OneDrive/Documents/Master
2/Mémoire/resultats_comptage_ESG_Final.xlsx"

# Noms exacts des 3 fichiers français, avec extension .txt
french_files <- c(
  "CBS.FR.2022.txt",
  "CBS.FR.2023.txt",
  "CBS.FR.2024.txt"
)

# =====
# 2) LECTURE DES DICTIONNAIRES
# =====

load_dict <- function(path) {
  read_excel(path) %>%
    mutate(
      Words = tolower(trimws(Words)),
      Words = stringi::stri_trans_general(Words, "Latin-ASCII"),
      Topic = trimws(Topic)
    ) %>%
    filter(!is.na(Words), Words != "") %>%
    distinct(Words, Topic)
}

dict_en <- load_dict(dict_en_path)
dict_fr <- load_dict(dict_fr_path)

# =====
# 3) NETTOYAGE TEXTE
# =====

clean_text <- function(text) {
  text %>%
    tolower() %>%
    stringi::stri_trans_general("Latin-ASCII") %>%
    str_replace_all("[[:punct:]]", " ") %>%
    str_replace_all("[[:digit:]]", " ") %>%
    str_replace_all("\\s+", " ") %>%
    trimws()
}

```

```

# =====
# 4) FONCTION DE COMPTAGE
# =====

count_esg_terms_in_file <- function(file_path, group_name, dict, dict_language) {

  text <- paste(readLines(file_path, warn = FALSE, encoding = "UTF-8"), collapse = " ")
  text_clean <- clean_text(text)

  total_words <- str_count(text_clean, "\\S+")

  count_term <- function(term, text_clean) {
    escaped_term <- str_replace_all(
      term,
      "([.|()\\^{}+$*?]|\\[|\\]|\\\\\\\\)",
      "\\\\"
    )
    pattern <- paste0("\\b", escaped_term, "\\b")
    str_count(text_clean, pattern)
  }

  term_counts <- dict %>%
    mutate(count = map_int(Words, ~ count_term(.x, text_clean)))

  counts_by_cat <- term_counts %>%
    group_by(Topic) %>%
    summarise(cat_count = sum(count), .groups = "drop")

  get_cat_count <- function(cat_name) {
    value <- counts_by_cat %>%
      filter(Topic == cat_name) %>%
      pull(cat_count)

    ifelse(length(value) == 0, 0, value)
  }

  env_count <- get_cat_count("Environmental")
  soc_count <- get_cat_count("Social")
  gov_count <- get_cat_count("Governance")

  total_esg_words <- sum(term_counts$count)

  esg_ratio <- ifelse(total_words > 0, total_esg_words / total_words, NA)
  esg_per_10000_words <- esg_ratio * 10000

  file_name <- basename(file_path)
  file_name_no_ext <- tools::file_path_sans_ext(file_name)

  data.frame(

```

```

file_name = file_name_no_ext,
file_name_with_ext = file_name,
group = group_name,
dictionary_used = dict_language,
total_words = total_words,
environmental_count = env_count,
social_count = soc_count,
governance_count = gov_count,
total_esg_words = total_esg_words,
esg_ratio = esg_ratio,
esg_per_10000_words = esg_per_10000_words,
stringsAsFactors = FALSE
)
}

# =====
# 5) LISTE DES FICHIERS
# =====

files_non_esg <- list.files(folder_non_esg, pattern = "\\..txt$", full.names = TRUE)
files_bel_esg <- list.files(folder_bel_esg, pattern = "\\..txt$", full.names = TRUE)

all_files <- c(files_non_esg, files_bel_esg)

# =====
# 6) COMPTAGE AVEC DICTIONNAIRE ADAPTÉ
# =====

results_all <- map_dfr(all_files, function(file_path) {

  file_name <- basename(file_path)

  group_name <- ifelse(
    dirname(file_path) == normalizePath(folder_bel_esg, winslash = "/", mustWork = FALSE),
    "BEL ESG",
    "Hors BEL ESG"
  )

  if (file_name %in% french_files) {
    count_esg_terms_in_file(file_path, group_name, dict_fr, "FR")
  } else {
    count_esg_terms_in_file(file_path, group_name, dict_en, "EN")
  }
})

