

**Louvain School of Management**

# **Quel est l'impact des décisions et comportements des responsables politiques sur les marchés financiers américains ?**

Et plus spécifiquement, une stratégie de copiage du congrès américain est-elle viable ?

Auteur·e(s) : Lucas Dinis Pires, Ysaline Koch  
Promoteur : Bertrand Candelon  
Année académique 2025-2026  
Travail de fin d'études (TFE) en vue d'obtenir le titre de  
Master en sciences de gestion à finalité spécialisée  
Horaire de jour

## Statement on responsible use of generative AI

Au cours de la préparation de ce mémoire master, les auteurs Dinis Pires Lucas et Koch Ysaline ont eu recours à de l'intelligence artificielle (Claude Anthropic et Chat GPT Open AI). Ces intelligences artificielles ont été utilisées comme des outils d'assistance rédactionnelle (texte et code). Principalement pour structurer les idées des deux auteurs, mais surtout afin d'améliorer la qualité du texte et la qualité du code (python, R).

Exemple de certains prompts utilisés :

- « Reformule cette phrase »
- « Modifie le paragraphe afin qu'il soit grammaticalement cohérent avec le reste du texte »
- « Rend ce paragraphe plus simple tout en restant académique »
- « Vérifie les erreurs d'orthographe et de grammaire »
- « Remets ces paragraphes dans l'ordre afin de structurer nos idées »
- « Reformule ce paragraphe afin qu'il soit plus académique »
- Etc.

Les auteurs demeurent entièrement responsables du contenu de ce mémoire.

|                                                                                                                                            |                                                                        |
|--------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------|
| The content of our master's thesis is our original output, even if we used generative AI to help prepare it.                               | <input checked="" type="checkbox"/> Yes<br><input type="checkbox"/> No |
| We master the theories, tools and methods that we use for the thesis.                                                                      | <input checked="" type="checkbox"/> Yes<br><input type="checkbox"/> No |
| We verified that the output presented in the thesis is accurate and valid, and that it contains no hallucination, no false reference, etc. | <input checked="" type="checkbox"/> Yes<br><input type="checkbox"/> No |
| We acknowledge that we master that final output and that we are able to explain it and answer questions about it.                          | <input checked="" type="checkbox"/> Yes<br><input type="checkbox"/> No |
| We made sure not to provide generative AI tools with access to confidential data or information.                                           | <input checked="" type="checkbox"/> Yes<br><input type="checkbox"/> No |

### SIGNATURE

| Author last name, first name | Date       | Signature                                                                             |
|------------------------------|------------|---------------------------------------------------------------------------------------|
| DINIS PIRES Lucas            | 27-05-2026 |  |
| KOCH Ysaline                 | 27-05-2026 |  |

## Résumé :

Ce mémoire analyse les transactions boursières réalisées par les membres du Congrès américain entre 2021 et 2025. L'objectif principal est de déterminer si les élus bénéficient d'un avantage informationnel leur permettant de générer des rendements anormaux. Mais également d'évaluer la viabilité des stratégies de copy trading basées sur la publication de leurs opérations financières.

L'étude s'appuie sur près de 30 000 transactions issues de la plateforme Quiver Quantitative et sur des données de marché provenant de Yahoo Finance. Deux hypothèses sont testées : d'une part, l'existence de rendements anormaux après les transactions des élus ; d'autre part, la possibilité pour des investisseurs extérieurs de reproduire ces performances après publication des opérations dans le cadre du STOCK Act.

Les rendements analysés sont, dans l'ensemble, faibles, voire négatifs sur certains horizons temporels. Les résultats traduisent que les élus ne surperforment pas significativement le marché. Ces constats nuancent donc l'idée selon laquelle les élus américains bénéficieraient systématiquement d'un avantage informationnel.

Des différences selon, les partis, les secteurs et les délais de publication sont mises en exergue par notre analyse. Ainsi, les transactions publiées rapidement semblent conserver davantage de valeur informationnelle. Tandis qu'un délai de publication allongé réduit fortement l'intérêt du copy trading. En ce sens que l'information paraît déjà intégrée par les marchés.

En conclusion, le STOCK Act a probablement limité l'exploitabilité des transactions parlementaires. Certains signaux ponctuels subsistent, cependant aucune stratégie de copy trading systématique ne semble capable de battre durablement le marché.

**UNIVERSITÉ CATHOLIQUE DE LOUVAIN**  
**Louvain School of Management**

Place des Doyens, 1 bte L2.01.01, 1348 Louvain-la-Neuve  
Boulevard Emile Devreux 6, 6000 Charleroi, Belgique  
Chaussée de Binche 151, 7000 Mons, Belgique

[www.uclouvain.be/lsm](http://www.uclouvain.be/lsm)

## Table des matières

|                                                                                     |           |
|-------------------------------------------------------------------------------------|-----------|
| <b>I. INTRODUCTION.....</b>                                                         | <b>1</b>  |
| <b>II. REVUE DE LITTÉRATURE .....</b>                                               | <b>4</b>  |
| 1. EFFICIENCE DES MARCHÉS ET ASYMÉTRIE D'INFORMATION .....                          | 4         |
| 2. INSIDER TRADING ET PERFORMANCES ANORMALES .....                                  | 4         |
| 3. TRADING POLITIQUE ET CONGRÈS AMÉRICAIN .....                                     | 5         |
| 4. DÉVELOPPEMENT DU COPY TRADING POLITIQUE .....                                    | 6         |
| 5. POSITIONNEMENT DU MÉMOIRE .....                                                  | 7         |
| <b>III. MÉTHODOLOGIE .....</b>                                                      | <b>8</b>  |
| 1. SOURCES DE DONNÉES .....                                                         | 8         |
| 2. CONSTRUCTION DE L'ÉCHANTILLON .....                                              | 8         |
| 3. VARIABLES PRINCIPALES .....                                                      | 9         |
| 4. NETTOYAGE DES DONNÉES .....                                                      | 9         |
| 5. CONSTRUCTION DES RENDEMENTS .....                                                | 10        |
| <i>H1 : Rendements à partir de la date réelle de transaction.....</i>               | <i>10</i> |
| <i>H2 - Rendements à partir de la date de publication.....</i>                      | <i>11</i> |
| 6. MÉTHODE D'ANALYSE .....                                                          | 11        |
| 7. LIMITES MÉTHODOLOGIQUES .....                                                    | 13        |
| <b>IV. PARTIE 0 - ANALYSE DESCRIPTIVE ET EXPLORATOIRE.....</b>                      | <b>15</b> |
| 1. CONSTITUTION DE L'ÉCHANTILLON.....                                               | 15        |
| 2. STATISTIQUES DESCRIPTIVES SUR LES RENDEMENTS EXCÉDENTAIRES .....                 | 15        |
| 3. ACTIVITÉ DE TRADING .....                                                        | 15        |
| 4. CONCENTRATION DE L'ACTIVITÉ.....                                                 | 16        |
| <i>Par les élus.....</i>                                                            | <i>16</i> |
| <i>Par les titres.....</i>                                                          | <i>16</i> |
| 5. PERFORMANCES INDIVIDUELLES.....                                                  | 17        |
| <i>Par les élus.....</i>                                                            | <i>17</i> |
| <i>Par les titres.....</i>                                                          | <i>17</i> |
| 6. REMARQUES MÉTHODOLOGIQUES.....                                                   | 17        |
| <b>V. RÉSULTATS EMPIRIQUES .....</b>                                                | <b>18</b> |
| 1. PARTIE 1 - H1.....                                                               | 18        |
| <i>Return +1/+7/+30 jours vs. date de la transaction.....</i>                       | <i>18</i> |
| <i>Buy vs. Sell.....</i>                                                            | <i>18</i> |
| <i>Analyse politique.....</i>                                                       | <i>19</i> |
| <i>Discussion H1.....</i>                                                           | <i>20</i> |
| 2. PARTIE 2 - H2.....                                                               | 21        |
| <i>Rendements après publication.....</i>                                            | <i>21</i> |
| <i>Effet du reporting lag.....</i>                                                  | <i>22</i> |
| <i>Application au copy trading.....</i>                                             | <i>23</i> |
| <i>Discussion H2.....</i>                                                           | <i>25</i> |
| <b>VI. CONCLUSION .....</b>                                                         | <b>26</b> |
| 1. RÉPONSE À H1 : LES ÉLUS SURPERFORMENT-ILS LE MARCHÉ ? .....                      | 26        |
| 2. RÉPONSE À H2 : LE COPY TRADING EST-IL VIABLE ? .....                             | 27        |
| 3. IMPLICATIONS RÉGLEMENTAIRES.....                                                 | 28        |
| 4. LIMITES DU MÉMOIRE .....                                                         | 29        |
| 5. PERSPECTIVES DE RECHERCHE .....                                                  | 29        |
| 6. CONCLUSION GÉNÉRALE .....                                                        | 30        |
| <b>VII. REFERENCES.....</b>                                                         | <b>31</b> |
| <b>VIII. ANNEXES .....</b>                                                          | <b>34</b> |
| 1. GRAPHIQUES DES STATISTIQUES .....                                                | 34        |
| 1) <i>Distribution des rendements excédentaires – transactions du Congrès .....</i> | <i>34</i> |
| 2) <i>Rendements excédentaire médian par chambre.....</i>                           | <i>34</i> |

|     |                                                                       |    |
|-----|-----------------------------------------------------------------------|----|
| 3)  | <i>Rendements excédentaire par parti</i> .....                        | 35 |
| 4)  | <i>Rendement excédentaire médian par année</i> .....                  | 35 |
| 5)  | <i>Rendement excédentaire moyen par délai de déclaration</i> .....    | 36 |
| 6)  | <i>Volume de transactions estimés par année</i> .....                 | 36 |
| 7)  | <i>Distribution des délai de déclaration</i> .....                    | 37 |
| 8)  | <i>Élus les plus actifs par nombre de transactions</i> .....          | 37 |
| 9)  | <i>20 actions les plus échangées par les membres du Congrès</i> ..... | 38 |
| 10) | <i>Top 20 élus par rendement excédentaire médian</i> .....            | 38 |
| 11) | <i>Rendement excédentaire moyen par délai de déclaration</i> .....    | 39 |
| 12) | <i>ANOVA</i> .....                                                    | 39 |
| 13) | <i>Bonferroni, Comparaison par paires</i> .....                       | 40 |
| 14) | <i>Rendement moyen par stratégie de copy trading</i> .....            | 40 |
| 2.  | <b>DATA</b> .....                                                     | 41 |
| 1)  | <i>ANNEXE 1</i> .....                                                 | 41 |
| 2)  | <i>ANNEXE 2</i> .....                                                 | 41 |
| 3)  | <i>ANNEXE 3</i> .....                                                 | 41 |
| 4)  | <i>ANNEXE 4</i> .....                                                 | 41 |
| 3.  | <b>CODES</b> .....                                                    | 41 |
| 1)  | <i>ANNEXE 5</i> .....                                                 | 41 |
| 2)  | <i>ANNEXE 6</i> .....                                                 | 42 |
| 3)  | <i>ANNEXE 7</i> .....                                                 | 47 |

## I. Introduction

Depuis de nombreuses années, les transactions financières réalisées par les membres du Congrès américain suscitent une attention croissante, tant dans le débat public que dans la littérature académique. Cette attention s'est particulièrement intensifiée depuis l'adoption du Stop Trading on Congressional Knowledge Act (STOCK Act) en 2012 (United States Congress, 2012), qui impose aux élus américains de déclarer publiquement leurs opérations boursières dans un délai maximal de 45 jours (PTR, Periodic Transaction Reports) (United States Congress, 2012 ; U.S. Department of the Interior, 2017). L'optimisation de la transparence ainsi que la limitation des risques d'utilisation d'informations privilégiées, sont clairement les objectifs de cette réglementation, qui s'est imposée dans le cadre des activités politiques et législatives. Celle-ci a permis une étude empirique de la performance boursière des élus américains et d'un potentiel avantage lié à leur position (Eggers & Hainmueller, 2013 ; Hanousek Jr., et al., 2021)

Il apparaît cependant que l'asymétrie d'information soit toujours au cœur de nombreuses controverses au sein du Congrès. Les médias ne sont pas pour rien dans le renforcement de l'idée selon laquelle, des élus pourraient bénéficier d'informations susceptibles d'influencer les marchés financiers. En effet, durant l'après-covid ainsi que lors des périodes de forte volatilité des marchés, des opérations particulièrement médiatisées ont renforcé cette idée.

Les déclarations publiques des opérations financières des élus ont vu l'émergence de plateformes spécialisées dans le suivi des transactions parlementaires. Des plateformes telles que Quiver Quantitative (Hall, 2025) ainsi que l'hypermédiatisation de certains portefeuilles politiques ne sont pas étrangères au phénomène de « copy trading ». Le « copy trading » étant la reproduction des transactions réalisées par les membres du Congrès (United States Congress, 2012 ; U.S. Department of the Interior, 2017).

Une double interrogation apparaît dès lors. D'une part, les membres du Congrès réalisent-ils effectivement des rendements anormaux susceptibles de refléter un avantage informationnel ? D'autre part, même si un tel avantage existe, reste-t-il exploitable par des investisseurs extérieurs une fois les transactions rendues publiques, compte tenu du délai de déclaration imposé par le STOCK Act ?

La littérature académique existante apporte des résultats contrastés. En effet, selon Alan J. Ziobrowski et al., les élus américains obtiennent des performances anormalement positives de leurs transactions (Ziobrowski et al., 2004, 2011). Les auteurs indiquent : *“Senators are likely to have knowledge of forthcoming government actions before the information becomes public. This could provide them with an informational advantage over other investors”* (Ziobrowski et al., 2004, p. 661). Ils suggèrent donc l’existence d’un avantage informationnel potentiel. D’autres travaux, tels que ceux de Andrew C. Eggers et Jens Hainmueller (2013), concluent à une absence de surperformance significative. Les auteurs soutiennent même que les élus sous-performeraient même le marché. Eggers et Hainmueller (2013, p. 535) : *“We find no evidence of either informed trading or above-market portfolio returns for Congress as a whole or any subset of members. In fact, the average investor in Congress underperformed the market [...]”*. Toutefois, peu d’études récentes exploitent les données postérieures à la pandémie Covid-19, période marquée par une hausse importante de l’activité de trading des élus et par une visibilité accrue des divulgations parlementaires auprès du grand public.

Dans ce contexte, notre mémoire poursuit un double objectif. Premièrement, il vise à évaluer si les transactions réalisées par les membres du Congrès américain génèrent des rendements anormaux à court terme après leur exécution (H1). Deuxièmement, notre travail tente d’observer si, malgré le délai de publications des transactions parlementaires imposé par le STOCK Act, les transactions conservent une valeur informationnelle exploitable. Cette question amène à une application empirique, sous forme de simulation simplifiée, de stratégie de copy trading (H2).

Pour répondre à ces questions, ce travail s’appuie sur une base de données de plus de 31 000 transactions issues de Quiver Quantitative et réalisées entre 2021 et 2025 par des membres du Congrès américain. Les rendements post-transaction sont construits à partir de données de marché provenant de Yahoo Finance. Ceux-ci nous permettront par la suite de tester plusieurs horizons temporels de performance (+1 jour, +7 jours et +30 jours). Une attention particulière est également portée à l’hétérogénéité des comportements de trading ainsi qu’aux secteurs les plus fréquemment associés aux meilleures performances observées.

Nous espérons contribuer à la littérature existante grâce à ce travail. Nous proposerons dans un premier temps, une actualisation empirique de la question du trading parlementaire à partir de données récentes couvrant la période post-Covid. Nous examinerons dans un second temps, la viabilité économique réelle d’une stratégie de copy trading basée sur les obligations de

publications du Congrès. Nous démontrerons, dans un troisième temps, les limites méthodologiques et réglementaires associées à l'étude empirique des transactions politiques.

Ce mémoire s'articule autour de deux hypothèses principales.

H1 : Les membres du Congrès américain réalisent des rendements anormaux après leurs transactions boursières.

H2 : Une stratégie de copy trading fondée sur la publication des transactions parlementaires permet de générer des rendements anormaux malgré le délai imposé par le STOCK Act.

## II. Revue de littérature

### 1. Efficience des marchés et asymétrie d'information

L'analyse des transactions financières réalisées par les membres du Congrès américain s'inscrit dans le cadre plus large de la littérature sur l'efficience des marchés financiers et l'asymétrie d'information. Selon l'Efficient Market Hypothesis développée par Eugene Fama (1970), les prix des actifs financiers incorporent instantanément l'ensemble des informations disponibles. La publication induirait une diminution de l'asymétrie d'information (Akerlof, 1970). Impliquant donc qu'aucun investisseur ne devrait pouvoir surperformer à partir d'informations publiques.

Toutefois, plusieurs travaux ont remis en question cette théorie en mettant en exergue, que les différents acteurs économiques ne disposent pas du même accès à l'information. George Akerlof (1970) dans son concept d'Information Asymmetry, met en avant les potentiels déséquilibres lorsqu'un agent dispose d'informations supérieures à celles des autres participants du marché. Dans ce contexte, certains individus ou institutions peuvent être en mesure d'obtenir des avantages économiques significatifs grâce à un accès privilégié à l'information.

Les responsables politiques constituent un cas particulier d'acteurs potentiellement exposés à ce type d'asymétrie. En raison de leur participation aux processus législatifs, de leur accès à certaines informations macroéconomiques ou réglementaires et de leurs interactions avec de nombreux acteurs institutionnels et privés, les élus peuvent théoriquement bénéficier d'informations pertinentes avant leur diffusion au public.

### 2. Insider trading et performances anormales

La littérature financière a largement étudié la capacité de certains acteurs à générer des rendements anormaux grâce à des informations privilégiées ou à des positions informationnelles favorables. Les travaux portant sur l'insider trading ou délit d'initié, montrent généralement que les dirigeants d'entreprise réalisent des performances supérieures à celles du marché lorsqu'ils négocient les titres de leur propre société.

Les recherches de Josef Lakonishok et Inmoo Lee (2001) mettent notamment en évidence l'existence de rendements anormaux positifs associés aux transactions d'initiés, surtout dans le

cas des achats. L'idée que certains agents bénéficiant d'informations privilégiées anticipent effectivement les évolutions futures des marchés financiers est clairement suggérée par ces résultats.

Cette littérature a suscité l'intérêt, et plusieurs études se sont penchées sur les transactions financières d'acteurs disposant d'un accès privilégié aux informations publiques ou/et réglementaires. Parmi ces acteurs, on retrouve notamment des responsables politiques et des hauts fonctionnaires. Les travaux de Mara Faccio (2002) montrent par exemple que les connexions politiques peuvent influencer significativement les comportements d'investissement et les performances financières observées sur les marchés.

### 3. Trading politique et Congrès américain

Le cas du Congrès américain a fait l'objet d'une attention particulière dans la littérature académique en raison de l'importante quantité d'informations économiques et réglementaires auxquelles les élus peuvent avoir accès dans le cadre de leurs fonctions (Hanousek et al., 2023).

Les travaux de Alan J. Ziobrowski et al. concluent que les sénateurs américains auraient historiquement réalisé des performances supérieures à celles du marché (Ziobrowski et al., 2004). Les auteurs suggèrent l'existence d'un avantage informationnel potentiel. Ils soulignent entre autre des rendements anormaux positifs associés aux portefeuilles des sénateurs américains, peu ou pas compatibles avec une stricte efficience des marchés.

Des recherches plus récentes ont contredit ces résultats. Andrew C. Eggers et Jens Hainmueller (2013) montrent notamment que les performances observées dans certaines études antérieures pourraient être fortement influencées par des biais méthodologiques, des problèmes de sélection des données ou encore des effets de composition des portefeuilles.

La question du trading parlementaire a connu un regain d'attention après l'adoption du Stop Trading on Congressional Knowledge Act (STOCK Act) (United States Congress, 2012), qui impose aux membres du Congrès américain de déclarer publiquement leurs transactions financières dans un délai maximal de 45 jours. Cette réforme visait principalement à renforcer la transparence et à limiter les risques d'utilisation d'informations privilégiées par les élus (Harnisch et al., 2020).

Toutefois, malgré cette réglementation, les controverses médiatiques autour des transactions parlementaires sont restées nombreuses, particulièrement après la pandémie de Covid-19. La forte volatilité des marchés, les plans de relance successifs et les débats réglementaires autour des secteurs technologiques, pharmaceutiques ou industriels ont contribué à renforcer l'attention portée aux transactions réalisées par certains élus américains.

#### 4. Développement du copy trading politique

Parallèlement à la médiatisation croissante des transactions parlementaires, plusieurs plateformes spécialisées, telles que Quiver Quantitative, ont facilité l'accès du grand public aux déclarations financières des membres du Congrès. Cette accessibilité accrue a favorisé le développement de stratégies dites de « copy trading », consistant à reproduire les transactions réalisées par certains élus considérés comme particulièrement performants (Erdős et al., 2022).

La propagation des stratégies de copy trading politique a été renforcée par les réseaux sociaux et par le lancement de produits financiers inspirés des transactions parlementaires (Hall & Garcia, 2025; QuiverQuant, s.d.).

L'idée implicite ici, repose sur la supposition de la part des investisseurs potentiels, que certains membres du Congrès disposent de plus d'informations ou d'anticipations quant aux évolutions des marchés. Dès lors, imiter leurs transactions pourrait théoriquement permettre aux investisseurs extérieurs de bénéficier indirectement de cet avantage informationnel.

