Essays in Information in Financial Markets

Yifeng Guo

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy under the Executive Committee of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2020
Abstract

Essays in Information in Financial Markets

Yifeng Guo

This dissertation studies topics in the areas of information in financial markets. In the first chapter, *Should Information be Sold Separately? Evidence from MiFID II*, we examine whether selling information separately improves its production. We use a recent regulation in Europe (MiFID II) that unbundles research from transactions to investigate this question. We show that unbundling causes fewer research analysts to cover a firm. This decrease does not come from small- or mid-cap firms but is concentrated in large firms. Contrary to conventional wisdom, the reduction in the coverage quantity is accompanied by an increase in the coverage quality. Further analyses suggest that the enhancement of analyst competition could drive the results: inaccurate analysts drop out (extensive margin) and analysts who stay produce better-quality research (intensive margin). Our findings suggest that selling information separately improves information quality at the cost of reducing information quantity.

The second chapter, *Going Public or Staying Private: The Cost of Mandated Transparency*, focuses on how transparency requirements in public markets affect firms’ decisions to go public or stay private. Public markets are transparent institutions, where disclosure is mandatory, and order flows observable. We show that transparency can lead to insufficient information acquisition and inefficient investment. Transparency of order flows in public markets discourages information acquisition. Insufficient information acquisition then exacerbates the cost of imperfect disclosure.
When the short-term disclosable signal diverges from the long-run value of a project, entrepreneurs prefer opaque private markets where investors can bargain over the costs of acquiring information. Our model links a firm’s preference for public markets to the quality of disclosure metrics. Imperfect communication between investors and entrepreneurs caused by market transparency is a mechanism by which mandatory disclosure may destroy value, leading firms to remain private.

In the third chapter, *Active and Passive Funds: An Equilibrium Analysis*, we provide a benchmark model to analyze investors’ equilibrium choices and the welfare consequences of active and passive investing. Active investing is costly, but it brings two benefits: investors can better hedge by freely trading each asset in the portfolio and can acquire information about the possible state of the world. Information acquisition decisions are strategic substitutes. Investors will become active until the net value of being active shrinks to zero. We show that when the cost of acquiring information is low, equilibrium features the coexistence of informed active investors and passive investors. When the cost of acquiring information rises, informed active, uninformed active and passive investors could coexist. Finally, if the cost of being uninformed active is sufficiently low, passive investing is dominated by active investing. The benchmark model allows future research to explore whether the market equilibrium induces the optimal level of information acquisition and active investment.
# Table of Contents

List of Tables ......................................................... v

List of Figures .......................................................... vii

Acknowledgments ......................................................... ix

Chapter 1: Should Information be Sold Separately? Evidence from MiFID II .................. 1

1.1 Introduction ......................................................... 1

1.2 Regulation background ........................................... 7

1.3 Hypotheses development ......................................... 10

1.4 Empirical design and data ........................................ 13

1.4.1 Empirical design .............................................. 13

1.4.2 Data .............................................................. 14

1.5 Investigating the impact of unbundling: firm level analyses ......................... 17

1.5.1 The impact of unbundling on analyst coverage .................. 19

1.5.2 The impact of unbundling on analysts forecast quality ............ 20

1.6 Investigating the channel: analyst level analyses .......................... 22

1.6.1 Intensive margin ............................................... 22

1.6.2 Extensive margin .............................................. 25

1.7 Robustness check and extensions ................................... 30
1.7.1 Firm’s brokerage house coverage ............................................. 30
1.7.2 Other quality measures ......................................................... 31
1.7.3 Alternatives ........................................................................... 35
1.7.4 Brokerage house level employment ......................................... 37
1.7.5 Capital market effects ............................................................ 39
1.8 Conclusion ................................................................................. 41
1.9 Figures ...................................................................................... 42
1.10 Tables ...................................................................................... 47
1.11 Data construction ................................................................. 63
  1.11.1 Firm level observations ....................................................... 63
  1.11.2 Choice of data frequency .................................................... 64
  1.11.3 Variable definitions ............................................................ 65
1.12 Additional figures ............................................................... 67
1.13 Other summary statistics ...................................................... 71
1.14 Additional regression results .................................................. 76
  1.14.1 Dynamic coefficients point estimates .................................. 76
  1.14.2 Additional results for firm level coverage ......................... 78
  1.14.3 Additional results for analyst informativeness .................... 81
1.15 Propensity score matching .................................................... 82