# =====
# 7) EXPORT EXCEL
# =====

```

```

# =====
# 7) EXPORT EXCEL
# =====

write_xlsx(
    list(
        Comptage_ESG = results_all,
        Dictionnaire_EN = dict_en,
        Dictionnaire_FR = dict_fr
    ),
    "Base_Données_ESG_Fin.xlsx"
)

print("Excel Base_Données_ESG_Fin.xlsx créé avec succès")

```

```

#
=====
# STATISTIQUES DESCRIPTIVES — Mémoire BEL ESG
#
=====

if (!require("readxl")) install.packages("readxl")
if (!require("dplyr")) install.packages("dplyr")
library(readxl); library(dplyr)

# — 1. Chargement
=====

df <- read_excel("C:/Users/bouil/OneDrive/Documents/Master
2/Mémoire/Base_Donnée_Mémoire_Final.xlsx")
df <- df %>% rename(BEL_ESG = `BEL ESG`)
cat("Total observations :", nrow(df), "\n\n")

# — 2. Deux sous-échantillons
=====

# 2022 : année pré-indice (BEL_ESG=0 pour tout le monde par définition)
# 2023-2024 : période inférentielle (BEL_ESG varie)
df_pre <- df %>% filter(year == 2022)
df_inf <- df %>% filter(year %in% c(2023, 2024))

cat("Période pré-indice 2022 :", nrow(df_pre), "observations (toutes BEL_ESG=0)\n")
cat("Période inférentielle 2023-2024 :", nrow(df_inf), "observations\n")
cat(" dont BEL_ESG=1 :", sum(df_inf$BEL_ESG == 1), "\n")
cat(" dont BEL_ESG=0 :", sum(df_inf$BEL_ESG == 0), "\n\n")

vars <- c("ESG_LING", "ESGC", "GHG_INT", "Perf_ESG")

# — 3. Stats descriptives — période inférentielle par groupe
=====
cat("===== \n")
cat(" STATS DESCRIPTIVES — 2023-2024 PAR GROUPE || \n")
cat(" (période où BEL_ESG est défini et varie) || \n")
cat("===== \n\n")

for (v in vars) {
  cat(sprintf("—— %s —— \n", v))
  stats <- df_inf %>%
    group_by(BEL_ESG) %>%
    summarise(
      Groupe = ifelse(first(BEL_ESG)==1, "BEL ESG ", "Hors BEL ESG"),
      N = sum(!is.na(.data[[v]])),
      Moy = round(mean(.data[[v]], na.rm=TRUE), 3),
      Med = round(median(.data[[v]], na.rm=TRUE), 3),

```

```

SD   = round(sd(.data[[v]], na.rm=TRUE), 3),
Min   = round(min(.data[[v]], na.rm=TRUE), 3),
Max   = round(max(.data[[v]], na.rm=TRUE), 3),
.groups = "drop"
) %>% select(Groupe, N, Moy, Med, SD, Min, Max)
print(as.data.frame(stats))
cat("\n")
}

```

— 4. Stats descriptives — 2022 (référence pré-indice)

```

cat("||| \n")
cat("|| STATS DESCRIPTIVES — 2022 (PRÉ-INDICE, référence) || \n")
cat("||| \n\n")

```

```

for (v in vars) {
  x <- df_pre[[v]]
  x <- x[!is.na(x)]
  cat(sprintf("%-12s N=%d Moy=%.3f Méd=%.3f SD=%.3f Min=%.3f Max=%.3f\n",
    v, length(x), mean(x), median(x), sd(x), min(x), max(x)))
}

```

— 5. Test t — différence entre groupes (2023-2024)

```

cat("\n||| \n")
cat("|| TEST T — DIFFÉRENCE DE MOYENNES (2023-2024) || \n")
cat("||| \n\n")

```

```

for (v in vars) {
  bel <- df_inf[[v]][df_inf$BEL_ESG == 1 & !is.na(df_inf[[v]])]
  hors <- df_inf[[v]][df_inf$BEL_ESG == 0 & !is.na(df_inf[[v]])]
  if (length(bel) > 1 & length(hors) > 1) {
    tt <- t.test(bel, hors)
    sig <- ifelse(tt$p.value < 0.01, "***",
      ifelse(tt$p.value < 0.05, "**",
        ifelse(tt$p.value < 0.10, "*", "ns")))
    dir <- ifelse(mean(bel) > mean(hors), "BEL > Hors", "BEL < Hors")
    cat(sprintf("%-12s BEL=%.3f Hors=%.3f Diff=%.3f p=%.4f %s %s\n",
      v, mean(bel), mean(hors),
      mean(bel)-mean(hors), tt$p.value, sig, dir))
  }
}
cat("\n*** p<0.01 ** p<0.05 * p<0.10 ns=non significatif\n")

```

— 6. Distribution par année