Néanmoins, la viabilité réelle de ces stratégies demeure incertaine. Le délai de publication imposé par le STOCK Act peut réduire considérablement l'exploitabilité des informations disponibles, dans la mesure où les marchés financiers peuvent avoir déjà intégré l'information pertinente avant la publication officielle des transactions.

Cette problématique reste relativement peu étudiée dans la littérature académique récente, notamment sur des données postérieures à la pandémie de Covid-19 et dans un contexte de démocratisation massive des plateformes de suivi des transactions politiques.

La question centrale ne porte toutefois pas uniquement sur la rentabilité du copy trading, mais également sur l'impact du délai de publication imposé par le STOCK Act sur la dissipation progressive de l'information potentiellement contenue dans les transactions parlementaires.

## 5. Positionnement du mémoire

Dans ce contexte, le présent mémoire s'inscrit dans la continuité des travaux portant sur les performances financières des responsables politiques américains tout en cherchant à apporter une contribution plus récente et plus appliquée à la question du trading parlementaire.

Premièrement, cette étude propose une actualisation empirique de la littérature à partir de données couvrant la période 2021-2025, marquée par une forte volatilité des marchés et une médiatisation accrue des transactions du Congrès.

Deuxièmement, elle distingue explicitement deux dimensions souvent confondues dans les travaux antérieurs :

- l'existence éventuelle de rendements anormaux associés aux transactions des élus eux-mêmes ;
- la possibilité pour des investisseurs extérieurs d'exploiter ces informations via des stratégies de copy trading après publication des transactions.

Enfin, ce mémoire accorde une attention particulière aux limites méthodologiques associées à l'étude empirique des transactions parlementaires, notamment en matière de qualité des données, de délais de publication et d'hétérogénéité des comportements de trading observés entre élus et secteurs d'activité.

### III. Méthodologie

#### 1. Sources de données

L'étude empirique des transactions nécessite une base de données cohérente, de qualité et exploitable. C'est la base de Quiver Quantitative qui a été retenue comme principale source pour ce travail.

Quiver Quantitative est une plateforme spécialisée dans l'agrégation des déclarations financières des membres du Congrès américain imposées par le Stop Trading on Congressional Knowledge Act (STOCK Act). La base utilisée recense les transactions boursières déclarées publiquement par les élus américains ainsi que plusieurs informations associées aux opérations, notamment la date de transaction, la date de publication, le type d'opération et le titre concerné.

Yahoo Finance nous a procuré les données de marché nécessaires au calcul des rendements post-transaction. La totalité des prix historiques journaliers des titres ont été utilisés afin de reconstituer les variations des rendements observés en fonction des transactions réalisées par les membres du Congrès. Le S&P 500 est utilisé comme benchmark de marché pour le calcul des rendements excédentaires.

#### 2. Construction de l'échantillon

L'échantillon initial couvre l'ensemble des transactions parlementaires disponibles entre janvier 2021 et avril 2026. Plusieurs filtres ont ensuite été appliqués afin d'obtenir un échantillon homogène et cohérent avec les objectifs de l'étude.

Seules les transactions réalisées entre 2021 et 2025 ont été conservées afin d'éviter les biais liés à une année 2026 incomplète. Ensuite, les opérations de type *purchase* (achat) et *sale* (vente) ont été choisies et donc extraites.

Le choix d'exclure ETF, options, obligations, cryptomonnaies et autres instruments dérivés a été posé pour nous permettre une analyse concentrée sur des actions cotées traditionnelles.

Après nettoyage, l'échantillon final comprend 29 859 transactions réalisées par 224 membres du Congrès américain sur 2 646 titres distincts.

Certaines valeurs extrêmes ont fait l'objet d'un traitement spécial. Afin de limiter l'influence des valeurs aberrantes les rendements excédentaires ont été winsorisés (plafonnés) aux 1er et 99e percentiles (bornes retenues : -162 % et +501 %).

Enfin, le représentant Ro Khanna a été exclu de l'échantillon final. Celui-ci totalise à lui seul 15 218 transactions sur la période étudiée, principalement attribuées à une activité algorithmique automatisée qui ne reflète pas directement des décisions discrétionnaires d'investissement. Son inclusion aurait introduit un biais de concentration important dans les analyses descriptives et empiriques.

### 3. Variables principales

Les principales variables utilisées dans cette étude sont les suivantes :

- rendement brut post-transaction ;
- rendement excédentaire par rapport au S&P 500 ;
- type de transaction (achat/vente) ;
- délai de déclaration (*reporting lag*) ;
- affiliation politique ;
- chambre parlementaire ;
- horizons temporels (+1, +7 et +30 jours).

Le délai de déclaration constitue une variable particulièrement importante dans le cadre de l'hypothèse H2, puisqu'il permet d'évaluer dans quelle mesure le temps écoulé entre la transaction réelle et sa publication affecte l'exploitabilité potentielle des transactions parlementaires.

### 4. Nettoyage des données

L'utilisation des données de l'échantillon construit nécessite un nettoyage. Ce nettoyage en plusieurs points permet l'analyse empirique.

Dans un premier temps, les doublons éventuels ont été supprimés. Les tickers (symbole boursier) non exploitables ou ne disposant pas d'historique de prix suffisant ont été exclus. Dans un second temps, les dates de transactions et de publications ont été standardisées. Cette

standardisation nous permet d'une part de calculer les rendements post-transactions mais également d'analyser le délai de publication.

Nous avons porté une attention particulière aux incohérences présentes dans certaines déclarations parlementaires, notamment concernant les types d'actifs ou les informations manquantes. Dans un troisième et dernier temps, nous avons exclu les données précitées, qui ne permettaient pas de construire correctement les rendements.

## 5. Construction des rendements

L'analyse empirique repose sur deux approches complémentaires correspondant aux deux hypothèses principales du mémoire.

### H1 : Rendements à partir de la date réelle de transaction

Cette première approche vise à évaluer si les transactions des membres du Congrès génèrent des rendements anormaux statistiquement significatifs sur différents horizons temporels.

Pour cette partie, les rendements sont calculés à partir de la date réelle d'exécution de la transaction déclarée par l'élue. Les prix historiques journaliers des actions sont récupérés via Yahoo Finance afin de reconstruire les performances observées après chaque opération.

Trois horizons temporels sont retenus :

- +1 jour ;
- +7 jours ;
- +30 jours.

Ces 3 horizons temporels nous permettent de distinguer les effets immédiats (+1), de court terme (+7) et plus persistants (+30) potentiellement associés aux transactions parlementaires.

Les rendements post-transaction sont calculés selon la formule suivante :

$R_{\{i,t+n\}} = \frac{P_{\{i,t+n\}} - P_{\{i,t\}}}{P_{\{i,t\}}}$  ; où :

- $(P_{\{i,t\}})$  représente le prix du titre à la date de transaction ;
- $(P_{\{i,t+n\}})$  représente le prix observé (n) jours après la transaction.

Les rendements excédentaires sont ensuite obtenus en soustrayant le rendement du S&P 500 observé sur la même période.

Cette première approche vise à mesurer l'existence éventuelle d'un avantage informationnel associé aux transactions des membres du Congrès.

## H2 - Rendements à partir de la date de publication

La seconde analyse cherche à évaluer dans quelle mesure le délai de publication des transactions parlementaires affecte l'exploitabilité potentielle des informations rendues publiques.

Dans un premier temps, les rendements sont calculés à partir de la date officielle de publication des transactions afin de reproduire les conditions auxquelles ferait face un investisseur cherchant à imiter les opérations des élus après publication.

Une simulation simplifiée de stratégie de copy trading est ensuite réalisée. La stratégie suppose qu'un investisseur prend position le lendemain de la publication officielle de la transaction et conserve le titre sur différents horizons temporels (+1, +7 et +30 jours).

Dans un second temps, les transactions sont regroupées selon leur délai de déclaration (*reporting lag*) :

- 0-10 jours ;
- 11-30 jours ;
- 31-45 jours ;
- plus de 45 jours.

Les différences de rendement entre groupes sont analysées à l'aide d'une analyse de variance (ANOVA) ainsi que de tests pairwise (post hoc) de Bonferroni.

Cette approche permet d'évaluer si les transactions déclarées rapidement présentent des performances différentes des transactions publiées plus tardivement.

## 6. Méthode d'analyse

L'analyse empirique se déroule en trois étapes.

La première consiste en une analyse descriptive et exploratoire des transactions parlementaires. Celle-ci porte notamment sur :

- la distribution des rendements ;
- les différences entre achats et ventes ;
- les comportements observés selon les chambres et les partis politiques ;
- les délais de déclaration ;
- la concentration de l'activité de trading.

La deuxième étape vise à tester l'existence éventuelle de rendements anormaux associés aux transactions des membres du Congrès. Nous avons mobilisé quatre procédures statistiques, choisies en fonction de la structure des données et de la nature des comparaisons effectuées.

- Le test t de Student à un échantillon nous a permis de tester si le rendement excédentaire moyen de la population est statistiquement différent de zéro. Les hypothèses nulle et alternative s'écrivent :  $H_0 : \mu = 0$  ;  $H_1 : \mu \neq 0$

La statistique de test est :  $t = (\bar{x} - \mu_0) / (s / \sqrt{n})$ , où  $\bar{x}$  est la moyenne empirique des rendements excédentaires,  $s$  l'écart-type de l'échantillon,  $n$  la taille de l'échantillon et  $\mu_0 = 0$  la valeur testée sous  $H_0$ . La statistique suit une loi de Student à  $n-1$  degrés de liberté. Étant donné le nombre important d'observations mobilisé pour ce travail, nous avons utilisé un théorème central limite afin de garantir une normalité asymptotique de la statistique de test. Dès lors, l'hypothèse de normalité des rendements n'est plus une condition nécessaire à la validité des résultats. Le test est bilatéral et le seuil de signification retenu est  $\alpha = 0,05$ .

- Afin de comparer les rendements entre sous-groupes (démocrates vs républicains, Chambre vs Sénat, achats vs ventes), nous avons eu recours au test t de Welch (Welch, 1947). Cette variante du test t de Student est adaptée dans le cas où les variances de population sont inégales entre groupes. Sous  $H_0 : \mu_1 = \mu_2$ , la statistique de Welch est :  $t_W = (\bar{x}_1 - \bar{x}_2) / \sqrt{(s_1^2/n_1 + s_2^2/n_2)}$ , où  $\bar{x}_1$ ,  $\bar{x}_2$  désignent les moyennes des deux groupes,  $s_1^2$ ,  $s_2^2$  leurs variances empiriques et  $n_1$ ,  $n_2$  leurs effectifs. Les degrés de liberté sont estimés par l'approximation de Satterthwaite (1946) :

$$v = (s_1^2/n_1 + s_2^2/n_2)^2 / [(s_1^2/n_1)^2/(n_1 - 1) + (s_2^2/n_2)^2/(n_2 - 1)]$$

Le recours à cette variante est justifié par l'hétéroscédasticité observée entre sous-groupes (achats et ventes, partis politiques) et garantit un contrôle rigoureux du risque de conclure qu'il existe un effet alors qu'il n'en existe pas.

- Une ANOVA à un facteur et correction de Bonferroni. Cette analyse de l'effet du délai de déclaration sur les rendements post-publication implique une comparaison simultanée entre quatre catégories. Une analyse de variance à un facteur (ANOVA) est d'abord conduite afin de tester l'égalité globale des moyennes :  $H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$
- La statistique F nous a permis de comparer la variance inter-groupes à la variance intra-groupe. En cas de rejet de l'hypothèse nulle ( $H_0$ ), des comparaisons par paires post-hoc sont effectuées avec la correction de Bonferroni. Celle-ci ajuste le seuil de signification à  $\alpha/m$ , où  $m$  désigne le nombre de comparaisons réalisées. Cette correction nous a permis de contrôler le risque d'erreur de type I lié aux tests multiples.
- Le test de Mann-Whitney U (robustesse non paramétrique) (Mann & Whitney, 1947), en complément des tests paramétriques, est utilisé comme vérification de robustesse. Ce test ne requiert aucune hypothèse distributionnelle. Il est donc particulièrement adapté aux distributions de rendements asymétriques et leptokurtiques observées dans notre échantillon. Celui-ci nous a permis de confirmer les conclusions obtenues par les tests paramétriques sans présupposer la normalité.

La dernière et la troisième étape se penche sur la viabilité économique d'une stratégie de copy trading. Compte tenu du fait que celle-ci serait fondée sur les publications des transactions. Les performances calculées à partir des dates réelles de transaction sont comparées à celles obtenues après publication officielle. Cette comparaison nous permet d'observer un éventuel impact du délai de publication. L'impact portant sur la possible exploitation d'informations financières parlementaires.

Compte tenu des limites des données disponibles, cette analyse sectorielle conserve un caractère exploratoire. En effet, nous avons observé une grande hétérogénéité dans les comportements de trading. Il apparaît également que certains secteurs offrent potentiellement de meilleures performances et ce notamment dans les domaines technologiques et industriels.

## 7. Limites méthodologiques

Plusieurs limites doivent être prises en considération dans l'interprétation des résultats.

Premièrement, les données issues des publications parlementaires présentent certaines imprécisions. Les montants exacts des transactions ne sont pas publiés mais communiqués sous

forme d'intervalles. De plus, certaines opérations peuvent être déclarées avec retard ou comporter des approximations concernant les dates d'exécution.

Deuxièmement, l'affiliation à un secteur précis par les élus, tout comme les liens directs entre leurs activités législatives et les entreprises des secteurs concernés ne sont pas systématiquement disponibles dans les bases de données utilisées. L'analyse sectorielle réalisée dans ce mémoire doit donc être interprétée comme une analyse exploratoire et non comme une tentative de démonstration causale stricte.

Troisièmement, des approches économétriques plus avancées, telles que les modèles Difference-in-Differences (DiD), auraient éventuellement pu permettre l'identification plus précise de certains effets causaux associés aux transactions parlementaires.

Cependant, la structure des données disponibles, l'hétérogénéité importante des transactions et les contraintes méthodologiques de notre travail, nous ont conduit à opter pour une approche principalement descriptive et comparative.

Pour conclure, les rendements observés ne permettent pas d'établir une relation causale directe entre activité politique et performance boursière. Les résultats doivent être interprétés avec prudence. Ceux-ci doivent être replacés dans le cadre plus large des limites propres aux études empiriques sur les comportements financiers et politiques.

## IV. Partie 0 - Analyse descriptive et exploratoire

### 1. Constitution de l'échantillon

Les données proviennent de Quiver Quantitative, qui agrège les déclarations de transactions imposées par le STOCK Act. Après nettoyage, l'échantillon retenu couvre 29 859 transactions réalisées par 224 membres du Congrès sur 2 646 titres distincts, entre janvier 2021 et avril 2026.

Nous avons appliqué, à l'aide d'un code R, les filtres suivants: (i) fenêtre temporelle 2021-2025, (ii) transactions de type achat ou vente uniquement, (iii) actions cotées seulement et (iv) winsorisation des rendements excédentaires aux 1er et 99e percentiles (bornes : -162 % / +501 %) (v) exclusion de Ro Khanna (activité algorithmique).

### 2. Statistiques descriptives sur les rendements excédentaires

Force est de constater que seul 33,6% des transactions affichent un rendement excédentaire positif<sup>1</sup>. Pris dans leur ensemble, les membres du Congrès sous-performent. En effet, le rendement excédentaire médian est à -21,4 %, tandis que la moyenne est à -11,3 %.

Ces constatations viennent contrarier l'idée populaire selon laquelle les élus profiteraient d'une asymétrie informationnelle systématique.

Nous pourrions suggérer que la sous-performance est liée au sens de la transaction. Ce qui n'est pas le cas, les achats (médiane : -21,6 %) et les ventes (médiane : -21,3 %) évoluent de manière quasi identique. Cela reflèterait davantage des biais de sélection ou de timing communs aux deux types d'opérations.

Par chambre<sup>2</sup>, le Sénat (-20,3 %) devance légèrement la Chambre des représentants (-21,5 %), les deux restant en territoire négatif. Par parti<sup>3</sup>, les républicains affichent une moyenne de -8,0 % contre -16,2 % pour les démocrates, soit un écart d'environ 8 points de pourcentage.

### 3. Activité de trading

L'activité est structurellement vendeuse<sup>4</sup>. Le ratio achats/ventes est inférieur à 1 pour chaque année de l'échantillon<sup>5</sup> : 0,94 en 2021 et 2022, 0,74 en 2023, 0,89 en 2024, 0,81 en 2025. Deux

<sup>1</sup> Voir annexe 1. 1) Distribution des rendements excédentaires – transactions du Congrès

<sup>2</sup> Voir annexe 1. 2) Rendements excédentaires médian par chambre

<sup>3</sup> Voir annexe 1. 3) Rendements excédentaires médian par parti

<sup>4</sup> Voir annexe 1. 4) Rendements excédentaires médian par année

<sup>5</sup> Voir annexe 1. 5) Activité de trading du Congrès (2021-2025)

hypothèses sont plausibles pour expliquer cette tendance à la vente. D'une part, les pressions réglementaires liées à la transparence et d'autre part, des besoins de liquidités pour pouvoir faire face aux coûts des campagnes électorales.

Les transactions en 2021 performant avec un pic<sup>6</sup> de volume de 198 M\$ d'achats et 202 M\$ de ventes, soit 400 M\$ combinés. Le volume des transactions décline jusqu'en 2024 pour atteindre 190 M\$ (72 M\$ en achats et 118 M\$ en ventes). Nous observons un rebond en 2025, le volume transactionnel s'élevant à 289 M\$ combinés.

Nous constatons que les élus fournissent les données de leurs transactions boursières dans un délai moyen de 45,2 jours. La médiane est de 29 jours. Malgré le délai maximal légal, près de 12% des transactions sont déposées après les 45 jours<sup>7</sup>. Ceci soulève une question de conformité non négligeable.

#### 4. Concentration de l'activité

##### Par les élus

Michael T. McCaul (républicain, Chambre) concentre fortement l'activité et domine avec 6 988 transactions après nettoyage<sup>8</sup>.

Josh Gottheimer (démocrate, Chambre) vient en deuxième position et totalise 2 183 transactions, tandis que Lisa McClain (républicain, Chambre) arrive en troisième position avec 1 394 transactions.

Cette très forte concentration suppose des stratégies très différentes d'un élu à l'autre.

##### Par les titres

Nous observons une polarisation vers la technologie<sup>9</sup>. Celle-ci est cohérente avec l'intérêt politique pour la régulation du numérique. En effet, les titres les plus échangés sont les grandes capitalisations technologiques. Ci-après et dans l'ordre du nombre transactions sur le titre : MSFT (435 transactions), NVDA (364), AAPL (359), AMZN (314), VRNG (270), GOOGL (241).

---

<sup>6</sup> Voir annexe 1. 6) Volume de transactions estimé par année

<sup>7</sup> Voir annexe 1. 7) Distribution du délai de déclaration

<sup>8</sup> Voir annexe 1. 8) Élus les plus actifs par nombre de transactions

<sup>9</sup> Voir annexe 1. 9) 20 actions les plus échangées par les membres du Congrès

## 5. Performances individuelles

### Par les élus

Ont été sélectionnés, les élus avec au moins 10 transactions. L'analyse met en exergue des écarts de performance considérables. Par exemple<sup>10</sup>, Mark Green républicain, à la Chambre affiche un rendement excédentaire médian de +39,8 % sur 316 transactions. Tandis que Brian Mast également républicain, à la Chambre enregistre lui, une médiane de -162 % sur 24 transactions. Cette hétérogénéité interroge et invite à une analyse au niveau individuel plutôt que global.

### Par les titres

Ont été sélectionnés les titres affichant une performance et un minimum de 20 transactions. Nous observons une polarisation dans les semi-conducteurs MU (+501 %), KLAC (+313 %), LRCX (+222 %) et dans les valeurs industrielles : CW (+263 %), GEV (+191 %).

## 6. Remarques méthodologiques

Deux limites importantes doivent toujours être gardées lors de l'interprétation des résultats descriptifs.

Première limite, le rendement excédentaire utilisé ici est précalculé par Quiver Quantitative selon une fenêtre de 60 jours glissants. Une robustesse sur une fenêtre de 30 jours construite à partir de Yahoo Finance (`excess_return_30d`) sera mobilisée dans la suite de notre travail.

Deuxième limite, notre hypothèse d'un potentiel avantage informationnel n'est pas contredite par la sous-performance globale observée. La sous-performance n'empêche pas les profits anormaux concentrés sur un sous-ensemble de transactions ciblées. Lesdites transactions seraient celles réalisées dans des secteurs directement liés aux activités législatives des intéressés. C'est précisément ce que nous chercherons à établir dans les parties 1 (H1) et 2 (H2).

Ces résultats descriptifs motivent les analyses empiriques présentées dans les sections suivantes, portant à la fois sur l'existence éventuelle de rendements anormaux et sur l'impact du délai de publication sur l'exploitabilité des transactions parlementaires.

---

<sup>10</sup> Voir annexe 1. 10) Top 20 élus par rendements excédentaire médian

## V. Résultats empiriques

Nous proposons dans ce chapitre de tester nos 2 hypothèses ci-après dénommées H1 et H2. Pour rappel nos deux hypothèses sont :

H1 : Les membres du Congrès américain réalisent des rendements anormaux après leurs transactions boursières.

H2 : Une stratégie de copy trading fondée sur la publication des transactions parlementaires permet de générer des rendements anormaux malgré le délai imposé par le STOCK Act.

### 1. Partie 1 - H1

Rendements anormaux des transactions congressionnelles.