```
cat("\n┌───────────────────────────────────┐\n")
cat("└─ DISTRIBUTION BEL_ESG PAR ANNÉE ─┘\n")
cat("┌───────────────────────────────────┐\n\n")

df %>%
  group_by(year, BEL_ESG) %>%
  summarise(n = n(), .groups="drop") %>%
  print()

# — 7. Corrélations (période inférentielle)

cat("\n┌───────────────────────────────────┐\n")
cat("└─ MATRICE DE CORRÉLATIONS — 2023-2024 ─┘\n")
cat("┌───────────────────────────────────┐\n\n")

cor_data <- df_inf %>%
  select(BEL_ESG, ESG_LING, ESGC, GHG_INT, Perf_ESG) %>%
  filter(complete.cases(.))
print(round(cor(cor_data), 3))

cat("\n✅ Terminé. Lance maintenant regressions_BEL_ESG.R pour les régressions.\n")
```

Annexe E – Script régressions OLS

RÉGRESSIONS OLS — Mémoire BEL ESG Index

```

# Hypothèses H1a, H1b, H2, H3
# 3 spécifications par hypothèse :
# (1) Sans contrôles
# (2) + Effets fixes année
# (3) + Effets fixes secteur + Log(Cap) [modèle principal]
#

```

```

# — 0. Packages

```

```

packages <- c("readxl", "dplyr", "lmtest", "sandwich", "stargazer")
for (p in packages) {
  if (!require(p, character.only = TRUE, quietly = TRUE)) {
    install.packages(p); library(p, character.only = TRUE)
  }
}

```

```

# — 1. Chargement et préparation

```

```

df <- read_excel("C:/Users/bouil/OneDrive/Documents/Master
2/Mémoire/Base_Donnée_Mémoire_Final.xlsx")

df <- df %>% rename(BEL_ESG = `BEL ESG`)

df$year_f <- as.factor(df$year)
df$year_f <- relevel(df$year_f, ref = "2022")
df$secteur_f <- as.factor(df$Secteur)
df$secteur_f <- relevel(df$secteur_f, ref = "Financier")
df$ln_cap <- df$`Log(Cap)`
df$ln_GHG_INT <- log1p(df$GHG_INT)

cat("Observations :", nrow(df), "| BEL_ESG=1 :", sum(df$BEL_ESG==1, na.rm=TRUE),
    "| BEL_ESG=0 :", sum(df$BEL_ESG==0, na.rm=TRUE), "\n\n")

```

```

# — Fonctions utilitaires

```

```

—
show_reg <- function(model, title = "") {
  cat("\n", strrep("=", 65), "\n", title, "\n", strrep("-", 65), "\n")
  print(coeftest(model, vcov = vcovHC(model, type = "HC3")))
  cat(sprintf("R2=%.4f | R2 adj=%.4f | N=%d\n",
    summary(model)$r.squared, summary(model)$adj.r.squared, nobs(model)))
}

conclude <- function(model, var="BEL_ESG", hyp="", direction="pos") {
  ct <- coeftest(model, vcov = vcovHC(model, type = "HC3"))
  beta <- round(ct[var, "Estimate"], 4)
  pval <- round(ct[var, "Pr(>|t|)"], 4)
  sig <- ifelse(pval<0.01, "****", ifelse(pval<0.05, "***", ifelse(pval<0.10, "**", "ns")))
}

```

```

ok <- (direction=="pos" && beta>0 && pval<0.05) ||
  (direction=="neg" && beta<0 && pval<0.05)
cat(sprintf("\n► %s : beta=%0.4f | p=%0.4f %s\n", hyp, beta, pval, sig))
if (ok)      cat(" -> CONFIRMEE\n")
else if (pval<0.1) cat(" -> PARTIELLEMENT CONFIRMEE (10%)\n")
else        cat(" -> NON CONFIRMEE\n")
}