Return +1/+7/+30 jours vs. date de la transaction

Tableau 1 - Tests H1 sur les achats (excess return vs. zéro)

| Horizon   | N_achats | Moy. (%) | t-stat | p-value | Sig. |
|-----------|----------|----------|--------|---------|------|
| +1 jour   | 12 105   | -0,006   | -0,338 | 0,735   | Non  |
| +7 jours  | 12 105   | -0,028   | -0,708 | 0,479   | Non  |
| +30 jours | 12 105   | -0,128   | -1,608 | 0,108   | Non  |

Source : Yahoo Finance. Seuil de signification : 5 %.

Pour chacun des horizons testés, les rendements anormaux testés sont négatifs, tous les résultats sont statistiquement non-significatifs (seuil : 5%).

Ces résultats ne nous permettent pas de confirmer l'existence d'un avantage informationnel en ce qui concerne les achats analysés dans notre échantillon.

Buy vs. Sell

Tableau 2 - Tests t complets (excess\_return\_30d, Yahoo Finance, QuiverQuant)

| Groupe           | N      | Moy. (%) | Médiane | t-stat | p-value | Sig. |
|------------------|--------|----------|---------|--------|---------|------|
| Tous             | 26 525 | -0,181   | -0,191  | -3,41  | 0,001   | Oui  |
| Achats           | 12 105 | -0,128   | -0,201  | -1,61  | 0,108   | Non  |
| Ventes           | 14 420 | -0,227   | -0,183  | -3,164 | 0,002   | Oui  |
| Achats vs Ventes | -      | -        | -       | 0,928  | 0,353   | Non  |

L'analyse du tableau semble tendre vers un résultat statistiquement non-significatif. En effet, les titres ne surperforment pas durant les 30 jours après la transaction d'achat (moyenne : -0,128%).

Les titres vendus ont moyen de -0,227 %, significatif au seuil de 1 % ( $p = 0,002$ ). Concrètement, la valeur des titres que les élus du Congrès vendent continue de baisser en moyenne après la vente et ce durant les 30 jours suivants. Ce résultat suggère un mauvais timing et donc potentiellement l'absence d'un avantage informationnel.

La différence achats/ventes n'est pas significative ( $p = 0,353$ ). Ce résultat suggère que les élus ne disposent pas d'informations leur permettant d'anticiper les mouvements du marché, ni à l'achat, ni à la vente. Ceci pourrait suggérer un biais de sélection sectorielle commun aux deux types d'opération.

En conclusion, il n'existe pas de différence statistiquement significative entre les rendements post-achat et post-vente. Les élus ne sont pas systématiquement plus habiles dans un sens que dans l'autre.

En robustesse, la variable QuiverQuant indique une moyenne de -9,94 % pour les achats ( $p < 0,001$ ), soit un écart d'un facteur 75 par rapport à la mesure Yahoo Finance sur 30 jours (-0,13 %, non significatif). Le manque de documentation par Quiver Quantitative sur la fenêtre retenue, ne nous permet pas d'utiliser ce résultat directement. En effet, cette lacune empêche toute comparaison propre avec les autres mesures. De plus, sur la période 2021-2025, quelques méga-capitalisations technologiques ont concentré l'essentiel des gains du S&P 500. Un portefeuille diversifié sur l'ensemble des secteurs sous-performe ce type d'indice presque automatiquement, quelle que soit la qualité des décisions de trading. Le -9,94 % reflète donc probablement cette dynamique de marché autant qu'un signal propre aux transactions des élus.

### Analyse politique

#### a) House/Senate

Tableau 3 - Tests par sous-groupe (Achats, excess\_return\_30d,  $H_0 : \mu = 0$ )

| Groupe | N      | Moy. (%) | t-stat | p-value | Sig. |
|--------|--------|----------|--------|---------|------|
| House  | 10 992 | -0,139   | -1,66  | 0,097   | Non  |
| Senate | 1 113  | -0,017   | -0,068 | 0,946   | Non  |

Tableau 4 - Tests par sous-groupe (Achats, excess\_return\_30d,  $H_0 : \mu = 0$ )

| Groupe       | N     | Moy. (%) | t-stat | p-value | Sig. |
|--------------|-------|----------|--------|---------|------|
| Démocrates   | 4 675 | -0,414   | -3,19  | 0,001   | Oui  |
| Républicains | 7 402 | +0,056   | 0,562  | 0,574   | Non  |

Tableau 5 - Tests bilatéraux : chambre et parti (Achats, excess\_return\_30d)

| Comparaison <sup>11</sup>   | N_G1   | Moy_G1 | N_G2  | Moy_G2 | p-value | Sig. |
|-----------------------------|--------|--------|-------|--------|---------|------|
| House vs. Senate            | 10 992 | -0,139 | 1 113 | -0,017 | 0,647   | Non  |
| Démocrates vs. Républicains | 4 675  | -0,414 | 7 402 | +0,056 | 0,004   | Oui  |

Deux résultats se dégagent. Il n'existe pas de différence significative entre la Chambre et le Sénat ( $p = 0,647$ ). Il existe cependant une différence statistiquement significative ( $p = 0,004$ ) dans les rendements post-achat selon l'appartenance politique de l'élus. En effet, les démocrates sous-performent avec un rendement anormal moyen de -0,414%, statistiquement significatif au seuil de 1%. Les républicains quant à eux, affichent un rendement anormal moyen de +0,056% ( $>0$ ), statistiquement non-significatif.

Remarque : Les degrés de liberté des tests de Welch sont estimés par l'approximation de Satterthwaite. Pour les groupes aux effectifs similaires (achats vs ventes), l'approximation converge vers  $N_1+N_2-2$  ; pour les groupes très asymétriques (Chambre vs Sénat,  $N_1=10\,992$ ,  $N_2=1\,113$ ), le df est contraint par la strate la plus petite ( $df \approx 1\,112$ ), cohérent avec la p-value observée ( $p = .647$ ,  $t = -0.459$ ).

### Discussion H1

Sur la période 2021-2025, les achats des membres du Congrès américain ne génèrent pas de rendements anormaux positifs significatifs aux horizons +1, +7 et +30 jours. Ce résultat contraste avec Ziobrowski et al. (2004, 2011), qui documentaient des surperformances significatives avant le STOCK Act, mais s'aligne avec Eggers et Hainmueller (2013) et Belmont et al. (2020).

Différents facteurs expliquent ce résultat :

Premièrement, le STOCK Act a potentiellement porté ses fruits en réduisant l'avantage informationnel. En effet, le délai maximal de publication des transactions a permis de réduire la période d'exploitation d'une information privée. Eggers et Hainmueller (2013) arguent un début de déclin des rendements anormaux avant 2012. Ils suggèrent un probable effet de pression médiatique. Comme ils le soulignent, "*the disappearance of abnormal returns appears to coincide with increased public scrutiny of congressional stock trading*" (Eggers et Hainmueller, 2013, p. 546).

---

<sup>11</sup> G1= Représentants de la chambre - House  
G2= Sénateurs - Senat

Deuxièmement, pour les portefeuilles multisectoriels, le problème est amplifié par le contexte de marché 2021-2025 (post-covid). De fait, la concentration des gains du S&P 500 sur Apple, Microsoft, Nvidia et Alphabet a pour effet la sous performance des portefeuilles un tant soit peu diversifiés.

Enfin, 30 jours est peut-être trop court pour capturer un avantage informationnel lié à des cycles législatifs pluriannuels.

Ces facteurs s'accumulent. On ne peut pas conclure que l'avantage informationnel a disparu, on peut conclure qu'il n'est pas mesurable avec les horizons, le benchmark et la période retenus.

## 2. Partie 2 - H2

Le délai de publication affecte-t-il l'exploitabilité des transactions parlementaires ?

L'objectif de cette seconde analyse est d'évaluer dans quelle mesure le délai de publication imposé par le Stop Trading on Congressional Knowledge Act influence l'exploitabilité potentielle des transactions réalisées par les membres du Congrès américain.

Alors que l'hypothèse H1 reposait sur la date réelle de transaction, cette seconde approche adopte le point de vue d'un investisseur extérieur ne pouvant accéder aux informations qu'au moment de leur divulgation publique. L'enjeu est donc de déterminer si les transactions parlementaires conservent une valeur informationnelle après publication ou si celle-ci est progressivement dissipée par le délai de déclaration.

### Rendements après publication

Afin de reproduire les conditions réelles auxquelles ferait face un investisseur cherchant à imiter les transactions parlementaires, les rendements sont ici calculés à partir de la date officielle de publication des opérations et non plus à partir de leur date d'exécution réelle.

Les performances post-publication sont étudiées sur plusieurs horizons temporels :

- +1/+7/+30
- `excess_return_30d`

Les rendements observés sont comparés au rendement du S&P 500 sur la même période afin de construire un rendement excédentaire (*excess return*). La variable `excess_return_30d`,

reconstruite à partir des données Yahoo Finance, sert de mesure principale en raison de la transparence de sa méthodologie.

Tableau 6 - Rendements post-publication par horizon temporel

| Horizon   | N achats | Moy. achats (%) | p-value | Sig. | N ventes | Moy. Ventes (%) | p-value | Sig. |
|-----------|----------|-----------------|---------|------|----------|-----------------|---------|------|
| + 1 jour  | 12 105   | -0,006          | 0,735   | Non  | 14 422   | +0,003          | 0,853   | Non  |
| +7 jours  | 12 105   | -0,028          | 0,479   | Non  | 14 422   | +0,030          | 0,382   | Non  |
| +30 jours | 12 105   | -0,128          | 0,108   | Non  | 14 420   | -0,227          | 0,002   | Oui  |

L'analyse du tableau des résultats nous laisse observer une absence de surperformance robuste, après publication et ce peu importe l'horizon temporel testé. Les achats affichent un rendement excédentaire moyen de -0,128 % sur 30 jours, statistiquement non significatif ( $p = 0,108$ ). Les ventes présentent quant à elles un rendement légèrement négatif et statistiquement significatif ( $p\text{-value} = 0,002$ ).

Par ailleurs, aucune différence significative n'est observée entre les achats et les ventes ( $p\text{-value} = 0,353$ ). Ces résultats suggèrent que la publication des transactions parlementaires ne permet pas de générer des rendements anormaux systématiques.

#### Effet du reporting lag

Afin d'analyser plus spécifiquement l'impact du délai de publication, les transactions ont été regroupées selon plusieurs catégories de *reporting lag* correspondant au nombre de jours écoulés entre la transaction réelle et sa publication :

- 0–10 jours ;
- 11–30 jours ;
- 31–45 jours ;
- plus de 45 jours.

L'analyse porte uniquement sur les transactions d'achat, dans la mesure où celles-ci sont généralement considérées comme plus informatives dans la littérature financière.

Tableau 7 - Tests par catégorie de délai de déclaration

| Délai       | N achats | Moy. (%) | Médiane (%) | t-stat | p-value | Sig. |
|-------------|----------|----------|-------------|--------|---------|------|
| 0-10 jours  | 1 275    | +1,59    | -10,95      | 0,61   | 0,540   | Non  |
| 11-30 jours | 5 941    | -9,26    | -21,69      | -7,25  | < 0,001 | Oui  |
| 31-45 jours | 3 625    | -14,90   | -25,66      | -8,97  | < 0,001 | Oui  |
| 45+ jours   | 1 963    | -9,99    | -20,92      | -4,40  | < 0,001 | Oui  |

Notre analyse du tableau des résultats met en avant des différences significatives entre les 4 catégories de reporting lag. L'analyse montre un rendement positif de +1,59% dans le premier délai (0-10 jours), statistiquement non-significatif ( $p = 0.540$ ). Les trois autres délais affichent quant à eux des rendements moyens négatifs, statistiquement significatif<sup>12</sup>.

L'analyse de variance (ANOVA<sup>13</sup>) valide ces différences statistiquement significatives entre les groupes ( $F = 8,95$  ;  $p < 0,001$ ).

Les tests post-hoc de Bonferroni<sup>14</sup> illustrent notamment que les transactions déclarées dans les 0–10 jours se distinguent significativement des autres catégories.

Ces résultats laissent entendre que les transactions divulguées rapidement conservent davantage de valeur informationnelle que celles publiées plus tardivement. À l'inverse, l'allongement du délai de publication semble progressivement réduire l'exploitabilité potentielle des informations contenues dans les transactions parlementaires.

#### Application au copy trading

La simulation suivante traduit ces résultats en termes économiques concrets. L'investisseur est supposé prendre position le lendemain de la publication officielle d'une transaction d'achat, puis détenir le titre pendant 30 jours. Les performances sont comparées au S&P 500 sur la même période.

Les rendements excédentaires sont conditionnels<sup>15</sup> au délai de publication (simulation de copy trading, `excess_return_30d` depuis la date de transaction).

<sup>12</sup> Voir annexe 1. 11) Rendement excédentaire moyen par délai de déclaration

<sup>13</sup> Voir annexe 1. 12) ANOVA

<sup>14</sup> Voir annexe 1. 13) Bonferroni, Comparaison par paires

<sup>15</sup> Voir annexe 1. 14) Rendement moyen par stratégie de copy trading

Tableau 8 - Rendements excédentaires conditionnels au délai de publication

| Reporting lag (cumulé)      | N      | Lag Moy (j) | Moy. (%) | p-value | % positifs | IR     |
|-----------------------------|--------|-------------|----------|---------|------------|--------|
| Ultra-rapide ( $\leq 10$ j) | 1 196  | 5,4         | +0,477   | 0,056   | 51,0%      | 0,055  |
| Rapide ( $\leq 20$ j)       | 3 246  | 12,0        | +0,154   | 0,306   | 50,6%      | 0,018  |
| Standard ( $\leq 30$ j)     | 6 790  | 19,2        | -0,016   | 0,881   | 49,8%      | -0,002 |
| Légale ( $\leq 45$ j)       | 10 211 | 25,2        | -0,062   | 0,477   | 49,2%      | -0,007 |
| Aveugle (tous les achats)   | 11 807 | 41,8        | -0,144   | 0,074   | 48,7%      | -0,016 |

La stratégie ultra-rapide ( $\leq 10$  jours) affiche la meilleure performance (+0,477 %) mais n'atteint pas le seuil de significativité de 5 % ( $p = 0,056$ ). Cette stratégie est la seule à obtenir une moyenne positive et un ratio d'information positif. Elle ne couvre cependant qu'approximativement 10 % des achats totaux. Cette faible couverture réduit la diversification du portefeuille simulé et augmente la variance des estimations.

La stratégie aveugle reproduit par construction le résultat H1 sur les achats (même population, même variable, même test). Dès lors, l'ensemble des déclarations copiées sans filtre, ne bat pas le marché.

L'analyse par année montre une forte hétérogénéité temporelle. L'année 2022, enregistre 216 transactions avec un rendement moyen de +1,921 % ( $p\text{-value} < 0,001$ ). L'année 2025, recense 457 transactions qui performant à +0,857 % ( $p\text{-value} = 0,032$ ). Ces deux exercices positifs contrastent avec les années 2021, 2023 et 2024 qui sont non significatives et ont des rendements légèrement négatifs. La performance agrégée masque donc le signal positif concentré sur les années 2022 et 2025.

Un point méthodologique mérite d'être souligné. Les transactions dont le reporting\_lag dépasse 30 jours, ont une fenêtre de détention qui est expirée avant la publication. Une stratégie de copy trading à 30 jours est dès lors inapplicable. Le tableau ci-dessous décompose les rendements selon ce critère:

Tableau 9 - Rendements selon la disponibilité de la fenêtre de détention

| Groupe                                   | N     | Moy. (%) | p-value |
|------------------------------------------|-------|----------|---------|
| Fenêtre active ( $\text{lag} \leq 30$ j) | 6 790 | -0,016   | 0,881   |
| Fenêtre expirée ( $\text{lag} > 30$ j)   | 5 017 | -0,317   | 0,010   |

Les transactions dont la fenêtre est expirée sous-performent significativement (-0,317 %,  $p = 0,010$ ), tandis que celles dont la fenêtre est encore active ne se distinguent pas significativement de zéro. Ce résultat confirme que le délai de publication est le principal déterminant de l'exploitabilité des transactions parlementaires.

La rentabilité agrégée de la stratégie reste limitée. La simulation est volontairement simplifiée et ne tient pas compte des coûts de transaction, des problèmes de liquidité ni des contraintes d'exécution réelles.

### Discussion H2

Les résultats obtenus dans cette seconde analyse permettent d'apporter plusieurs enseignements importants concernant l'exploitabilité des transactions parlementaires américaines.

Premièrement, l'absence de surperformance significative après publication suggère que les mesures imposées par le STOCK Act, réduisent fortement la possibilité pour les investisseurs extérieurs, de reproduire systématiquement des rendements anormaux via des stratégies de copy trading.

Deuxièmement, l'existence possible d'un avantage informationnel, concentré sur certaines transactions, à savoir celles publiées rapidement, est révélée. Les opérations des élus au Congrès américain, déclarées dans un délai de 0 à 10 jours arborent des résultats supérieurs à celles des autres catégories de délai.

Enfin, l'avantage informationnel potentiel des élus pourrait donc exister de manière localisée et temporaire. Cet avantage se dissiperait progressivement dans le laps de temps qui sépare la transaction réelle et sa publication officielle.

Pour conclure, le STOCK Act ne semble toujours pas empêcher les soupçons d'avantage informationnel dont bénéficieraient les élus.

## VI. Conclusion

Ce mémoire portait sur les transactions boursières des membres du Congrès américain. La question de départ était double : ces élus tirent-ils un avantage informationnel de leur position pour réaliser des rendements anormaux, et cet avantage, s'il existe, peut-il être exploité par un investisseur extérieur une fois les transactions rendues publiques ?

Le cadre réglementaire autour de ces transactions a évolué depuis le STOCK Act de 2012, qui impose aux membres du Congrès de déclarer leurs opérations boursières dans un délai de 45 jours. Deux objectifs de cette réglementation étaient, d'une part, réduire le risque de délit d'initié et d'autre part, améliorer la transparence envers les citoyens. Cependant, des transactions suspectes continuent d'être relatées par les médias et les académistes. Notre travail s'est donc appuyé sur les données de déclaration disponibles pour la période 2021-2025 afin de tester empiriquement nos deux hypothèses.

H1 : Les membres du Congrès américain réalisent des rendements anormaux après leurs transactions boursières.

H2 : Une stratégie de copy trading fondée sur la publication des transactions parlementaires permet de générer des rendements anormaux malgré le délai imposé par le STOCK Act.

### 1. Réponse à H1 : les élus surperforment-ils le marché ?

H1 supposait que les membres du Congrès réalisent des rendements anormaux positifs à la suite de leurs transactions boursières. Les résultats ne permettent pas de valider cette hypothèse en agrégé.

Sur les horizons de +1, +7 et +30 jours, les achats parlementaires n'affichent pas de rendement excédentaire significativement différent de zéro. La mesure principale, construite à partir des données Yahoo Finance sur 30 jours (`excess_return_30d`), donne une moyenne de -0,128 % pour les achats, non significative ( $t(12104) = -1,608$ ,  $p = .108$ ). Les ventes ressortent à -0,227 % ( $p = .002$ ), ces titres continuent de baisser après la cession. Nous pouvons supposer qu'il s'agit plus d'un mauvais timing que d'un avantage informationnel sur les sorties.

Ces résultats agrégés s'accordent avec les travaux d'Eggers et Hainmueller (2013). En revanche, ils contrastent avec ceux de Ziobrowski et al. (2004, 2011). Ces derniers documentaient des

surperformances nettes dans les années 1990 et au début des années 2000, avant l'adoption du STOCK Act.

L'absence de surperformance au niveau agrégé ne signifie pourtant pas une absence totale d'avantage informationnel. L'analyse par sous-groupes révèle une hétérogénéité importante. Les achats des démocrates sous-performent nettement (-0,414 %,  $p = .001$ ). Ceux des républicains affichent +0,056 %, un résultat non significatif ( $p = .574$ ). Nous observons également des différences selon les secteurs (liés aux activités de l'élus) et les années (grands événements législatifs). Certaines transactions (liées aux activités du secteur ou aux grands événements) pourraient révéler un signal informationnel que l'approche agrégée ne capte pas. S'il existe un avantage, il serait localisé plutôt que généralisé à l'ensemble des élus.

H1 est donc rejetée. Sur la période et les horizons retenus, les membres du Congrès ne perçoivent pas de rendements anormaux positifs significatifs. Ce premier résultat reste néanmoins à nuancer. Des effets ponctuels et ciblés restent à identifier.

## 2. Réponse à H2 : le copy trading est-il viable ?

H2 supposait qu'une stratégie consistant à copier les transactions des élus permet de générer des rendements anormaux, et ce malgré le délai de publication imposé par le STOCK Act. Nos résultats ne permettent pas de défendre cette viabilité.

Aucune des stratégies que nous avons simulées ne bat le marché de façon robuste. La stratégie aveugle (réplique tous les achats sans filtre) offre un résultat de -0,144 % ( $p = .074$ ). La stratégie standard (limitée aux déclarations reçues dans les 30 jours suivant la transaction) est quasiment nulle (-0,016 %,  $p = .881$ ). Seule la stratégie ultra-rapide (limitée aux déclarations reçues dans les 10 premiers jours) affiche un rendement positif (+0,477 %). Cependant, la p-value reste à 0,056 et cette catégorie ne représente qu'environ 10 % des achats totaux. Ceci fragilise l'estimation par la variance élevée.