#
=====
=====

# H1a — ESGC
#
=====
=====

cat("\n\nH1a : ESGC ~ BEL_ESG\n", strrep("=", 40), "\n")

h1a_1 <- lm(ESGC ~ BEL_ESG, data = df)
show_reg(h1a_1, "H1a (1) — sans controles")

h1a_2 <- lm(ESGC ~ BEL_ESG + year_f, data = df)
show_reg(h1a_2, "H1a (2) — + effets fixes annee")

h1a_3 <- lm(ESGC ~ BEL_ESG + ln_cap + year_f + secteur_f, data = df)
show_reg(h1a_3, "H1a (3) — + Log(Cap) + secteur [PRINCIPAL]")
conclude(h1a_3, "BEL_ESG", "H1a", "pos")

#
=====
=====

# H1b — ln(GHG_INT)
#
=====
=====

cat("\n\nH1b : ln(GHG_INT) ~ BEL_ESG\n", strrep("=", 40), "\n")

h1b_1 <- lm(ln_GHG_INT ~ BEL_ESG, data = df)
show_reg(h1b_1, "H1b (1) — sans controles")

h1b_2 <- lm(ln_GHG_INT ~ BEL_ESG + year_f, data = df)
show_reg(h1b_2, "H1b (2) — + effets fixes annee")

h1b_3 <- lm(ln_GHG_INT ~ BEL_ESG + ln_cap + year_f + secteur_f, data = df)
show_reg(h1b_3, "H1b (3) — + Log(Cap) + secteur [PRINCIPAL]")
conclude(h1b_3, "BEL_ESG", "H1b", "neg")

#
=====
=====

# H2 — ESG_LING

```

```

#
=====
cat("\n\nH2 : ESG_LING ~ BEL_ESG\n", strrep("=", 40), "\n")

h2_1 <- lm(ESG_LING ~ BEL_ESG, data = df)
show_reg(h2_1, "H2 (1) — sans controles")

h2_2 <- lm(ESG_LING ~ BEL_ESG + year_f, data = df)
show_reg(h2_2, "H2 (2) — + effets fixes annee")

h2_3 <- lm(ESG_LING ~ BEL_ESG + ln_cap + year_f + secteur_f, data = df)
show_reg(h2_3, "H2 (3) — + Log(Cap) + secteur [PRINCIPAL]")
conclude(h2_3, "BEL_ESG", "H2", "pos")

#
=====
# H3 — ESG_LING conditionnel
#
=====
cat("\n\nH3 : ESG_LING ~ Perf_ESG + BEL_ESG\n", strrep("=", 40), "\n")

h3_1 <- lm(ESG_LING ~ Perf_ESG + BEL_ESG, data = df)
show_reg(h3_1, "H3 (1) — sans controles")

h3_2 <- lm(ESG_LING ~ Perf_ESG + BEL_ESG + year_f, data = df)
show_reg(h3_2, "H3 (2) — + effets fixes annee")

h3_3 <- lm(ESG_LING ~ Perf_ESG + BEL_ESG + ln_cap + year_f + secteur_f, data = df)
show_reg(h3_3, "H3 (3) — + Log(Cap) + secteur [PRINCIPAL]")
conclude(h3_3, "Perf_ESG", " Perf_ESG (controle)", "pos")
conclude(h3_3, "BEL_ESG", " BEL_ESG (test central H3)", "pos")

#
=====
# TABLEAU RECAPITULATIF — modeles principaux
#
=====
cat("\n\nTABLEAU FINAL — 4 modeles principaux\n", strrep("=", 40), "\n\n")

stargazer(
  h1a_3, h1b_3, h2_3, h3_3,
  type      = "text",
  title     = "Resultats OLS — modeles principaux (spec. 3)",
  column.labels = c("H1a:ESGC", "H1b:ln(GHG)", "H2:LING", "H3:LING cond."),
  se = list(
    sqrt(diag(vcovHC(h1a_3,"HC3"))),

```

```

sqrt(diag(vcovHC(h1b_3,"HC3"))),
sqrt(diag(vcovHC(h2_3, "HC3"))),
sqrt(diag(vcovHC(h3_3, "HC3")))
),
star.cutoffs = c(0.1, 0.05, 0.01),
notes      = "HC3 robust SE. * p<0.1, ** p<0.05, *** p<0.01",
notes.append = FALSE, omit.stat = c("f","ser"), digits = 3, no.space = TRUE
)