L'analyse de variance par catégorie de délai engendre le résultat le plus probant. Les rendements des transactions déclarées entre 0 et 10 jours (groupe 1) se distinguent significativement des trois autres groupes ( $F(3, 12800) = 8,95$ ,  $p < .001$  ; comparaisons post-hoc de Bonferroni significatives). Ce résultat suggère que la valeur informationnelle d'une transaction parlementaire se dissipe rapidement entre la date d'exécution et celle de publication. Passé un

délai de 10 jours, l'information perd de sa valeur. Notre analyse annuelle vient confirmer cette lecture. En effet, les transactions en 2022 et 2025 affichent des rendements positifs pour la stratégie ultra-rapide. Or les rendements des transactions des années 2021, 2023 et 2024 sont neutres ou légèrement négatifs. Ceci signale une viabilité fortement dépendante du contexte macroéconomique.

H2 est donc partiellement rejetée. Les divulgations publiques ne génèrent pas de rendements anormaux systématiques. Un signal résiduel existe pour les déclarations les plus rapides, mais il est marginalement significatif, restreint à une fraction minoritaire des achats et instable dans le temps. Une stratégie de copy trading systématique ne semble pas viable en pratique.

### 3. Implications réglementaires

Ces résultats permettent d'évaluer partiellement l'efficacité du STOCK Act. L'obligation de publication des transactions dans un délai de 45 jours, semble avoir réduit la fenêtre d'exploitation de l'information. La transparence améliore donc bien l'intégrité apparente des marchés.

Néanmoins, les transactions déclarées dans les 0-10 jours surperforment les autres catégories. Indiquant qu'un avantage résiduel subsiste. Cet avantage résiduel profite aux élus ainsi qu'aux investisseurs les plus réactifs. Ne faut-il pas y voir un plaidoyer pour un raccourcissement du délai de déclaration. Plusieurs propositions législatives récentes plaident pour une obligation de déclaration en temps réel (United States Congress, 2024; Gillibrand, 2023; CongressFlow, 2026). Le PELOSI Act proposé en 2025, va plus loin en proposant l'interdiction de détention d'actions individuelles par les membres du Congrès et leur conjoint (Congressman Mark Alford, 2025).

Ce constat souligne également la nécessité d'un mécanisme de sanction effectif. En pratique, près de 12 % des transactions traitées dans notre mémoire sont déclarées hors délai (irrégularité sans conséquence systématique visible).

Le défaut de qualité et le manque de standardisation des données publiques posent par ailleurs un problème majeur (National Academy of Public Administration, 2013). En effet, les formats de déclaration sont multiples, les données sur les montants imprécises et certaines informations manquent de manière récurrente. Une amélioration des exigences de forme permettrait une surveillance plus efficace. Celle-ci faciliterait dès lors les travaux de recherche indépendants.

#### 4. Limites du mémoire

Ce travail se heurte à plusieurs limites qu'il convient de signaler.

La qualité des données est la plus contraignante. Les déclarations publiques souffrent d'imprécisions importantes. Les montants sont renseignés par fourchettes (\$1 000-\$15 000, \$15 001-\$50 000, etc.) plutôt qu'en valeur exacte. Les délais de déclaration ne sont pas toujours respectés et les identifiants de titres sont parfois manquants ou ambigus. Ces limites réduisent la précision des mesures de rendement et introduisent inévitablement du bruit dans les estimations.

Sur le plan méthodologique, l'approche retenue pour notre travail est corrélationnelle. La causalité ne peut être mise en avant entre les rendements anormaux observés et un avantage informationnel. Des problèmes de sélection des titres achetés ou des effets de composition du portefeuille liés à l'appartenance politique constituent des explications alternatives que nos résultats ne permettent pas d'écarter.

L'analyse sectorielle a été traitée de manière exploratoire. Une démarche rigoureuse nécessiterait de croiser les transactions avec les affectations en commission de chaque élu. En raison de la qualité des données récoltées, nous n'avons pas pu réaliser cela. Cette démarche constituerait cependant un bon niveau d'analyse pour détecter un avantage informationnel localisé.

Enfin, la simulation de copy trading ne tient pas compte de la réalité. En effet, elle repose sur des hypothèses simplificatrices. Nous supposons l'absence de coûts de transaction, une liquidité parfaite et un horizon de rendement uniforme. Ces facteurs réduiraient en pratique les performances des stratégies simulées, ce qui renforce la conclusion de non-viabilité.

#### 5. Perspectives de recherche

Les résultats obtenus ouvrent plusieurs pistes pour des travaux futurs.

La principale limite identifiée dans ce mémoire est l'absence d'identification causale. Un design en différences de différences (DiD), exploitant l'adoption du STOCK Act comme choc exogène ou comparant les membres des commissions sectorielles à ceux qui n'y siègent pas, permettrait

de remédier à cette faiblesse. Ce prisme semble le plus optimal pour isoler un effet propre à la position législative de l'écu.

Une étude d'événement formel, ancrée sur des dates précises (votes en commission, auditions, annonces réglementaires), serait mieux adaptée pour localiser temporellement le signal informationnel. Celle-ci permettrait ainsi de tester si les transactions parlementaires précèdent chaque fois les événements susceptibles d'influencer les cours.

La récolte et l'utilisation de données intra-journalières permettrait de mesurer les rendements sur des fenêtres beaucoup plus courtes. Cette approche serait particulièrement utile pour les transactions déclarées rapidement. En effet, celles-ci semblent profiter d'un signal concentré dans les premiers jours.

L'intégration des coûts de transaction et des contraintes de liquidité dans la simulation de copy trading. En tenant compte de ces données, les simulations permettraient une évaluation de la viabilité réelle des stratégies.

## 6. Conclusion générale

Sur la période 2021-2025, l'avantage informationnel des membres du Congrès américain ne semble ni généralisé ni systématique. Les rendements anormaux positifs documentés avant le STOCK Act ne se retrouvent pas dans les données récentes, ce qui suggère que la loi a effectivement réduit la fenêtre d'exploitation d'une information privée.

Certains résultats invitent cependant à ne pas conclure trop rapidement. Les transactions déclarées très rapidement gardent une valeur informationnelle résiduelle. Des effets localisés, (au niveau d'élus particuliers, de secteurs ou de périodes spécifiques,) restent plausibles sans avoir été pleinement mesurés. L'absence de preuve n'est pas la preuve de l'absence.

Le délai de publication est la variable centrale de notre mémoire. Il conditionne l'exploitabilité des transactions parlementaires par les investisseurs extérieurs. Nous pensons que c'est sur ce point que les efforts réglementaires devraient continuer à progresser.

## VII. Références

- Akerlof, G. A. (1970). The Market for “Lemons”: Quality Uncertainty and the Market Mechanism. *The Quarterly Journal of Economics*, 84(3), 488–500.  
<https://doi.org/10.2307/1879431>
- Belmont, W., Sacerdote, B., Sehgal, R., & Van Hoek, I. (2020). Relief Rally : Senators As Feckless As the Rest of Us at Stock Picking. National Bureau Of Economic Research.  
<https://doi.org/10.3386/w26975>
- Congressman Mark Alford. (2025, 14 mai). *Alford Introduces PELOSI Act to Ban Members of Congress from Owning, Selling Individual Stocks*.  
<https://alford.house.gov/news/documentsingle.aspx?DocumentID=1187>
- Eggers, A. C., & Hainmueller, J. (2013). Capital Losses: The Mediocre Performance of Congressional Stock Portfolios. *The Journal Of Politics*, 75(2), 535-551.  
doi:10.1017/S0022381613000194
- Erdős, S., Papp, T., & Vörös, Z. (2022). The effects of community-based signals on investment decisions in copy trading. *Journal Of Behavioral And Experimental Economics*, 97, 101847. <https://doi.org/10.1016/j.socec.2022.101847>
- Faccio, M. (2002). Politically Connected Firms: Can They Squeeze the State? *SSRN Electronic Journal*. DOI:10.2139/ssrn.305099
- Fama, E. F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal Of Finance*, 25(2), 383-417. <https://doi.org/10.2307/2325486>
- Gillibrand, K. (2022, 9 février). *Senator Gillibrand, Representative Porter reintroduce STOCK Act 2.0 bicameral legislation would expand STOCK Act and issue new restrictions to prevent government officials from lining their pockets with insider knowledge*. U.S. Senate. <https://www.gillibrand.senate.gov/news/press/release/senator-gillibrand-representative-porter-reintroduce-stock-act-20-bicameral-legislation-would-expand-stock-act-and-issue-new-restrictions-to-prevent-government-officials-from-lining-their-pockets-with/>

- Hall, R. & Garcia, E. (2025, 13 janvier). *Dozens of Congress members outperformed the stock market in 2024. Here's who gained the most-And why*. Independent. <https://www.independent.co.uk/news/world/americas/us-politics/congress-stock-trading-nancy-pelosi-b2678752.html>
- Hanousek, J., Jo, H., Pantzalis, C., & Park, J. C. (2023). A Dilemma of Self-interest vs. Ethical Responsibilities in Political Insider Trading. *Journal of Business Ethics*, 187(1), 137-167. doi: 10.1007/s10551-022-05265-0
- Hanousek Jr., J., Park, J. C., & Pantzalis, C. (2021). Political Insider Trading: A narrow versus comprehensive approach. *SSRN Electronic Journal*. DOI: 10.2139/ssrn.3808702
- Harnisch, K. J., Rosen, K. M., Ramer, S., Tang, W., Harnisch, K. J., Rosen, K. M., Ramer, S., Tang, W., Harnisch, K. J., Rosen, K. M., Ramer, S., & Tang, W. (2020, 25 mars). *What is the STOCK Act?* Global Regulation Tomorrow. Consulté le 15 avril 2026 sur <https://www.regulationtomorrow.com/2020/03/what-is-the-stock-act/>
- Lakonishok, J., & Lee, I. (2001). Are Insider Trades Informative? *Review of Financial Studies*, 14(1), 79-111. DOI: 10.1093/rfs/14.1.79
- Mann, H. B., & Whitney, D. R. (1947). On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other. *The Annals of Mathematical Statistics*, 18(1), 50–60. DOI: 10.1214/aoms/1177730491
- National Academy of Public Administration. (2013, mars). *The STOCK Act: An independent review of the impact of providing personally identifiable financial information online*. <https://s3.us-west-2.amazonaws.com/napa-2021/studies/the-stock-act-an-independent-review-of-the-impact-of-providing-personally-i/STOCKactFinal1.pdf>
- Satterthwaite, F. E. (1946). An Approximate Distribution of Estimates of Variance Components. *Biometrics Bulletin*, 2(6), 110–114. <https://doi.org/10.2307/3002019>
- The delay problem. (2026, 26 mars). *CongressFlow*. <https://congressflow.com/learn/delay-problem>
- United States Congress. (2024). *Immediate Stock Disclosure Act* (H.R. 7485, 118th Congress). Congress.gov. <https://www.congress.gov/bill/118th-congress/house-bill/7485/text>

United States Congress. (2012). *Stop Trading on Congressional Knowledge Act of 2012* (S. 2038, 112th Congress). Congress.gov.

<https://www.congress.gov/bill/112th-congress/senate-bill/2038>

U.S. Department of the Interior. (2017). *STOCK Act summary*.

[https://www.doi.gov/sites/doi.gov/files/migrated/ethics/forms/upload/stock\\_act\\_summary\\_5\\_22\\_2017.pdf](https://www.doi.gov/sites/doi.gov/files/migrated/ethics/forms/upload/stock_act_summary_5_22_2017.pdf)

Ziobrowski, A. J., Boyd, J. W., Cheng, P., & Ziobrowski, B. J. (2011). Abnormal returns from the common stock investments of members of the US House of Representatives.

*Business and Politics*, 13(1), 1-22. DOI:10.2202/1469-3569.1308

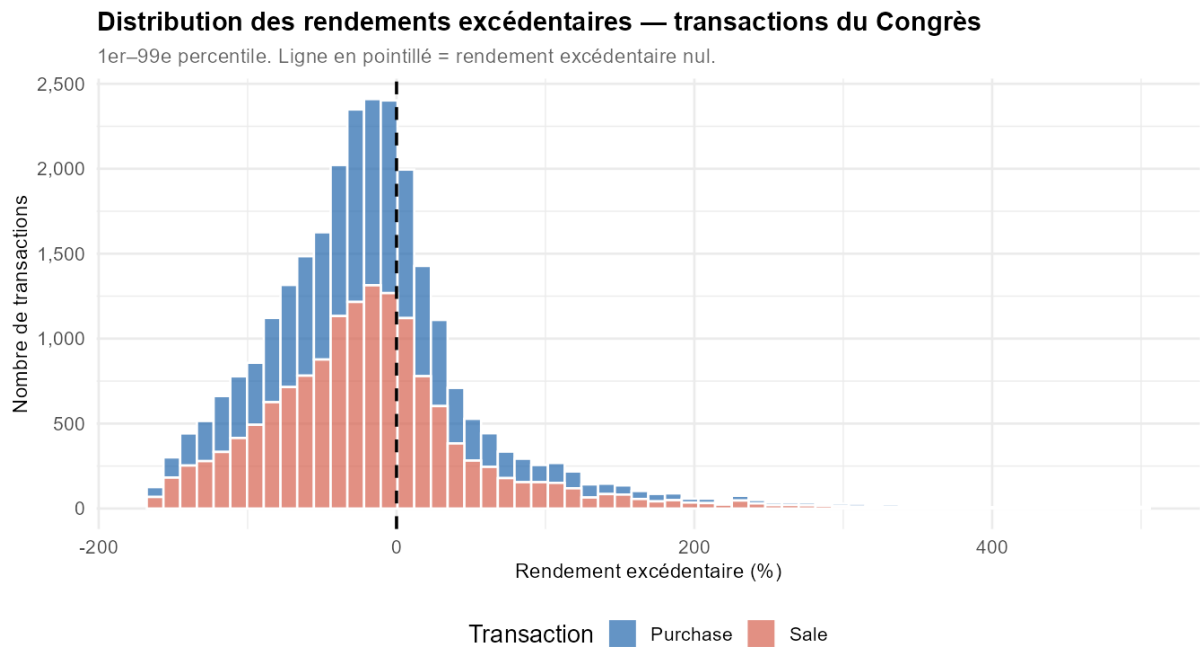
Ziobrowski, A. J., Cheng, P., Boyd, J. W., & Ziobrowski, B. J. (2004). *Abnormal Returns from the Common Stock Investments of the U.S. Senate*.

<https://doi.org/10.1017/S0022109000003161>

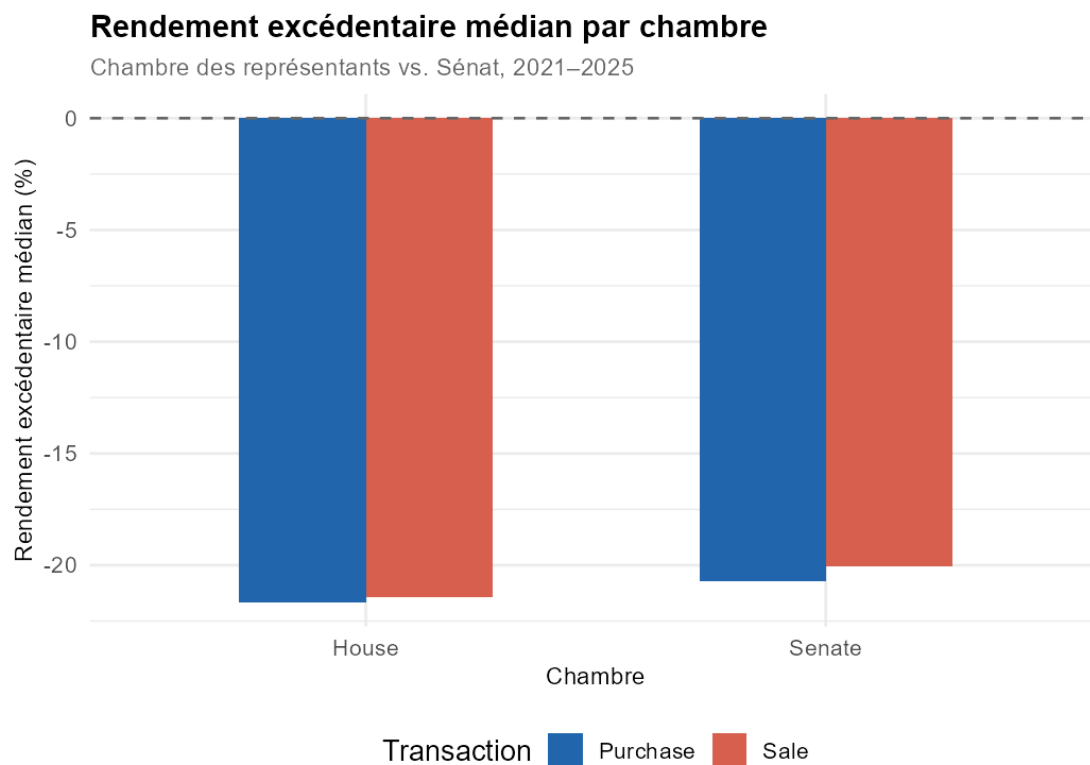
## VIII. Annexes

### 1. Graphiques des statistiques

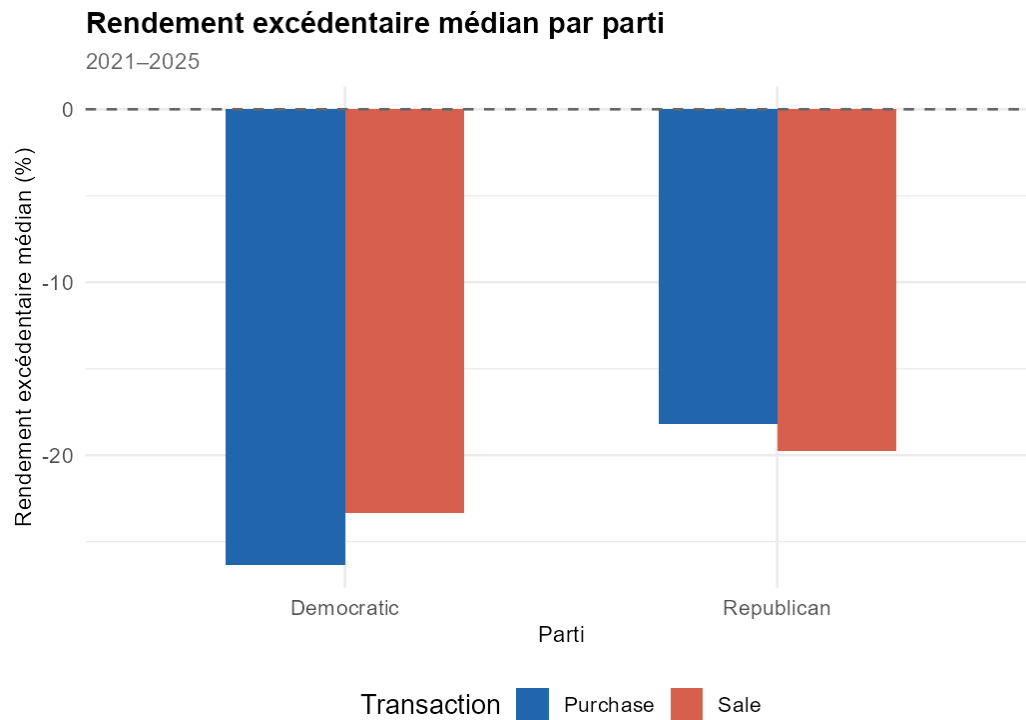
#### 1) Distribution des rendements excédentaires – transactions du Congrès



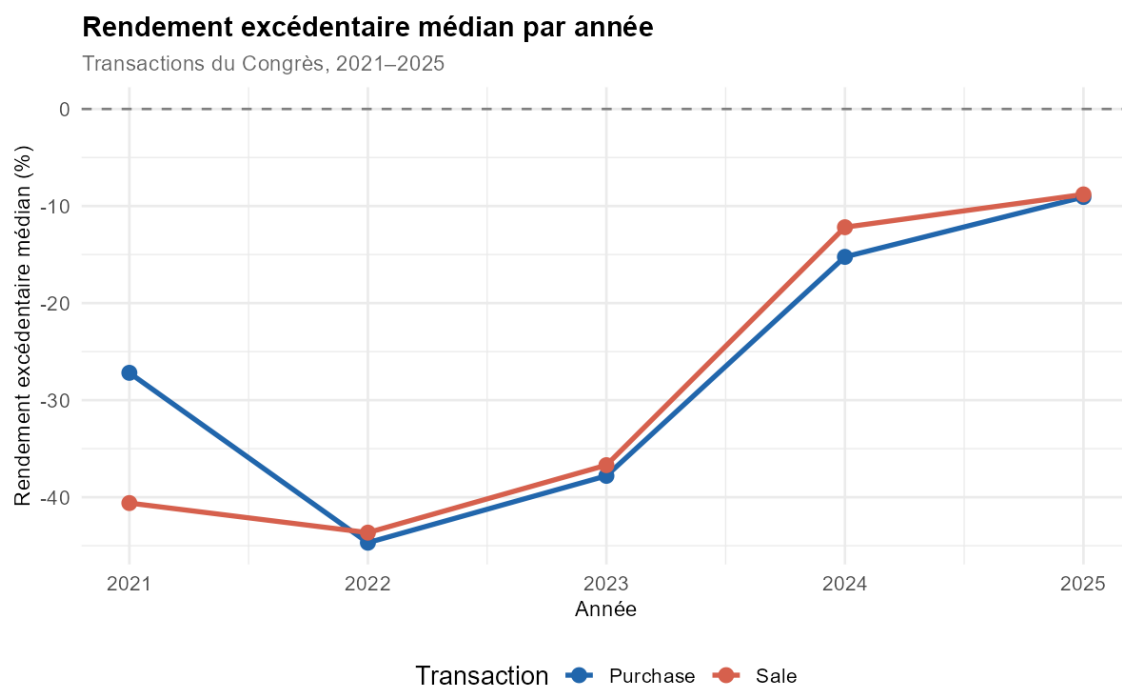
#### 2) Rendements excédentaires médian par chambre