#

```

```

# TABLEAUX 3 SPECIFICATIONS — H1a, H1b, H2, H3
#

```

```

for (specs in list(
  list(m=list(h1a_1,h1a_2,h1a_3), title="H1a : ESGC"),
  list(m=list(h1b_1,h1b_2,h1b_3), title="H1b : ln(GHG_INT)"),
  list(m=list(h2_1, h2_2, h2_3), title="H2 : ESG_LING"),
  list(m=list(h3_1, h3_2, h3_3), title="H3 : ESG_LING conditionnel")
)) {
  cat("\n\nTABLEAU —", specs$title, "\n", strep("-", 50), "\n\n")
  stargazer(
    specs$m[[1]], specs$m[[2]], specs$m[[3]],
    type      = "text",
    title     = specs$title,
    column.labels = c("(1) Sans controles", "(2) + Annee", "(3) + Secteur + Taille"),
    se = list(
      sqrt(diag(vcovHC(specs$m[[1]], "HC3"))),
      sqrt(diag(vcovHC(specs$m[[2]], "HC3"))),
      sqrt(diag(vcovHC(specs$m[[3]], "HC3")))
    ),
    star.cutoffs = c(0.1, 0.05, 0.01),
    notes      = "HC3 robust SE. * p<0.1, ** p<0.05, *** p<0.01",
    notes.append = FALSE, omit.stat = c("f","ser"), digits = 3, no.space = TRUE
  )
}

#

```

```

# ROBUSTESSE — Residus GW_RESID
#

```

```

cat("\n\nROBUSTESSE — Residus GW_RESID (H3)\n", strep("=", 40), "\n\n")

df$GW_RESID <- NA
idx <- as.integer(rownames(model.frame(h3_3)))
df$GW_RESID[idx] <- residuals(h3_3)

```

```

df %>%
  filter(!is.na(GW_RESID)) %>%
  group_by(BEL_ESG) %>%
  summarise(
    Groupe = ifelse(first(BEL_ESG)==1,"BEL ESG","Hors BEL ESG"),
    N      = n(),
    Moy    = round(mean(GW_RESID),4),
    Median = round(median(GW_RESID),4),
    SD     = round(sd(GW_RESID),4),
    .groups = "drop"
  ) %>% select(Groupe,N,Moy,Median,SD) %>% print()

bel_r <- df$GW_RESID[df$BEL_ESG==1 & !is.na(df$GW_RESID)]
hors_r <- df$GW_RESID[df$BEL_ESG==0 & !is.na(df$GW_RESID)]
tt <- t.test(bel_r, hors_r)
cat(sprintf("\nTest t residus : t=%0.3f | p=%0.4f\n", tt$statistic, tt$p.value))
if (tt$p.value < 0.05 && mean(bel_r) > mean(hors_r)) {
  cat("-> Residus BEL ESG > Hors BEL ESG : confirmation H3\n")
} else {
  cat("-> Pas de difference significative de residus\n")
}

cat("\n\nScript termine.\n")

```

Annexe F – Tableaux de régressions détaillés

H1a (1) – sans contrôles

t test of coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	48.9717	1.1988	40.849	< 2e-16 ***
BEL_ESG	4.7983	1.8891	2.540	0.01238 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R2=0.0453 | R2 adj=0.0372 | N=120

H1a (2) – + effets fixes année

t test of coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	49.53825	1.73854	28.4941	< 2e-16 ***
BEL_ESG	5.24515	2.26423	2.3165	0.02229 *
year_f2023	-1.58819	2.67125	-0.5945	0.55330
year_f2024	-0.52482	2.71438	-0.1933	0.84702

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R2=0.0489 | R2 adj=0.0243 | N=120

H1a (3) – + Log(Cap) + secteur [PRINCIPAL]

t test of coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	38.670651	2.493899	15.5061	< 2.2e-16 ***
BEL_ESG	2.822079	1.920555	1.4694	0.144656
ln_cap	9.489454	1.755769	5.4047	3.969e-07 ***
year_f2023	-0.663824	2.036578	-0.3260	0.745098
year_f2024	1.309067	2.240162	0.5844	0.560207
secteur_fConsommation	1.992495	3.777725	0.5274	0.598985
secteur_fImmobilier	11.210602	2.305571	4.8624	4.003e-06 ***
secteur_fIndustriel	-2.210337	2.619321	-0.8439	0.400631
secteur_fMatériaux de base	18.435381	3.144381	5.8630	5.115e-08 ***
secteur_fSanté	-0.046995	3.161163	-0.0149	0.988166
secteur_fServices Publics	5.096219	2.997741	1.7000	0.092032 .
secteur_fTechnologie	12.223437	4.374302	2.7944	0.006164 **
secteur_fTélécommunication	12.611382	2.743702	4.5965	1.182e-05 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R2=0.5101 | R2 adj=0.4551 | N=120

H1b (1) – sans contrôles

t test of coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	3.10664	0.22479	13.8204	< 2e-16 ***
BEL_ESG	-0.70235	0.37172	-1.8895	0.06128 .

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R2=0.0271 | R2 adj=0.0188 | N=120

H1b (2) – + effets fixes annee

t test of coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	2.89709	0.33069	8.7609	1.87e-14 ***
BEL_ESG	-0.90625	0.44127	-2.0537	0.04225 *
year_f2023	0.34914	0.49380	0.7070	0.48096
year_f2024	0.46813	0.52118	0.8982	0.37093

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R2=0.0350 | R2 adj=0.0100 | N=120

H1b (3) – + Log(Cap) + secteur [PRINCIPAL]

t test of coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	2.39857	0.48302	4.9658	2.603e-06 ***
BEL_ESG	-0.35340	0.36357	-0.9720	0.3332195
ln_cap	-1.08067	0.29143	-3.7082	0.0003327 ***
year_f2023	0.12618	0.37140	0.3397	0.7347138
year_f2024	0.12083	0.40221	0.3004	0.7644398
secteur_fConsommation	0.46421	0.65890	0.7045	0.4826365
secteur_fImmobilier	0.65331	0.44725	1.4607	0.1470229
secteur_fIndustriel	2.34835	0.92373	2.5423	0.0124455 *
secteur_fMatériaux de base	3.43324	0.55172	6.2228	9.702e-09 ***
secteur_fSanté	0.58005	0.35910	1.6153	0.1091932
secteur_fServices Publics	3.94048	0.41342	9.5313	5.877e-16 ***
secteur_fTechnologie	0.43570	0.44534	0.9784	0.3301011
secteur_fTélécommunication	0.18884	0.53270	0.3545	0.7236679

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R2=0.4935 | R2 adj=0.4368 | N=120

```
=====
H2 (1) - sans controles
-----
```

t test of coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	334.144	8.980	37.2100	< 2e-16 ***
BEL_ESG	42.989	19.424	2.2131	0.02881 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R2=0.0486 | R2 adj=0.0405 | N=120

```
=====
H2 (2) - + effets fixes annee
-----
```

t test of coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	323.65346	11.07081	29.2349	< 2e-16 ***
BEL_ESG	29.83169	21.78849	1.3691	0.17360
year_f2023	-0.70468	19.69874	-0.0358	0.97152
year_f2024	44.34817	21.32206	2.0799	0.03974 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R2=0.0987 | R2 adj=0.0754 | N=120

```
=====
H2 (3) - + Log(Cap) + secteur [PRINCIPAL]
-----
```

t test of coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	261.2578	16.9816	15.3847	< 2.2e-16 ***
BEL_ESG	36.4970	20.7636	1.7577	0.081652 .
ln_cap	15.7304	16.9711	0.9269	0.356067
year_f2023	-3.7122	17.7620	-0.2090	0.834847
year_f2024	42.0472	20.0945	2.0925	0.038762 *
secteur_fConsommation	45.4047	52.4874	0.8651	0.388941
secteur_fImmobilier	23.9162	18.3471	1.3035	0.195187
secteur_fIndustriel	79.4717	23.7251	3.3497	0.001118 **
secteur_fMatériaux de base	124.4118	26.6927	4.6609	9.121e-06 ***
secteur_fSanté	112.9940	33.4395	3.3791	0.001016 **
secteur_fServices Publics	79.0847	37.2196	2.1248	0.035904 *
secteur_fTechnologie	159.1274	31.5466	5.0442	1.872e-06 ***
secteur_fTélécommunication	50.1333	33.9303	1.4775	0.142469

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R2=0.4125 | R2 adj=0.3466 | N=120

```
=====
H3 (1) – sans contrôles
-----
```

t test of coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	335.9963	9.2064	36.4959	<2e-16 ***
Perf_ESG	14.7610	11.7835	1.2527	0.2128
BEL_ESG	36.9831	20.3193	1.8201	0.0713 .

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R2=0.0586 | R2 adj=0.0425 | N=120