### 3) Rendements excédentaires par parti



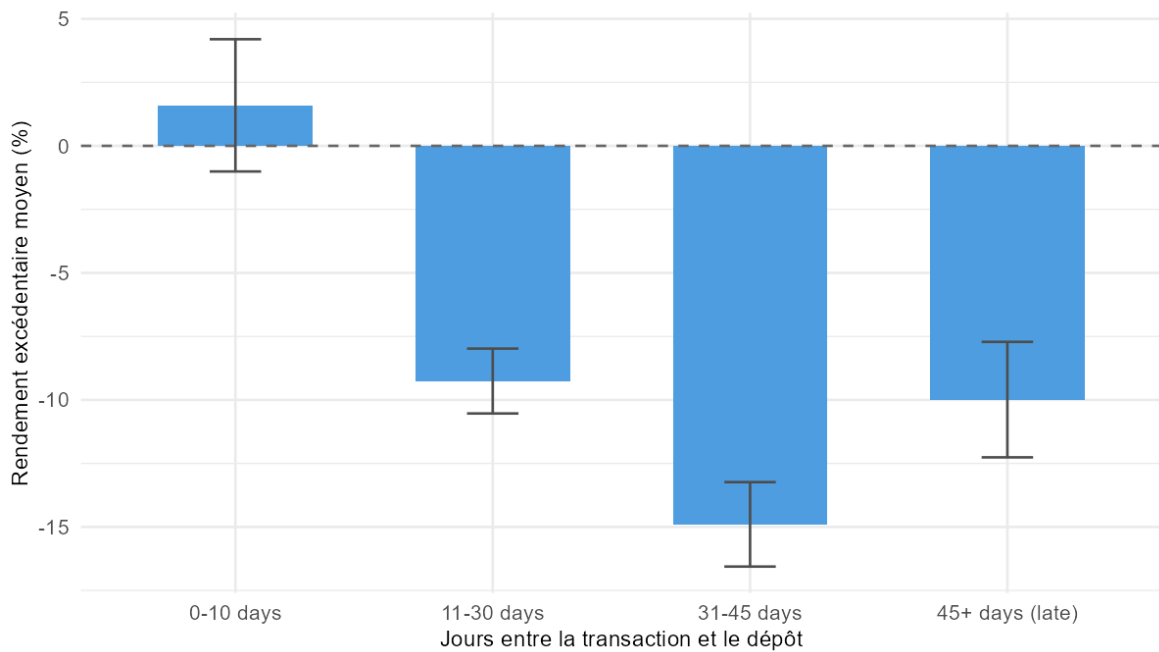
### 4) Rendement excédentaire médian par année



### 5) Rendement excédentaire moyen par délai de déclaration

#### Rendement excédentaire moyen par délai de déclaration (achats)

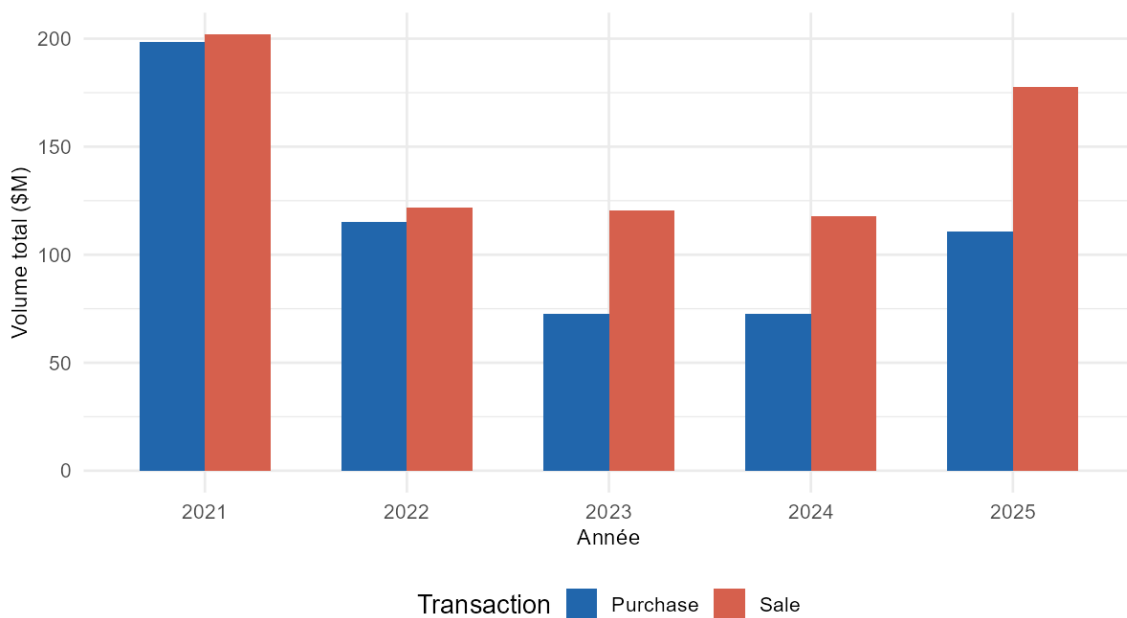
Barres d'erreur =  $\pm 1$  ET.  $H_0$  : rendement moyen = 0.



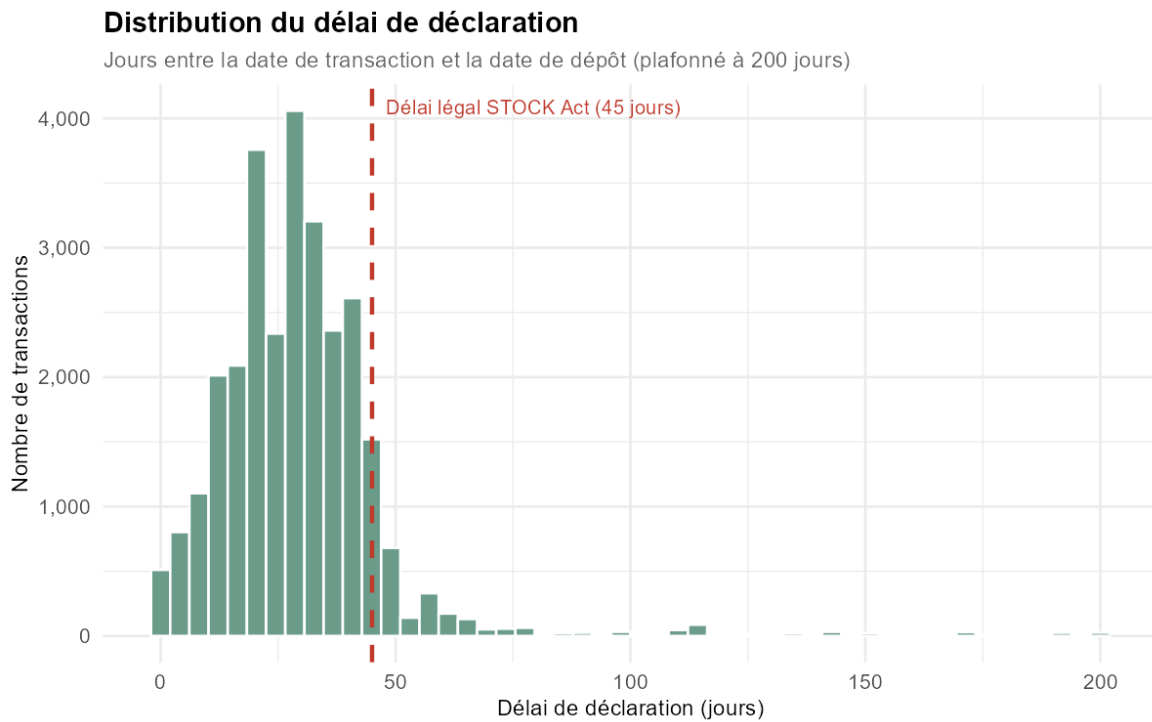
### 6) Volume de transactions estimés par année

#### Volume de transactions estimé par année (\$M)

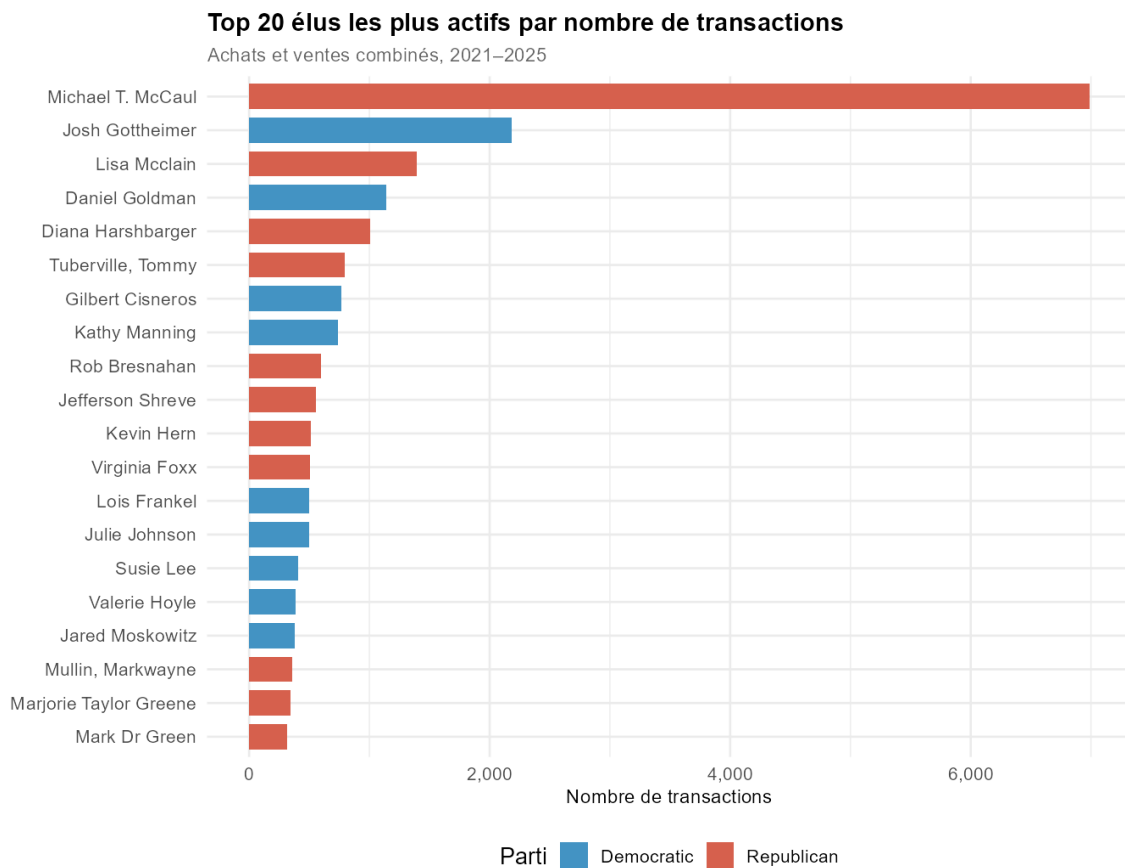
Basé sur la valeur médiane des fourchettes de taille déclarées



## 7) Distribution des délai de déclaration



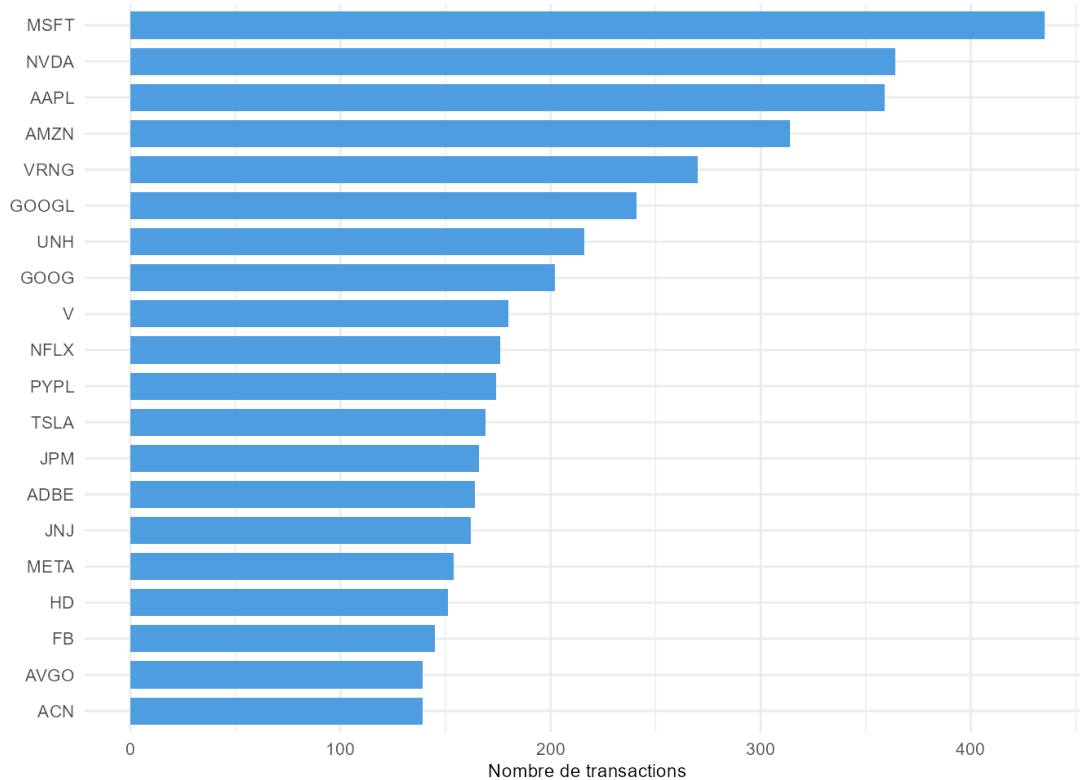
## 8) Élus les plus actifs par nombre de transactions



### 9) 20 actions les plus échangées par les membres du Congrès

#### Les 20 actions les plus échangées par les membres du Congrès

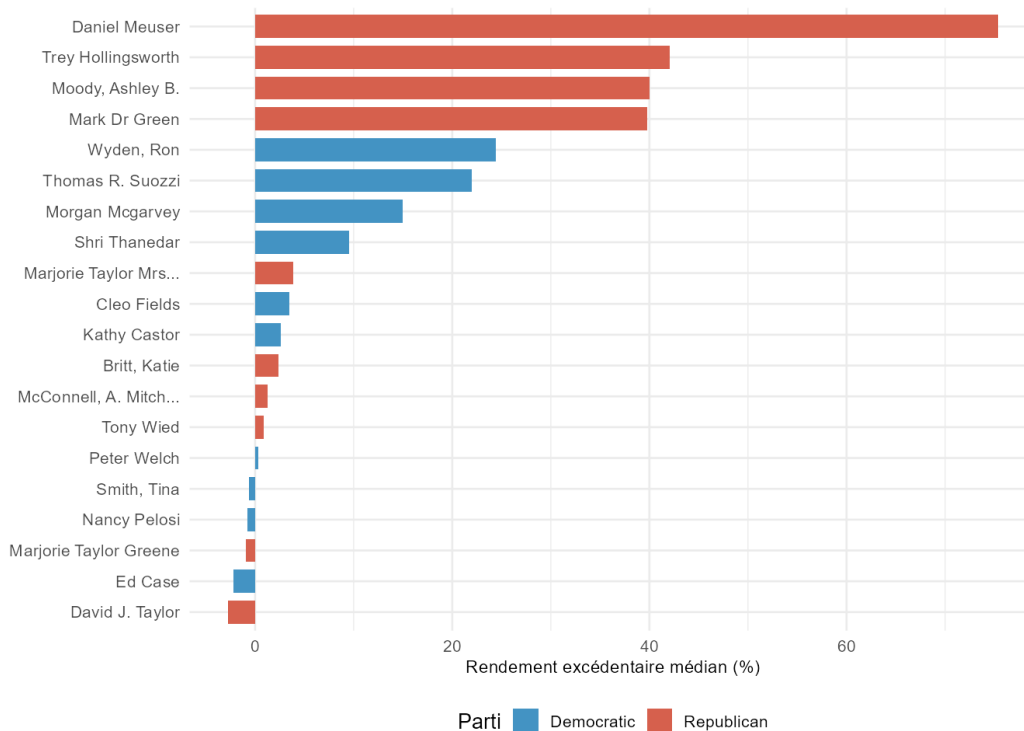
2021–2025 (achats et ventes combinés)



### 10) Top 20 élus par rendement excédentaire médian

#### Top 20 élus par rendement excédentaire médian

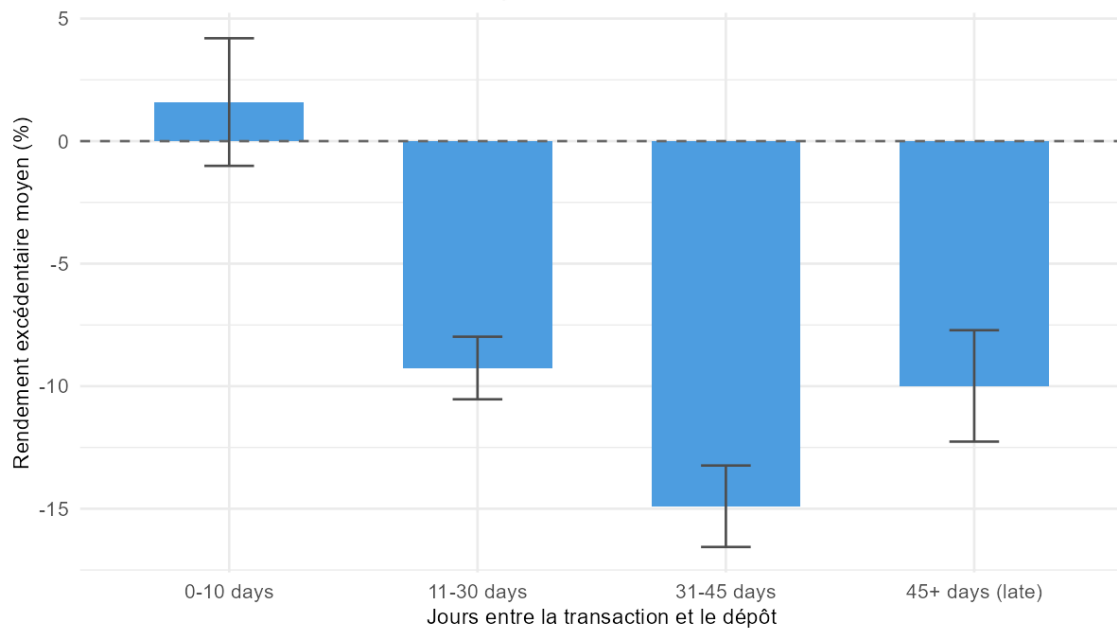
Minimum 10 transactions, 2021–2025



### 11) Rendement excédentaire moyen par délai de déclaration

#### Rendement excédentaire moyen par délai de déclaration (achats)

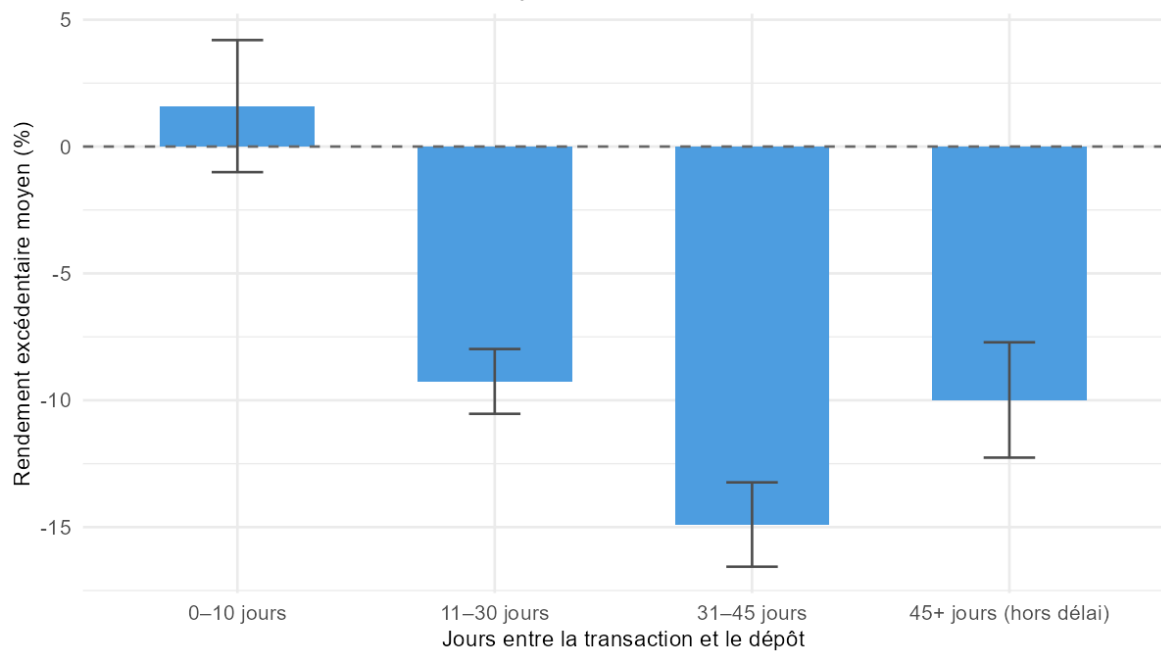
Barres d'erreur =  $\pm 1$  ET. H0 : rendement moyen = 0.



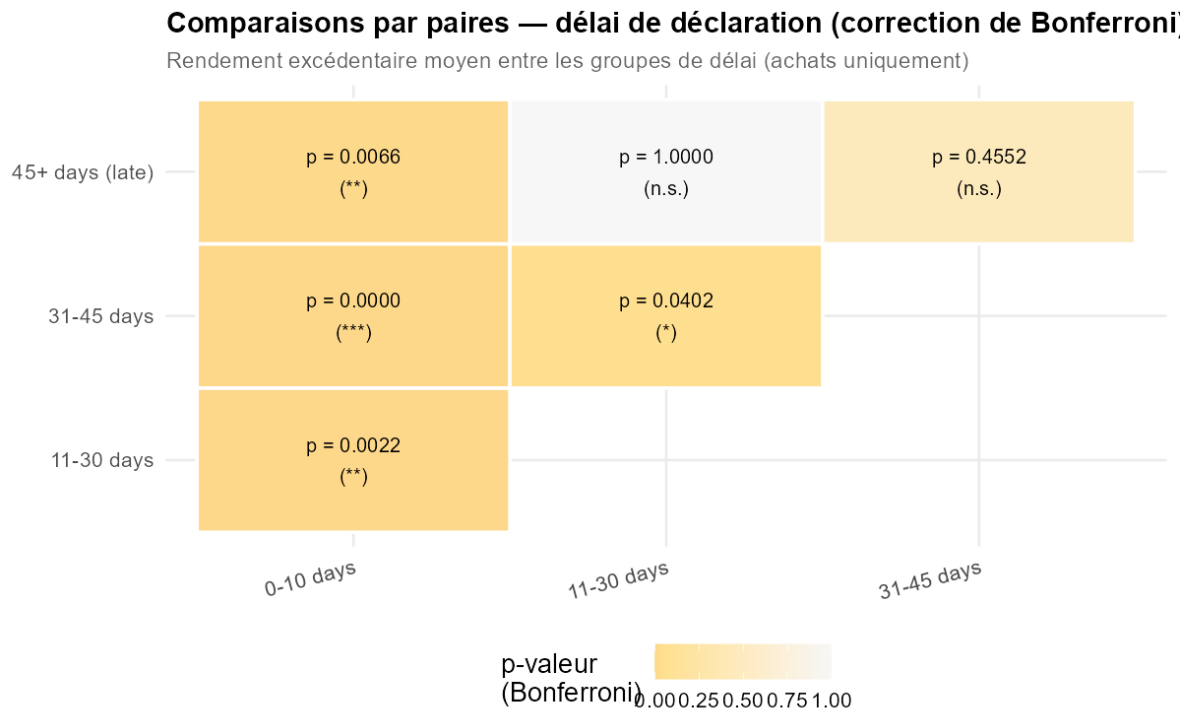
### 12) ANOVA

#### ANOVA — Rendement excédentaire moyen par délai de déclaration (achats)

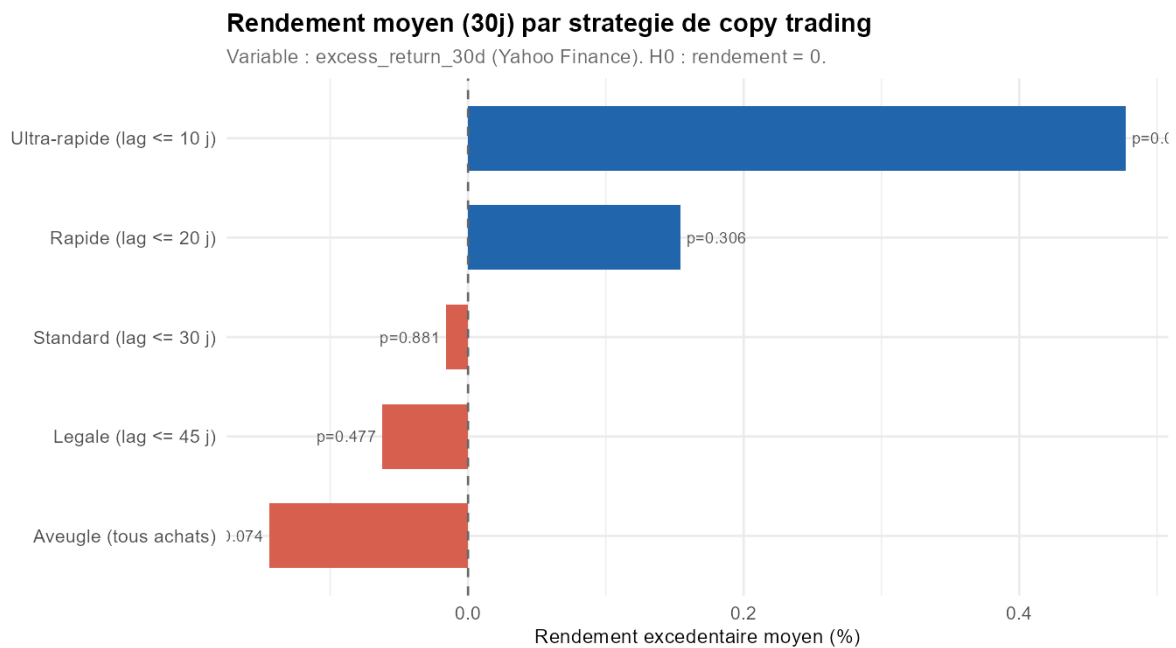
Barres d'erreur =  $\pm 1$  ET. H0 : rendement moyen = 0.



### 13) Bonferroni, Comparaison par paires



### 14) Rendement moyen par stratégie de copy trading



## 2. DATA

Les annexes de cette section reprennent toutes les données mobilisées dans ce mémoire et sont disponibles sur la plateforme DIAL. Ces annexes ont été déposées en même temps que le mémoire.

### 1) ANNEXE 1

Données brutes de Quiver Quantitative

### 2) ANNEXE 2

Quiver Quantitative V2 (Première version modifiée et nettoyée)

### 3) ANNEXE 3

Quiver Quantitative +Yahoo (Données enrichies avec les rendements du S&P 500)

### 4) ANNEXE 4

Data final version

## 3. Codes

Les annexes de cette section reprennent les codes qui ont été utilisé pour mener à bien ce mémoire. Ces codes/scripts ont été réalisés avec l'aide de Claude (Anthropic) afin de faciliter le bon fonctionnement de ceux-ci.

### 1) ANNEXE 5

Le script ci-dessous nous à permis de télécharger les données de Quiver Quantitative.

```
Quiver Quantitative - Bulk Congress Trading Fetcher
=====
Endpoint: GET https://api.quiverquant.com/beta/bulk/congresstrading
Docs: https://api.quiverquant.com/docs/#/operations/bulk_congresstrading_retrieve
Usage:
    pip install requests pandas
    export QUIVER_TOKEN="your_token"
    python quiver_congress_trades.py
"""
import os
import requests
import pandas as pd
from datetime import datetime, timedelta

# --- Config -----
QUIVER_TOKEN = os.getenv("QUIVER_TOKEN", "YOUR QUIVER_TOKEN_HERE")
YEARS_BACK = 5 # set None to keep all records
OUTPUT_CSV = "quiver_congressional_trades.csv"

# --- Fetch -----
def fetch_trades() -> pd.DataFrame:
```

```

url = "https://api.quiverquant.com/beta/bulk/congresstrading"
headers = {
    "Accept": "application/json",
    "Authorization": f"Bearer {QUIVER_TOKEN}",
}

print("Fetching bulk congress trading data...")
resp = requests.get(url, headers=headers, timeout=60)

if resp.status_code == 401:
    raise SystemExit("✗ Unauthorised – check your QUIVER_TOKEN.")
if resp.status_code == 403:
    raise SystemExit("✗ Forbidden – your plan may not include bulk access.")
resp.raise_for_status()

df = pd.DataFrame(resp.json())
print(f" → {len(df):,} records received.")
return df

# — Filter —
def filter_by_date(df: pd.DataFrame, years: int) -> pd.DataFrame:
    if years is None:
        return df
    df["Date"] = pd.to_datetime(df["Date"], errors="coerce")
    cutoff = datetime.now() - timedelta(days=365 * years)
    filtered = df[df["Date"] >= cutoff].copy()
    print(f" → {len(filtered):,} records within the last {years} years.")
    return filtered

# — Main —
def main():
    df = fetch_trades()
    df = filter_by_date(df, YEARS_BACK)
    df.to_csv(OUTPUT_CSV, index=False)
    print(f"\n✅ Saved {len(df):,} records to '{OUTPUT_CSV}'")
    print(df.head(3).to_string())

if __name__ == "__main__":
    main()

```

## 2) ANNEXE 6

Le script ci-dessous nous a permis de télécharger et ajouter les rendements (de Yahoo) sur base des tickers et des dates.

```

# Mémoire LSM 2025-2026 — Lucas Dinis Pires & Ysaline Koch
# Téléchargement des prix Yahoo Finance et calcul des rendements (1j, 7j, 30j)
# Stratégie : un téléchargement par ticker unique (~2 800 appels au lieu de ~30 000)
# Output : DATAMEMOIRE2_with_returns.csv

import pandas as pd
import yfinance as yf
import time
import os

```

```

# — Paramètres —————
INPUT_FILE = "C:/Users/LDP/Downloads/DATAMEMOIRE2.csv"
OUTPUT_FILE = "C:/Users/LDP/Downloads/DATAMEMOIRE2_with_returns.csv"
CACHE_FILE = "C:/Users/LDP/Downloads/price_cache.pkl"
DATE_COL = "Traded"
TICKER_COL = "Ticker"
WINDOWS = [1, 7, 30]
BATCH_SIZE = 50
SLEEP_SEC = 3.0
MAX_RETRIES = 3
START_BUFFER = 5
END_BUFFER = 45

# — 1. Chargement des données —————
print("Chargement des données...")
df = pd.read_csv(INPUT_FILE)
df[DATE_COL] = pd.to_datetime(df[DATE_COL])

print(f" Lignes      : {len(df):,}")
print(f" Tickers uniques: {df[TICKER_COL].nunique():,}")
print(f" Période       : {df[DATE_COL].min().date()} → {df[DATE_COL].max().date()}")

df = df[df[TICKER_COL].notna()]
df[TICKER_COL] = df[TICKER_COL].str.strip().str.upper()

BAD_TICKERS = {"571903BM4", "3.MONTH", "MATURE", "BRK.B"}
df[TICKER_COL] = df[TICKER_COL].replace("BRK.B", "BRK-B") # Yahoo Finance utilise BRK-B
df = df[~df[TICKER_COL].isin(BAD_TICKERS)]

unique_tickers = sorted(df[TICKER_COL].unique())
print(f" Tickers valides: {len(unique_tickers):,}")

# — 2. Téléchargement des prix (par lots, avec cache) —————
if os.path.exists(CACHE_FILE):
    print(f"\nReprise depuis le cache {CACHE_FILE}...")
    price_cache = pd.read_pickle(CACHE_FILE)
    cached_tickers = set(
        price_cache.columns.get_level_values(0).unique()
        if isinstance(price_cache.columns, pd.MultiIndex)
        else price_cache.columns
    )
    tickers_to_download = [t for t in unique_tickers if t not in cached_tickers]
    print(f" En cache : {len(cached_tickers):,} | À télécharger : {len(tickers_to_download):,}")
else:
    price_cache = pd.DataFrame()
    tickers_to_download = unique_tickers
    print(f"\nAucun cache — téléchargement de {len(tickers_to_download):,} tickers...")

global_start = (df[DATE_COL].min() - pd.Timedelta(days=START_BUFFER)).strftime("%Y-%m-%d")
global_end = (df[DATE_COL].max() + pd.Timedelta(days=END_BUFFER)).strftime("%Y-%m-%d")
print(f" Fenêtre : {global_start} → {global_end}")

batches = [tickers_to_download[i:i+BATCH_SIZE]]

```

```

        for i in range(0, len(tickers_to_download), BATCH_SIZE)]

all_prices = []
failed_tickers = []

for i, batch in enumerate(batches):
    print(f" Lot {i+1}/{len(batches)}: {batch[:5]} {'...' if len(batch) > 5 else ''}", end=" ")
    success = False
    for attempt in range(MAX_RETRIES):
        try:
            raw = yf.download(batch, start=global_start, end=global_end,
                              progress=False, auto_adjust=True, threads=True)
            if raw.empty:
                print("→ vide")
                failed_tickers.extend(batch)
            else:
                if isinstance(raw.columns, pd.MultiIndex):
                    close = raw["Close"]
                else:
                    close = raw[["Close"]].rename(columns={"Close": batch[0]})
                all_prices.append(close)
                print(f"→ ok ( {len(close)} jours, {close.shape[1]} tickers)")
                success = True
                break
        except Exception as e:
            err_str = str(e)
            if "Rate" in err_str or "429" in err_str or "Too Many" in err_str:
                wait = SLEEP_SEC * (attempt + 1) * 5
                print(f"→ rate limit, attente {wait:.0f}s (essai {attempt+1}/{MAX_RETRIES})...", end=" ")
                time.sleep(wait)
            else:
                print(f"→ ERREUR : {e}")
                failed_tickers.extend(batch)
                break
    if not success:
        print(f"→ abandon après {MAX_RETRIES} tentatives")
        failed_tickers.extend(batch)
        time.sleep(SLEEP_SEC)

if all_prices:
    new_prices = pd.concat(all_prices, axis=1)
    new_prices = new_prices.loc[:, ~new_prices.columns.duplicated()]
    if not price_cache.empty:
        price_cache = pd.concat([price_cache, new_prices], axis=1)
        price_cache = price_cache.loc[:, ~price_cache.columns.duplicated()]
    else:
        price_cache = new_prices

# Sauvegarde avant SPY — évite de tout reperdre si le download SPY échoue
price_cache.to_pickle(CACHE_FILE)
print(f"\nCache sauvegardé : {price_cache.shape[1]:,} tickers × {len(price_cache):,} jours")

if failed_tickers:
    print(f"Tickers échoués ( {len(failed_tickers)} ) : {failed_tickers[:20]}")

print("\nTéléchargement du benchmark SPY...")
spy_raw = yf.download("SPY", start=global_start, end=global_end,

```

```

        progress=False, auto_adjust=True)
if isinstance(spy_raw.columns, pd.MultiIndex):
    spy_prices = spy_raw["Close"].squeeze().rename("SPY")
else:
    spy_prices = spy_raw["Close"].squeeze().rename("SPY")
spy_prices = spy_prices.squeeze()
spy_prices.index = pd.to_datetime(spy_prices.index)
print(f" SPY : {len(spy_prices)} jours de bourse chargés")

# — 3. Calcul des rendements —————
print("\nCalcul des rendements...")

def get_next_available_price(prices_series, from_date, offset_days):
    """Prix de clôture au jour from_date + offset_days (saute WE et jours fériés)."""
    target = from_date + pd.Timedelta(days=offset_days)
    for extra in range(6):
        candidate = target + pd.Timedelta(days=extra)
        if candidate in prices_series.index:
            val = prices_series.loc[candidate]
            if pd.notna(val):
                return float(val)
    return None

def get_entry_price(prices_series, trade_date):
    """Prix d'entrée à la date de transaction (±2 jours pour WE/fériés)."""
    for offset in [0, 1, -1, 2, -2]:
        candidate = trade_date + pd.Timedelta(days=offset)
        if candidate in prices_series.index:
            val = prices_series.loc[candidate]
            if pd.notna(val):
                return float(val), candidate
    return None, None

price_cache.index = pd.to_datetime(price_cache.index)
spy_prices.index = pd.to_datetime(spy_prices.index)

for w in WINDOWS:
    df[f"return_{w}d"] = None
    df[f"excess_return_{w}d"] = None
df["entry_price"] = None
df["entry_date"] = None
df["ticker_found"] = False

total = len(df)
for idx, row in df.iterrows():
    ticker = row[TICKER_COL]
    trade_date = row[DATE_COL]

    if idx % 500 == 0:
        print(f" Ligne {idx:,} / {total:,}...")

    if ticker not in price_cache.columns:
        continue

    prices = price_cache[ticker]
    p0, actual_entry = get_entry_price(prices, trade_date)

```

```

if p0 is None or p0 == 0:
    continue

df.at[idx, "entry_price"] = p0
df.at[idx, "entry_date"] = actual_entry
df.at[idx, "ticker_found"] = True

spy0, _ = get_entry_price(spy_prices, trade_date)

for w in WINDOWS:
    p_exit = get_next_available_price(prices, trade_date, w)
    spy_exit = get_next_available_price(spy_prices, trade_date, w)

    if p_exit is not None and p_exit != 0:
        raw_ret = (p_exit - p0) / p0 * 100
        df.at[idx, f'return_{w}d"] = round(raw_ret, 4)
    if spy0 is not None and spy_exit is not None and spy0 != 0:
        spy_ret = (spy_exit - spy0) / spy0 * 100
        df.at[idx, f'excess_return_{w}d"] = round(raw_ret - spy_ret, 4)

print(f"\nCouverture :")
for w in WINDOWS:
    valid = df[f'return_{w}d"].notna().sum()
    print(f" return_{w}d : {valid:,} / {total:,} ({100*valid/total:.1f}%)")

# — 4. Winsorisation —————
print("\nWinsorisation au 1er-99e percentile...")

def winsorise(series):
    q_lo = series.quantile(0.01)
    q_hi = series.quantile(0.99)
    return series.clip(lower=q_lo, upper=q_hi)

return_cols = [c for c in df.columns if c.startswith("return_") or c.startswith("excess_return_")]
for col in return_cols:
    col_numeric = pd.to_numeric(df[col], errors="coerce")
    df[col + "_raw"] = col_numeric
    df[col] = winsorise(col_numeric)

print(f" Colonnes winsorisées : {return_cols}")

# — 5. Sauvegarde —————
df.to_csv(OUTPUT_FILE, index=False)
print(f"\n✓ Fichier sauvegardé : {OUTPUT_FILE}")
print(f" Lignes : {len(df):,} | Colonnes : {len(df.columns):,}")
print(f" Nouvelles colonnes : {[c for c in df.columns if 'return_' in c or c in ['entry_price', 'entry_date', 'ticker_found']]})")

print("\nÉchantillon de rendements calculés :")
sample_cols = [DATE_COL, TICKER_COL, "entry_price"] + [f'return_{w}d' for w in WINDOWS]
print(df[sample_cols].dropna(subset=["entry_price"]).head(10).to_string(index=False))

```

3) ANNEXE 7

Ce dernier Script R est le script principal du ce mémoire. Il nous a permis de nettoyer et analyser les données téléchargées via les deux scripts précédents.

```
# Mémoire LSM 2025-2026 – Lucas Dinis Pires & Ysaline Koch
# Analyse descriptive et exploratoire des transactions du Congrès (2021-2025)
# Sections : nettoyage · statistiques · visualisations · tests H1/H2 · copy trading

# — 0. Configuration —————

library(tidyverse)
library(scales)
library(knitr)
library(kableExtra)
library(lubridate)
library(RColorBrewer)

theme_set(
  theme_minimal(base_size = 13) +
  theme(
    plot.title = element_text(face = "bold", size = 14),
    plot.subtitle = element_text(colour = "grey40", size = 11),
    axis.title = element_text(size = 11),
    legend.position = "bottom"
  )
)

PLOTS_DIR <- "C:/Users/LDP/Documents/University-UCL/Master2 (2025-2026)/Thesis/congress_trading/data & scripts/plots"
TABLES_DIR <- "C:/Users/LDP/Documents/University-UCL/Master2 (2025-2026)/Thesis/congress_trading/data & scripts/tables"
DATA_DIR <- "C:/Users/LDP/Documents/University-UCL/Master2 (2025-2026)/Thesis/congress_trading/data & scripts/data"
dir.create(PLOTS_DIR, showWarnings = FALSE, recursive = TRUE)
dir.create(TABLES_DIR, showWarnings = FALSE, recursive = TRUE)

# — 1. Chargement et nettoyage —————
# Filtres : [F1] 2021-2025 · [F2] achat/vente · [F3] actions cotées
# [F4] winsorisation excess_return au 1er-99e percentile

raw <- read_csv(file.path(DATA_DIR, "DATAMEMOIRE2_with_returns.csv"),
  show_col_types = FALSE)

raw <- raw %>%
  mutate(
    Traded = as.Date(Traded),
    Filed = as.Date(Filed),
    excess_return = suppressWarnings(as.numeric(excess_return)),
    excess_return_30d = suppressWarnings(as.numeric(excess_return_30d))
  )

# [F1] Fenêtre temporelle
```

```

df <- raw %>%
  filter(year(Traded) >= 2021, year(Traded) <= 2025)
cat(sprintf("[F1] Après filtre dates (2021-2025)           : %d lignes\n", nrow(df)))

# Exclusion Ro Khanna (trading algorithmique, 15 218 opérations sur 2021-2025)
df <- df %>% filter(Name != "Ro Khanna")
cat(sprintf("[F1b] Après exclusion Ro Khanna           : %d lignes\n", nrow(df)))

# [F2] Type de transaction
df <- df %>%
  mutate(
    Transaction_clean = case_when(
      str_to_upper(Transaction) %in%
        c("PURCHASE", "Purchase") ~ "Purchase",
      str_to_upper(Transaction) %in%
        c("SALE", "SALE (FULL)", "SALE (PARTIAL)",
          "Sale", "Sale (Full)", "Sale (Partial)") ~ "Sale",
      TRUE ~ "Other"
    )
  ) %>%
  filter(Transaction_clean %in% c("Purchase", "Sale"))
cat(sprintf("[F2] Après filtre achat/vente           : %d lignes\n", nrow(df)))

# [F3] Actions uniquement
EQUITY_CODES <- c("ST", "Stock")
KNOWN_ETFS <- c(
  "SPY", "QQQ", "IWM", "GLD", "TLT", "VTI", "VOO", "VNQ", "AGG",
  "XLF", "XLE", "XLK", "XLV", "XLI", "XLB", "XLU", "XLP", "XLY",
  "EEM", "EFA", "HYG", "LQD", "SHY", "BND", "DIA", "MDY", "SLV"
)

df <- df %>%
  mutate(
    equity_status = case_when(
      TickerType %in% EQUITY_CODES ~ "equity_coded",
      is.na(TickerType) &
        !Ticker %in% KNOWN_ETFS &
        nchar(Ticker) <= 5 ~
        "equity_inferred",
      TRUE ~ "exclude"
    )
  ) %>%
  filter(equity_status != "exclude")

cat(sprintf("[F3] Après filtre actions           : %d lignes\n", nrow(df)))
cat(sprintf("      (codées: %d | inférées: %d)\n",
  sum(df$equity_status == "equity_coded"),
  sum(df$equity_status == "equity_inferred")))
cat(sprintf("      Lignes exclues (options/ETF/crypto) : %d\n\n",
  nrow(raw %>% filter(year(Traded) >= 2021, year(Traded) <= 2025,
    Name != "Ro Khanna",
    str_to_upper(Transaction) %in%
      c("PURCHASE", "SALE", "SALE (FULL)", "SALE
(PARTIAL)")) ) - nrow(df)))

# Taille des transactions → point médian numérique
parse_trade_size <- function(s) {
  if (is.na(s)) return(NA_real_)