```
=====
H3 (2) – + effets fixes annee
-----
```

t test of coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	324.3786	11.1942	28.9775	< 2e-16 ***
Perf_ESG	15.9673	12.0588	1.3241	0.18809
BEL_ESG	22.1713	22.9948	0.9642	0.33698
year_f2023	1.9163	19.6982	0.0973	0.92267
year_f2024	46.6375	21.2956	2.1900	0.03054 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R2=0.1104 | R2 adj=0.0794 | N=120

```
=====
H3 (3) – + Log(Cap) + secteur [PRINCIPAL]
-----
```

t test of coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	277.173	18.282	15.1605	< 2.2e-16 ***
Perf_ESG	36.250	16.642	2.1783	0.0316027 *
BEL_ESG	28.369	20.285	1.3985	0.1648758
ln_cap	-10.614	21.424	-0.4954	0.6213268
year_f2023	-1.406	17.397	-0.0808	0.9357397
year_f2024	40.885	20.057	2.0384	0.0439956 *
secteur_fConsommation	46.203	48.697	0.9488	0.3448776
secteur_fImmobilier	10.471	19.531	0.5361	0.5929907
secteur_fIndustriel	104.812	30.700	3.4141	0.0009077 ***
secteur_fMatériaux de base	123.908	27.109	4.5707	1.321e-05 ***
secteur_fSanté	118.388	32.203	3.6763	0.0003733 ***
secteur_fServices Publics	106.345	41.856	2.5408	0.0125095 *
secteur_fTechnologie	141.934	36.310	3.9090	0.0001637 ***
secteur_fTélécommunication	30.006	35.510	0.8450	0.4000139

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R2=0.4416 | R2 adj=0.3732 | N=120

Bibliographie

Austin, P. C. (2009). *Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples.* *Statistics in Medicine*, 28(25), 3083–3107. <https://doi.org/10.1002/sim.3697>

Baier, Berninger & Kiesel (2020). *Environmental, Social and Governance Reporting: A Textual Analysis*. [[Online appendix: ESG Word List](#)]

Barontini, R., & Gioja, L. (2025). The market power of ESG index providers: The effects of rebalancing ESG-themed indices. *Research in International Business and Finance*. DOI: 10.1016/j.ribaf.2025.102902

Euronext. (2023, 4 décembre). *Euronext launches the BEL® ESG Index*. Consulté le 30 octobre 2025 sur <https://www.euronext.com/en/about/media/euronext-press-releases/euronext-launches-belr-esg-index>

European Commission. (n.d.). *European Commission*. Retrieved Consulté le 30 octobre 2025, from https://commission.europa.eu/index_en

Fang, C., Wang, Z., & Zhao, L. (2025). Environmental regulations and the greenwashing of corporate ESG reports. *Economic Analysis and Policy*, 87, 1469–1481.

Friede, G., Busch, T., & Bassen, A. (2015). ESG and financial performance: Aggregated evidence from more than 2000 empirical studies. *Journal of Sustainable Finance & Investment*, 5(4), 210–233.

Hahn, R., & Kühnen, M. (2013). Determinants of sustainability reporting: A review of results, trends, and opportunities in an expanding field of research. *Journal of Cleaner Production*, 59, 5–21.

Khan, M., Serafeim, G., & Yoon, A. (2016). Corporate sustainability: First evidence on materiality. *The Accounting Review*, 91(6), 1697–1724.

Kolling, C., & Busch, T. (2025). Sustainability labels matter: Do ESG indices select sustainable firms? *Finance Research Letters*, 73, 106609.

Liu, G., Qian, H., Shi, Y., Zhang, Y., & Wu, F. (2024). Does ESG report greenwashing increase stock price crash risk? *China Journal of Accounting Studies*, 12(3), 615–639.

Petrenko, O. V., Aime, F., Ridge, J., & Hill, A. (2016). Corporate social responsibility or CEO narcissism? CSR as a determinant of CEO reputation. *Strategic Management Journal*, 37(2), 262-279.

Ruggeri, M., Vinci, G., Ruggieri, R., & Savastano, M. (2025). Facing the risk of greenwashing in the ESG report of global companies: The importance of life cycle thinking. *Corporate Social Responsibility and Environmental Management*. Advance online publication.

Stewart, R., Fantke, P., Bjørn, A., et al. (2018). *Life Cycle Assessment in Corporate Sustainability Reporting: Global, Regional, Sectoral, and Company-Level Trends*.

Xu, M., Tse, Y. K., Geng, R., Liu, Z., & Potter, A. (2025). Greenwashing and market value of firms: An empirical study. *International Journal of Production Economics*, 284, 10960