```

```

s <- str_replace_all(s, "[$,]", "")
if (str_detect(s, "(?i)over")) {
  val <- str_extract(s, "[0-9]+") %>% as.numeric()
  return(val * 1.5)
}
if (str_detect(s, "-")) {
  parts <- str_split(s, "-")[[1]] %>% str_trim() %>% as.numeric()
  if (length(parts) == 2 && !any(is.na(parts))) return(mean(parts))
}
suppressWarnings(as.numeric(str_extract(s, "[0-9.]+")))
}

df <- df %>%
  mutate(trade_size_mid = map_dbl(Trade_Size_USD, parse_trade_size))

# [F4] Winsorisation (valeur brute conservée dans excess_return_raw)
winsorise <- function(x, low = 0.01, high = 0.99) {
  x <- suppressWarnings(as.numeric(x))
  q <- quantile(x, probs = c(low, high), na.rm = TRUE)
  pmax(pmin(x, q[2]), q[1])
}

df <- df %>%
  mutate(
    excess_return          = suppressWarnings(as.numeric(excess_return)),
    excess_return_raw      = excess_return,
    excess_return          = winsorise(excess_return),
    excess_return_30d_raw  = excess_return_30d,
    excess_return_30d      = winsorise(excess_return_30d)
  )

q01 <- quantile(df$excess_return_raw, 0.01, na.rm = TRUE)
q99 <- quantile(df$excess_return_raw, 0.99, na.rm = TRUE)
cat(sprintf("[F4] Bornes winsorisation : [%.2f%%, %.2f%%]\n", q01, q99))
cat(sprintf("    Lignes affectées : %d (%.1f%%)\n\n",
  sum(!is.na(df$excess_return_raw) &
    (df$excess_return_raw < q01 | df$excess_return_raw > q99)),
  100 * mean(!is.na(df$excess_return_raw) &
    (df$excess_return_raw < q01 | df$excess_return_raw >
q99))))

df <- df %>%
  mutate(reporting_lag = as.numeric(Filed - Traded))

df <- df %>%
  mutate(
    year      = year(Traded),
    quarter   = paste0(year, " Q", quarter(Traded))
  )

df <- df %>%
  mutate(
    Party_clean = case_when(
      Party == "Republican" ~ "Republican",
      Party == "Democratic" ~ "Democratic",
      TRUE                  ~ "Other"
    )
  )

```

```

cat("== DATASET FINAL ==\n")
cat(sprintf(" Transactions      : %d\n",      nrow(df)))
cat(sprintf(" Élus uniques      : %d\n",      n_distinct(df$Name)))
cat(sprintf(" Tickers uniques   : %d\n",      n_distinct(df$Ticker)))
cat(sprintf(" Période          : %s à %s\n",
            format(min(df$Traded), "%Y-%m-%d"),
            format(max(df$Traded), "%Y-%m-%d"))))
cat(sprintf(" Achats            : %d\n",      sum(df$Transaction_clean ==
"Purchase"))))
cat(sprintf(" Ventes            : %d\n",      sum(df$Transaction_clean == "Sale"))))
cat(sprintf(" excess_return NA: %d\n",      sum(is.na(df$excess_return))))
cat("==\n\n")

write_csv(df, file.path(DATA_DIR, "congress_clean_2021_2025.csv"))
cat(sprintf("Données nettoyées : %s/congress_clean_2021_2025.csv\n\n", DATA_DIR))

# — 2. Statistiques descriptives —————

cat("==\n")
cat(" SECTION 2 – STATISTIQUES DESCRIPTIVES\n")
cat("==\n\n")

cat("— Résumé global – excess_return —\n")
summary_stats <- df %>%
  filter(!is.na(excess_return)) %>%
  summarise(
    N          = n(),
    Mean       = mean(excess_return),
    Median     = median(excess_return),
    SD         = sd(excess_return),
    Min        = min(excess_return),
    Max        = max(excess_return),
    Pct_positive = mean(excess_return > 0) * 100
  )
print(summary_stats)

cat("\n— Par type de transaction —\n")
summary_by_txn <- df %>%
  filter(!is.na(excess_return), Transaction_clean != "Other") %>%
  group_by(Transaction_clean) %>%
  summarise(
    N          = n(),
    Mean       = round(mean(excess_return), 3),
    Median     = round(median(excess_return), 3),
    SD         = round(sd(excess_return), 3)
  )
print(summary_by_txn)

cat("\n— Par chambre —\n")
summary_by_chamber <- df %>%
  filter(!is.na(excess_return)) %>%
  group_by(Chamber) %>%
  summarise(
    N          = n(),
    Mean       = round(mean(excess_return), 3),
    Median     = round(median(excess_return), 3),

```

```

    SD      = round(sd(excess_return), 3)
  )
print(summary_by_chamber)

cat("\n— Par parti —\n")
summary_by_party <- df %>%
  filter(!is.na(excess_return), Party_clean != "Other") %>%
  group_by(Party_clean) %>%
  summarise(
    N      = n(),
    Mean   = round(mean(excess_return), 3),
    Median = round(median(excess_return), 3),
    SD     = round(sd(excess_return), 3)
  )
print(summary_by_party)

cat("\n— Délai de déclaration (jours) —\n")
lag_desc <- df %>%
  filter(reporting_lag >= 0, reporting_lag <= 1000) %>%
  summarise(
    Mean_lag   = round(mean(reporting_lag), 1),
    Median_lag = median(reporting_lag),
    Max_lag    = max(reporting_lag),
    Pct_late   = round(mean(reporting_lag > 45) * 100, 1) # délai légal STOCK
    Act : 45 j
  )
print(lag_desc)

# — 3. Activité de trading —————

cat("\n===== \n")
cat("  SECTION 3 – ACTIVITÉ DE TRADING\n")
cat("===== \n\n")

trades_per_year <- df %>%
  filter(Transaction_clean != "Other") %>%
  count(year, Transaction_clean) %>%
  mutate(year = factor(year))

cat("— Transactions par année —\n")
print(trades_per_year %>% pivot_wider(names_from = Transaction_clean, values_from = n))

buy_sell_ratio <- df %>%
  filter(Transaction_clean %in% c("Purchase", "Sale")) %>%
  count(year, Transaction_clean) %>%
  pivot_wider(names_from = Transaction_clean, values_from = n) %>%
  mutate(BuySell_ratio = round(Purchase / Sale, 2))
cat("\n— Ratio achats/ventes —\n")
print(buy_sell_ratio)

volume_per_year <- df %>%
  filter(!is.na(trade_size_mid), Transaction_clean != "Other") %>%
  group_by(year, Transaction_clean) %>%
  summarise(Total_USD = sum(trade_size_mid, na.rm = TRUE), .groups = "drop") %>%
  mutate(Total_USD_M = round(Total_USD / 1e6, 1))
cat("\n— Volume estimé ($M) par année —\n")

```

```

print(volume_per_year)

# Groupé par Ticker uniquement – un même titre peut avoir plusieurs variantes de
# nom
cat("\n— Top 20 tickers (2021-2025) —\n")
top_tickers <- df %>%
  filter(Transaction_clean %in% c("Purchase", "Sale")) %>%
  count(Ticker, sort = TRUE) %>%
  head(20)
print(top_tickers)

cat("\n— Top 20 élus les plus actifs —\n")
top_traders_count <- df %>%
  filter(Transaction_clean %in% c("Purchase", "Sale")) %>%
  count(Name, Party_clean, Chamber, sort = TRUE) %>%
  head(20)
print(top_traders_count)

# — 4. Rendements excédentaires —————

cat("\n===== \n")
cat("  SECTION 4 – RENDEMENTS EXCÉDENTAIRES \n")
cat("===== \n\n")

cat("— Top 20 élus par rendement médian (≥10 transactions) —\n")
top_earners <- df %>%
  filter(!is.na(excess_return), Transaction_clean %in% c("Purchase", "Sale")) %>%
  group_by(Name, Party_clean, Chamber) %>%
  summarise(
    N_trades      = n(),
    Median_excess = round(median(excess_return), 2),
    Mean_excess   = round(mean(excess_return), 2),
    Pct_positive  = round(mean(excess_return > 0) * 100, 1),
    .groups = "drop"
  ) %>%
  filter(N_trades >= 10) %>%
  arrange(desc(Median_excess)) %>%
  head(20)
print(top_earners)

cat("\n— Bottom 20 élus par rendement médian (≥10 transactions) —\n")
bottom_earners <- df %>%
  filter(!is.na(excess_return), Transaction_clean %in% c("Purchase", "Sale")) %>%
  group_by(Name, Party_clean, Chamber) %>%
  summarise(
    N_trades      = n(),
    Median_excess = round(median(excess_return), 2),
    Mean_excess   = round(mean(excess_return), 2),
    .groups = "drop"
  ) %>%
  filter(N_trades >= 10) %>%
  arrange(Median_excess) %>%
  head(20)
print(bottom_earners)

cat("\n— Rendement médian par année —\n")
return_by_year <- df %>%

```

```

filter(!is.na(excess_return), Transaction_clean %in% c("Purchase", "Sale")) %>%
group_by(year) %>%
summarise(
  N = n(),
  Mean_excess = round(mean(excess_return), 3),
  Median_excess = round(median(excess_return), 3),
  .groups = "drop"
)
print(return_by_year)

cat("\n— Top 15 actions par rendement médian (≥20 transactions) —\n")
top_stocks_return <- df %>%
  filter(!is.na(excess_return), Transaction_clean %in% c("Purchase", "Sale")) %>%
  group_by(Ticker) %>%
  summarise(
    N_trades = n(),
    Median_excess = round(median(excess_return), 2),
    Mean_excess = round(mean(excess_return), 2),
    .groups = "drop"
  ) %>%
  filter(N_trades >= 20) %>%
  arrange(desc(Median_excess)) %>%
  head(15)
print(top_stocks_return)

# — 5. Visualisations —————

cat("\n===== \n")
cat("  SECTION 5 – VISUALISATIONS\n")
cat("===== \n\n")

p1 <- df %>%
  filter(Transaction_clean %in% c("Purchase", "Sale")) %>%
  count(year, Transaction_clean) %>%
  ggplot(aes(x = factor(year), y = n, fill = Transaction_clean)) +
  geom_col(position = "dodge", width = 0.65) +
  scale_fill_manual(values = c("Purchase" = "#2166ac", "Sale" = "#d6604d"),
    name = "Transaction") +
  scale_y_continuous(labels = comma) +
  labs(
    title = "Activité de trading au Congrès (2021–2025)",
    subtitle = "Nombre de transactions déclarées par année",
    x = "Année", y = "Nombre de transactions"
  )
ggsave(file.path(PLOTS_DIR, "fig_activite_trading_par_annee.png"), p1, width = 8,
  height = 5, dpi = 150)
print(p1)

q_low <- quantile(df$excess_return, 0.01, na.rm = TRUE)
q_high <- quantile(df$excess_return, 0.99, na.rm = TRUE)

p2 <- df %>%
  filter(!is.na(excess_return),
    excess_return >= q_low, excess_return <= q_high,
    Transaction_clean %in% c("Purchase", "Sale")) %>%
  ggplot(aes(x = excess_return, fill = Transaction_clean)) +
  geom_histogram(bins = 60, alpha = 0.7, colour = "white") +

```

```

geom_vline(xintercept = 0, linetype = "dashed", colour = "black", linewidth =
0.8) +
  scale_fill_manual(values = c("Purchase" = "#2166ac", "Sale" = "#d6604d"),
                    name = "Transaction") +
  scale_y_continuous(labels = comma) +
  labs(
    title = "Distribution des rendements excédentaires – transactions du
Congrès",
    subtitle = "1er-99e percentile. Ligne en pointillé = rendement excédentaire
nul.",
    x = "Rendement excédentaire (%)", y = "Nombre de transactions"
  )
ggsave(file.path(PLOTS_DIR, "fig_distribution_rendements_excedentaires.png"), p2,
width = 9, height = 5, dpi = 150)
print(p2)

p3 <- df %>%
  filter(!is.na(excess_return), Transaction_clean %in% c("Purchase", "Sale")) %>%
  group_by(Chamber, Transaction_clean) %>%
  summarise(Median_excess = median(excess_return), .groups = "drop") %>%
  ggplot(aes(x = Chamber, y = Median_excess, fill = Transaction_clean)) +
  geom_col(position = "dodge", width = 0.55) +
  geom_hline(yintercept = 0, linetype = "dashed", colour = "grey40") +
  scale_fill_manual(values = c("Purchase" = "#2166ac", "Sale" = "#d6604d"),
                    name = "Transaction") +
  labs(
    title = "Rendement excédentaire médian par chambre",
    subtitle = "Chambre des représentants vs. Sénat, 2021-2025",
    x = "Chambre", y = "Rendement excédentaire médian (%)"
  )
ggsave(file.path(PLOTS_DIR, "fig_rendement_par_chambre.png"), p3, width = 7, height
= 5, dpi = 150)
print(p3)

p4 <- df %>%
  filter(!is.na(excess_return),
         Party_clean != "Other",
         Transaction_clean %in% c("Purchase", "Sale")) %>%
  group_by(Party_clean, Transaction_clean) %>%
  summarise(Median_excess = median(excess_return), .groups = "drop") %>%
  ggplot(aes(x = Party_clean, y = Median_excess, fill = Transaction_clean)) +
  geom_col(position = "dodge", width = 0.55) +
  geom_hline(yintercept = 0, linetype = "dashed", colour = "grey40") +
  scale_fill_manual(values = c("Purchase" = "#2166ac", "Sale" = "#d6604d"),
                    name = "Transaction") +
  scale_x_discrete(labels = c("Democratic" = "Democrat", "Republican" =
"Republican")) +
  labs(
    title = "Rendement excédentaire médian par parti",
    subtitle = "2021-2025",
    x = "Parti", y = "Rendement excédentaire médian (%)"
  )
ggsave(file.path(PLOTS_DIR, "fig_rendement_par_parti.png"), p4, width = 7, height =
5, dpi = 150)
print(p4)

p5 <- df %>%
  filter(!is.na(excess_return), Transaction_clean %in% c("Purchase", "Sale")) %>%

```

```

group_by(year, Transaction_clean) %>%
summarise(Median_excess = median(excess_return), .groups = "drop") %>%
ggplot(aes(x = year, y = Median_excess, colour = Transaction_clean,
           group = Transaction_clean)) +
geom_line(linewidth = 1.2) +
geom_point(size = 3) +
geom_hline(yintercept = 0, linetype = "dashed", colour = "grey50") +
scale_colour_manual(values = c("Purchase" = "#2166ac", "Sale" = "#d6604d"),
                    name = "Transaction") +
scale_x_continuous(breaks = 2021:2025) +
labs(
  title = "Rendement excédentaire médian par année",
  subtitle = "Transactions du Congrès, 2021-2025",
  x = "Année", y = "Rendement excédentaire médian (%)"
)
ggsave(file.path(PLOTS_DIR, "fig_rendement_par_annee.png"), p5, width = 8, height =
5, dpi = 150)
print(p5)

p6 <- top_earners %>%
mutate(Name_short = str_trunc(Name, 22)) %>%
ggplot(aes(x = reorder(Name_short, Median_excess),
           y = Median_excess, fill = Party_clean)) +
geom_col(width = 0.75) +
coord_flip() +
scale_fill_manual(
  values = c("Republican" = "#d6604d", "Democratic" = "#4393c3", "Other" =
"grey60"),
  name = "Parti"
) +
labs(
  title = "Top 20 élus par rendement excédentaire médian",
  subtitle = "Minimum 10 transactions, 2021-2025",
  x = NULL, y = "Rendement excédentaire médian (%)"
)
ggsave(file.path(PLOTS_DIR, "fig_top20_elus_rendement.png"), p6, width = 9, height
= 7, dpi = 150)
print(p6)

p7 <- top_tickers %>%
ggplot(aes(x = reorder(Ticker, n), y = n)) +
geom_col(fill = "#4d9de0", width = 0.75) +
coord_flip() +
scale_y_continuous(labels = comma) +
labs(
  title = "Les 20 actions les plus échangées par les membres du Congrès",
  subtitle = "2021-2025 (achats et ventes combinés)",
  x = NULL, y = "Nombre de transactions"
)
ggsave(file.path(PLOTS_DIR, "fig_top20_tickers.png"), p7, width = 9, height = 7,
dpi = 150)
print(p7)

p8 <- df %>%
filter(reporting_lag >= 0, reporting_lag <= 200) %>%
ggplot(aes(x = reporting_lag)) +
geom_histogram(bins = 50, fill = "#6a9c89", colour = "white") +

```

```

    geom_vline(xintercept = 45, linetype = "dashed", colour = "#c0392b", linewidth
= 1) +
    annotate("text", x = 48, y = Inf, vjust = 2, hjust = 0,
            label = "Délai légal STOCK Act (45 jours)", colour = "#c0392b", size =
3.5) +
    scale_y_continuous(labels = comma) +
    labs(
      title    = "Distribution du délai de déclaration",
      subtitle = "Jours entre la date de transaction et la date de dépôt (plafonné
à 200 jours)",
      x = "Délai de déclaration (jours)", y = "Nombre de transactions"
    )
ggsave(file.path(PLOTS_DIR, "fig_distribution_delai_declaration.png"), p8, width =
8, height = 5, dpi = 150)
print(p8)

p9 <- volume_per_year %>%
  ggplot(aes(x = factor(year), y = Total_USD_M, fill = Transaction_clean)) +
  geom_col(position = "dodge", width = 0.65) +
  scale_fill_manual(values = c("Purchase" = "#2166ac", "Sale" = "#d6604d"),
                    name = "Transaction") +
  scale_y_continuous(labels = comma) +
  labs(
    title    = "Volume de transactions estimé par année ($M)",
    subtitle = "Basé sur la valeur médiane des fourchettes de taille déclarées",
    x = "Année", y = "Volume total ($M)"
  )
ggsave(file.path(PLOTS_DIR, "fig_volume_par_annee.png"), p9, width = 8, height = 5,
dpi = 150)
print(p9)

# — 7. Tests H1 et analyse du lag —————

cat("\n===== \n")
cat("  SECTION 7 – TESTS H1 & ANALYSE DU LAG \n")
cat("===== \n\n")

run_h1_tests <- function(data, ret_var, label) {
  cat(sprintf("— %s — \n", label))

  pull_ret <- function(txn) {
    if (txn == "All")
      data %>% filter(!is.na(.data[[ret_var]])) %>% pull(ret_var)
    else
      data %>% filter(Transaction_clean == txn, !is.na(.data[[ret_var]])) %>%
pull(ret_var)
  }

  x_all <- pull_ret("All")
  x_buy <- pull_ret("Purchase")
  x_sell <- pull_ret("Sale")

  t_all <- t.test(x_all, mu = 0)
  t_buy <- t.test(x_buy, mu = 0)
  t_sell <- t.test(x_sell, mu = 0)
  t_diff <- t.test(x_buy, x_sell, var.equal = FALSE)

```

```

results <- tibble(
  Group      = c("All", "Purchase (BUY)", "Sale (SELL)", "BUY vs SELL"),
  N          = c(length(x_all), length(x_buy), length(x_sell), NA_integer_),
  Mean       = round(c(mean(x_all), mean(x_buy), mean(x_sell), NA_real_), 3),
  Median     = round(c(median(x_all), median(x_buy), median(x_sell), NA_real_),
3),
  t_stat     = round(c(t_all$statistic, t_buy$statistic,
                      t_sell$statistic, t_diff$statistic), 3),
  p_value    = round(c(t_all$p.value, t_buy$p.value,
                      t_sell$p.value, t_diff$p.value), 4),
  Sig_5pct   = c(t_all$p.value < 0.05, t_buy$p.value < 0.05,
                  t_sell$p.value < 0.05, t_diff$p.value < 0.05)
)
print(results)
cat("\n")
invisible(results)
}

h1_primary <- run_h1_tests(df, "excess_return",
                          "excess_return - QuiverQuant (fenêtre courte)")

if ("excess_return_30d" %in% names(df) && any(!is.na(df$excess_return_30d))) {
  h1_robust <- run_h1_tests(df, "excess_return_30d",
                          "excess_return_30d - Yahoo Finance (30 jours)")
} else {
  cat("[WARN] excess_return_30d absent ou tout NA - robustesse ignorée\n\n")
}

for (rv in c("excess_return_1d", "excess_return_7d")) {
  if (rv %in% names(df) && any(!is.na(df[[rv]]))) {
    lbl <- switch(rv,
      excess_return_1d = "excess_return_1d - Yahoo Finance (+1 jour)",
      excess_return_7d = "excess_return_7d - Yahoo Finance (+7 jours)"
    )
    out <- run_h1_tests(df, rv, lbl)
    assign(paste0("h1_", sub("excess_return_", "", rv)), out)
  } else {
    cat(sprintf("[WARN] %s absent ou tout NA - ignoré\n\n", rv))
  }
}

cat("— Synthèse multi-horizons (Yahoo Finance) —\n")

build_horizon_summary <- function(rv, label) {
  if (!rv %in% names(df) || !any(!is.na(df[[rv]]))) return(NULL)
  x_buy <- df %>% filter(Transaction_clean == "Purchase", !is.na(.data[[rv]]))
  %>% pull(rv)
  x_sell <- df %>% filter(Transaction_clean == "Sale", !is.na(.data[[rv]]))
  %>% pull(rv)
  t_buy <- t.test(x_buy, mu = 0)
  t_sell <- t.test(x_sell, mu = 0)
  tibble(
    Horizon    = label,
    N_buy      = length(x_buy),
    Mean_buy   = round(mean(x_buy), 3),
    p_buy      = round(t_buy$p.value, 4),
    Sig_buy    = t_buy$p.value < 0.05,
    N_sell     = length(x_sell),

```

```

    Mean_sell = round(mean(x_sell), 3),
    p_sell    = round(t_sell$p.value, 4),
    Sig_sell  = t_sell$p.value < 0.05
  )
}

h1_multi <- bind_rows(
  build_horizon_summary("excess_return_1d", "+1 jour"),
  build_horizon_summary("excess_return_7d", "+7 jours"),
  build_horizon_summary("excess_return_30d", "+30 jours")
)
print(h1_multi)
cat("\n")
write_csv(h1_multi, file.path(TABLES_DIR, "table_h1_multi_horizons.csv"))

cat("— Sous-groupes politiques – excess_return_30d (achats) —\n\n")

run_subgroup_test <- function(data, group_var, group_val, label) {
  x <- data %>%
    filter(Transaction_clean == "Purchase",
           !is.na(excess_return_30d),
           .data[[group_var]] == group_val) %>%
    pull(excess_return_30d)
  if (length(x) < 10) return(NULL)
  tt <- t.test(x, mu = 0)
  tibble(
    Group    = label,
    N        = length(x),
    Mean     = round(mean(x), 3),
    Median   = round(median(x), 3),
    t_stat   = round(tt$statistic, 3),
    p_value  = round(tt$p.value, 4),
    Sig_5pct = tt$p.value < 0.05
  )
}

political_tests <- bind_rows(
  run_subgroup_test(df, "Chamber", "House", "House"),
  run_subgroup_test(df, "Chamber", "Senate", "Senate"),
  run_subgroup_test(df, "Party_clean", "Democratic", "Democrats"),
  run_subgroup_test(df, "Party_clean", "Republican", "Republicans")
)

cat("— H0 : rendement moyen = 0 par sous-groupe (achats) —\n")
print(political_tests)
cat("\n")
write_csv(political_tests, file.path(TABLES_DIR,
  "table_h1_sous_groupes_politiques.csv"))

cat("— Tests bilatéraux : Chambre vs. Sénat et Dém. vs. Rép. —\n")
buy_house <- df %>% filter(Transaction_clean=="Purchase", Chamber=="House",
  !is.na(excess_return_30d)) %>% pull(excess_return_30d)
buy_senate <- df %>% filter(Transaction_clean=="Purchase", Chamber=="Senate",
  !is.na(excess_return_30d)) %>% pull(excess_return_30d)
buy_dem <- df %>% filter(Transaction_clean=="Purchase",
  Party_clean=="Democratic", !is.na(excess_return_30d)) %>%
  pull(excess_return_30d)

```

```

buy_rep <- df %>% filter(Transaction_clean=="Purchase",
  Party_clean=="Republican", !is.na(excess_return_30d)) %>%
pull(excess_return_30d)

t_chamber <- t.test(buy_house, buy_senate, var.equal = FALSE)
t_party <- t.test(buy_dem, buy_rep, var.equal = FALSE)

two_sample <- tibble(
  Comparison = c("House vs. Senate", "Democrats vs. Republicans"),
  N_group1 = c(length(buy_house), length(buy_dem)),
  Mean_group1 = round(c(mean(buy_house), mean(buy_dem)), 3),
  N_group2 = c(length(buy_senate), length(buy_rep)),
  Mean_group2 = round(c(mean(buy_senate), mean(buy_rep)), 3),
  t_stat = round(c(t_chamber$statistic, t_party$statistic), 3),
  p_value = round(c(t_chamber$p.value, t_party$p.value), 4),
  Sig_5pct = c(t_chamber$p.value < 0.05, t_party$p.value < 0.05)
)
print(two_sample)
cat("\n")
write_csv(two_sample, file.path(TABLES_DIR,
  "table_h1_test_bilateral_politique.csv"))

cat("— Seuil Buffett : % de transactions avec excess_return > 11% —\n")
buffett_all <- df %>%
  filter(!is.na(excess_return), Transaction_clean %in% c("Purchase", "Sale")) %>%
  summarise(Group = "All", N = n(),
    Pct_above_buffett = round(mean(excess_return > 11) * 100, 1))

buffett_txn <- df %>%
  filter(!is.na(excess_return), Transaction_clean %in% c("Purchase", "Sale")) %>%
  group_by(Group = Transaction_clean) %>%
  summarise(N = n(),
    Pct_above_buffett = round(mean(excess_return > 11) * 100, 1),
    .groups = "drop")

print(bind_rows(buffett_txn, buffett_all))
cat("\n")

cat("— Test du lag : rendement excédentaire moyen par groupe de délai (achats)
—\n\n")

df <- df %>%
  mutate(
    lag_bin = case_when(
      reporting_lag <= 10 ~ "0-10 jours",
      reporting_lag <= 30 ~ "11-30 jours",
      reporting_lag <= 45 ~ "31-45 jours",
      reporting_lag > 45 ~ "45+ jours (hors délai)",
      TRUE ~ NA_character_
    ),
    lag_bin = factor(lag_bin,
      levels = c("0-10 jours", "11-30 jours",
        "31-45 jours", "45+ jours (hors délai)"))
  )

lag_buy <- df %>%
  filter(!is.na(lag_bin), !is.na(excess_return),
    Transaction_clean == "Purchase",

```

```

        reporting_lag >= 0, reporting_lag <= 1000)

lag_summary <- lag_buy %>%
  group_by(lag_bin) %>%
  summarise(
    N = n(),
    Mean_excess = round(mean(excess_return), 3),
    Median_excess = round(median(excess_return), 3),
    SD = round(sd(excess_return), 3),
    t_stat = round(t.test(excess_return, mu = 0)$statistic, 3),
    p_value = round(t.test(excess_return, mu = 0)$p.value, 4),
    .groups = "drop"
  )

print(lag_summary)
cat("\n")

cat("— ANOVA – excess_return ~ lag_bin (achats) —\n")
anova_lag <- aov(excess_return ~ lag_bin, data = lag_buy)
print(summary(anova_lag))
cat("\n")

cat("— Tests par paires (Bonferroni) —\n")
bonf_result <- pairwise.t.test(lag_buy$excess_return, lag_buy$lag_bin,
                              p.adjust.method = "bonferroni")

print(bonf_result)
cat("\n")

bonf_pmat <- as.data.frame(bonf_result$p.value) %>%
  rownames_to_column("Groupe_1") %>%
  pivot_longer(-Groupe_1, names_to = "Groupe_2", values_to = "p_adj") %>%
  filter(!is.na(p_adj)) %>%
  mutate(
    Sig_label = case_when(
      p_adj < 0.001 ~ "****",
      p_adj < 0.01 ~ "***",
      p_adj < 0.05 ~ "**",
      TRUE ~ "n.s."
    ),
    p_label = sprintf("p = %.4f\n(%)", p_adj, Sig_label)
  )

p_bonf <- bonf_pmat %>%
  ggplot(aes(x = Groupe_2, y = Groupe_1, fill = p_adj)) +
  geom_tile(colour = "white", linewidth = 0.8) +
  geom_text(aes(label = p_label), size = 3.4, lineheight = 1.3) +
  scale_fill_gradient2(
    low = "#d73027", mid = "#fee090", high = "#f7f7f7",
    midpoint = 0.05, limits = c(0, 1),
    name = "p-valeur\n(Bonferroni)"
  ) +
  labs(
    title = "Comparaisons par paires – délai de déclaration (correction de
Bonferroni)",
    subtitle = "Rendement excédentaire moyen entre les groupes de délai (achats
uniquement)",
    x = NULL, y = NULL
  ) +

```

```

    theme(axis.text.x = element_text(angle = 15, hjust = 1))
ggsave(file.path(PLOTS_DIR, "fig_anova_bonferroni_delai.png"),
       p_bonf, width = 8, height = 5, dpi = 150)
print(p_bonf)

p_lag <- lag_summary %>%
  ggplot(aes(x = lag_bin, y = Mean_excess)) +
  geom_col(fill = "#4d9de0", width = 0.6) +
  geom_hline(yintercept = 0, linetype = "dashed", colour = "grey40") +
  geom_errorbar(aes(ymin = Mean_excess - SD / sqrt(N),
                    ymax = Mean_excess + SD / sqrt(N)),
                width = 0.2, colour = "grey30") +
  labs(
    title = "ANOVA – Rendement excédentaire moyen par délai de déclaration
(achats)",
    subtitle = "Barres d'erreur =  $\pm 1$  ET.  $H_0$  : rendement moyen = 0.",
    x = "Jours entre la transaction et le dépôt",
    y = "Rendement excédentaire moyen (%)"
  )
ggsave(file.path(PLOTS_DIR, "fig_rendement_par_categorie_delai.png"),
       p_lag, width = 8, height = 5, dpi = 150)
print(p_lag)

write_csv(h1_primary %>% mutate(variable = "excess_return"),
         file.path(TABLES_DIR, "table_h1_ttest_principal.csv"))
write_csv(lag_summary, file.path(TABLES_DIR, "table_h2_anova_delai.csv"))
if (exists("h1_robust")) {
  write_csv(h1_robust %>% mutate(variable = "excess_return_30d"),
            file.path(TABLES_DIR, "table_h1_ttest_robustesse.csv"))
}
cat("Tableaux H1 sauvegardés dans output/tables/\n\n")

# — 8. Simulation de copy trading —————
# Variable : excess_return_30d (Yahoo Finance, fenêtre 30j depuis la
# transaction).
# Pour lag > 30 j, la fenêtre est expirée au moment de la publication.

cat("\n===== \n")
cat(" SECTION 8 – COPY TRADING \n")
cat("===== \n\n")

copy_df <- df %>%
  filter(
    Transaction_clean == "Purchase",
    !is.na(excess_return_30d),
    reporting_lag >= 0,
    reporting_lag <= 365
  ) %>%
  mutate(
    window_remaining_pct = pmax(0, (30 - reporting_lag) / 30),
    window_expired = reporting_lag > 30
  )

cat(sprintf("Base copy trading : %d achats avec excess_return_30d valide\n",
  nrow(copy_df)))
cat(sprintf(" dont fenêtre expirée (lag > 30 j) : %d (%.1f %)\n\n",
  sum(copy_df$window_expired),

```

```

100 * mean(copy_df$window_expired)))

run_copy_strategy <- function(data, max_lag, strategy_name) {
  filtered <- data %>% filter(reporting_lag <= max_lag)
  if (nrow(filtered) < 30) return(NULL)
  tt <- t.test(filtered$excess_return_30d, mu = 0)
  tibble(
    Strategie      = strategy_name,
    Lag_max_j      = max_lag,
    N_trades       = nrow(filtered),
    Lag_moyen_j    = round(mean(filtered$reporting_lag), 1),
    Mean_return_pct = round(mean(filtered$excess_return_30d), 3),
    Median_return_pct = round(median(filtered$excess_return_30d), 3),
    SD             = round(sd(filtered$excess_return_30d), 3),
    t_stat         = round(tt$statistic, 3),
    p_value        = round(tt$p.value, 4),
    Sig_5pct       = tt$p.value < 0.05,
    Pct_positif    = round(mean(filtered$excess_return_30d > 0) * 100, 1),
    IR             = round(mean(filtered$excess_return_30d) /
                          sd(filtered$excess_return_30d), 4)
  )
}

strategies <- bind_rows(
  run_copy_strategy(copy_df, 10, "Ultra-rapide (lag <= 10 j)"),
  run_copy_strategy(copy_df, 20, "Rapide (lag <= 20 j)"),
  run_copy_strategy(copy_df, 30, "Standard (lag <= 30 j)"),
  run_copy_strategy(copy_df, 45, "Legale (lag <= 45 j)"),
  run_copy_strategy(copy_df, 365, "Aveugle (tous achats)")
)

cat("— Performance par stratégie de copy trading —\n")
print(strategies)
cat("\n")
write_csv(strategies, file.path(TABLES_DIR,
  "table_h2_copy_trading_strategies.csv"))

yearly_copy <- copy_df %>%
  filter(reporting_lag <= 10) %>%
  group_by(year) %>%
  summarise(
    N          = n(),
    Mean_return = round(mean(excess_return_30d), 3),
    Median_return = round(median(excess_return_30d), 3),
    Pct_positif = round(mean(excess_return_30d > 0) * 100, 1),
    t_stat     = round(t.test(excess_return_30d, mu = 0)$statistic, 3),
    p_value    = round(t.test(excess_return_30d, mu = 0)$p.value, 4),
    .groups    = "drop"
  )

cat("— Ultra-rapide (lag <= 10 j) – par année —\n")
print(yearly_copy)
cat("\n")
write_csv(yearly_copy, file.path(TABLES_DIR,
  "table_h2_copy_trading_ultrarapide_annuel.csv"))

copy_window <- copy_df %>%
  group_by(window_expired) %>%

```

```

summarise(
  N = n(),
  Mean_return = round(mean(excess_return_30d), 3),
  Median_return = round(median(excess_return_30d), 3),
  Pct_positif = round(mean(excess_return_30d > 0) * 100, 1),
  t_stat = round(t.test(excess_return_30d, mu = 0)$statistic, 3),
  p_value = round(t.test(excess_return_30d, mu = 0)$p.value, 4),
  .groups = "drop"
) %>%
mutate(Groupe = ifelse(window_expired,
  "Fenetre expiree (lag > 30 j)",
  "Fenetre active (lag <= 30 j)")) %>%
select(Groupe, N, Mean_return, Median_return, Pct_positif, t_stat, p_value)

cat("— Fenêtre 30j active vs expirée au moment de la publication —\n")
print(copy_window)
cat("\n")
write_csv(copy_window, file.path(TABLES_DIR, "table_h2_copy_trading_fenetre.csv"))

p_copy <- strategies %>%
  mutate(
    Strategie = factor(Strategie, levels = rev(Strategie)),
    Positif = Mean_return_pct > 0
  ) %>%
  ggplot(aes(x = Strategie, y = Mean_return_pct, fill = Positif)) +
  geom_col(width = 0.65) +
  geom_hline(yintercept = 0, linetype = "dashed", colour = "grey40") +
  geom_text(aes(label = sprintf("p=%.3f", p_value),
    hjust = ifelse(Mean_return_pct >= 0, -0.1, 1.1)),
    size = 3.2, colour = "grey30") +
  coord_flip() +
  scale_fill_manual(values = c("TRUE" = "#2166ac", "FALSE" = "#d6604d"),
    guide = "none") +
  labs(
    title = "Rendement moyen (30j) par strategie de copy trading",
    subtitle = "Variable : excess_return_30d (Yahoo Finance). H0 : rendement =
0.",
    x = NULL,
    y = "Rendement excedentaire moyen (%)"
  )
ggsave(file.path(PLOTS_DIR, "fig_copy_trading_strategies.png"),
  p_copy, width = 9, height = 5, dpi = 150)
print(p_copy)
cat("Tableaux copy trading sauvegardés dans output/tables/\n\n")

# — 9. Tableaux de synthèse —————

cat("\n===== \n")
cat(" SECTION 9 – SYNTHÈSE \n")
cat("===== \n\n")

h1_rows <- list(h1_primary %>% mutate(Variable = "excess_return (QuiverQuant,
  fenetre courte)"))
if (exists("h1_robust")) {
  h1_rows[[2]] <- h1_robust %>% mutate(Variable = "excess_return_30d (Yahoo,
  30j)")
}

```

```

h1_synth <- bind_rows(h1_rows) %>%
  filter(Group %in% c("All", "Purchase (BUY)", "Sale (SELL)", "BUY vs SELL")) %>%
  mutate(
    Interpretation = case_when(
      Group == "Purchase (BUY)" & Sig_5pct & Mean > 0 ~ "Surperformance
significative",
      Group == "Purchase (BUY)" & Sig_5pct & Mean < 0 ~ "Sous-performance
significative",
      Group == "Purchase (BUY)" & !Sig_5pct ~ "Pas de rendement
anormal (H0 non rejetee)",
      Group == "Sale (SELL)" & Sig_5pct & Mean < 0 ~ "Ventes suivies de sous-
performance",
      Group == "BUY vs SELL" & Sig_5pct ~ "Difference achat/vente
significative",
      Group == "BUY vs SELL" & !Sig_5pct ~ "Pas de difference
achat/vente",
      TRUE ~ ""
    )
  ) %>%
  select(Variable, Group, N, Mean, Median, t_stat, p_value, Sig_5pct,
Interpretation)

cat("— Synthèse H1 —\n")
print(h1_synth)
cat("\n")
write_csv(h1_synth, file.path(TABLES_DIR, "table_synthese_h1.csv"))

h2_synth <- lag_summary %>%
  mutate(
    Interpretation = case_when(
      lag_bin == "0-10 jours" ~ "Rendement positif non sig. -
avantage informationnel residuel possible",
      lag_bin == "11-30 jours" ~ "Rendement negatif sig. - avantage
efface apres 10 jours",
      lag_bin == "31-45 jours" ~ "Rendement negatif sig. - fenetre 30j
presque expiree",
      lag_bin == "45+ jours (hors delai)" ~ "Rendement negatif sig. - hors delai
legal STOCK Act"
    ),
    Lien_litterature = case_when(
      lag_bin == "0-10 jours" ~ "Consistent avec Ziobrowski (2004) : avantage
informationnel residuel",
      TRUE ~ "Consistent avec STOCK Act : contrainte de
publication atteneue l'avantage"
    )
  )

cat("— Synthèse H2 / test lag —\n")
print(h2_synth)
cat("\n")
write_csv(h2_synth, file.path(TABLES_DIR, "table_synthese_h2.csv"))

cat("— Rappel ANOVA (H2) —\n")
cat("F = 8.95, df = 3, p = 6.45e-06 ***\n")
cat("Bonferroni (0-10j vs autres) : p <= 0.007 dans les 3 comparaisons\n\n")
cat("Tous les tableaux de synthèse sauvegardés dans output/tables/\n\n")

```

```
# — Résumé final —————

cat("\n===== \n")
cat("  SCRIPT TERMINÉ – PARTIE 0\n")
cat("===== \n")
cat(sprintf("Données nettoyées : %s/congress_clean_2021_2025.csv\n", DATA_DIR))
cat(sprintf("Graphiques      : %s/\n", PLOTS_DIR))
cat("Tableaux H1      : output/tables/h1_ttest_primary.csv\n")
cat("Robustesse H1     : output/tables/h1_ttest_robustness.csv\n")
cat("Test lag         : output/tables/h1_lag_test.csv\n")
cat("Copy trading      : output/tables/copy_trading_strategies.csv\n")
cat("Copy par année    : output/tables/copy_trading_ultrarapide_par_annee.csv\n")
cat("Copy fenêtre      : output/tables/copy_trading_fenetre.csv\n")
cat("Synthèse H1       : output/tables/synthese_h1.csv\n")
cat("Synthèse H2 lag   : output/tables/synthese_h2_lag.csv\n")
cat("Multi-horizon     : output/tables/h1_multi_horizon.csv\n")
cat("Sous-groupes pol. : output/tables/h1_political_subgroups.csv\n")
cat("Tests bilatéraux  : output/tables/h1_two_sample_political.csv\n\n")

cat("Chiffres clés :\n")
cat("  Transactions (2021-2025)      :", nrow(df), "\n")
cat("  Élus uniques                  :", n_distinct(df$Name), "\n")
cat("  Tickers uniques              :", n_distinct(df$Ticker), "\n")
cat("  Rendement excédentaire moyen  :", round(mean(df$excess_return, na.rm =
  TRUE), 3), "%\n")
cat("  Rendement excédentaire médian :", round(median(df$excess_return, na.rm =
  TRUE), 3), "%\n")
cat("  % transactions positives      :", round(mean(df$excess_return > 0, na.rm =
  TRUE) * 100, 1), "%\n")
```