Chapter 2: Going Public or Staying Private? The Cost of Mandated Transparency ...... 87
  2.1 Introduction ............................................................................. 87
  2.2 Baseline Model ....................................................................... 91
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.2.1</td>
<td>Agents</td>
<td>92</td>
</tr>
<tr>
<td>2.2.2</td>
<td>Markets</td>
<td>94</td>
</tr>
<tr>
<td>2.2.3</td>
<td>Definition of Equilibrium</td>
<td>96</td>
</tr>
<tr>
<td>2.3</td>
<td>Equilibrium Analysis</td>
<td>97</td>
</tr>
<tr>
<td>2.3.1</td>
<td>Public Market</td>
<td>97</td>
</tr>
<tr>
<td>2.3.2</td>
<td>Private Market</td>
<td>101</td>
</tr>
<tr>
<td>2.3.3</td>
<td>The Choice of Public vs. Private Markets</td>
<td>102</td>
</tr>
<tr>
<td>2.4</td>
<td>Extensions and Robustness Check</td>
<td>103</td>
</tr>
<tr>
<td>2.4.1</td>
<td>Short Supply of Informed Investors</td>
<td>103</td>
</tr>
<tr>
<td>2.4.2</td>
<td>Strategic Selling When a Liquidity Shock is Not Observable</td>
<td>105</td>
</tr>
<tr>
<td>2.4.3</td>
<td>Outside Financing</td>
<td>106</td>
</tr>
<tr>
<td>2.5</td>
<td>Mandatory Disclosure and Real Inefficiency</td>
<td>108</td>
</tr>
<tr>
<td>2.6</td>
<td>Conclusion</td>
<td>111</td>
</tr>
<tr>
<td>2.7</td>
<td>Figures</td>
<td>113</td>
</tr>
<tr>
<td>2.8</td>
<td>Proofs</td>
<td>118</td>
</tr>
</tbody>
</table>

Chapter 3: Active and Passive Funds: An Equilibrium Analysis | 122

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Introduction</td>
<td>122</td>
</tr>
<tr>
<td>3.2</td>
<td>Model</td>
<td>125</td>
</tr>
<tr>
<td>3.3</td>
<td>Equilibrium demand and prices at date 2</td>
<td>129</td>
</tr>
<tr>
<td>3.4</td>
<td>Equilibrium choice at date 1</td>
<td>134</td>
</tr>
<tr>
<td>3.4.1</td>
<td>Expected utility of informed, uninformed and passive investors</td>
<td>134</td>
</tr>
<tr>
<td>3.4.2</td>
<td>Equilibrium shares of informed, uniformed and passive investors</td>
<td>141</td>
</tr>
</tbody>
</table>
3.5 Conclusion ................................................................. 144
3.6 Figures ................................................................. 145
3.7 Proof ................................................................. 145

References ............................................................. 184
# List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Firm Level Summary Statistics (Sample for Coverage Quantity)</td>
<td>47</td>
</tr>
<tr>
<td>1.2</td>
<td>Firm Level Summary Statistics (Sample for Coverage Quality)</td>
<td>48</td>
</tr>
<tr>
<td>1.3</td>
<td>Analyst Level Summary Statistics (Sample for Intensive Margin)</td>
<td>49</td>
</tr>
<tr>
<td>1.4</td>
<td>Analyst Level Summary Statistics (Sample for Extensive Margin)</td>
<td>50</td>
</tr>
<tr>
<td>1.5</td>
<td>One Difference (Pre and Post)</td>
<td>51</td>
</tr>
<tr>
<td>1.6</td>
<td>Firm Level Outcomes</td>
<td>52</td>
</tr>
<tr>
<td>1.7</td>
<td>Analyst Level Outcomes (Intensive Margin)</td>
<td>53</td>
</tr>
<tr>
<td>1.8</td>
<td>Analyst Level Outcomes (Extensive Margin). Probability of Stop or Dropping Out after Unbundling</td>
<td>54</td>
</tr>
<tr>
<td>1.9</td>
<td>Analyst Level Outcomes (Extensive Margin). Probability of Stop or Dropping Out Before and After Unbundling</td>
<td>55</td>
</tr>
<tr>
<td>1.10</td>
<td>Placebo Test</td>
<td>56</td>
</tr>
<tr>
<td>1.11</td>
<td>Firms’ Brokerage House Coverage</td>
<td>57</td>
</tr>
<tr>
<td>1.12</td>
<td>Other Measures of Quality</td>
<td>58</td>
</tr>
<tr>
<td>1.13</td>
<td>Abnormal Return and Analyst Informativeness</td>
<td>59</td>
</tr>
<tr>
<td>1.14</td>
<td>Learning Effect (Lagged Forecast Error as Control)</td>
<td>60</td>
</tr>
<tr>
<td>1.15</td>
<td>Brokerage House Employment</td>
<td>61</td>
</tr>
<tr>
<td>1.16</td>
<td>Capital Market Effects</td>
<td>62</td>
</tr>
</tbody>
</table>
1.17 Summary Statistics (Country Information) ........................................... 71
1.18 Summary Statistics (Small vs. Large in Sample for Quantity) .............. 72
1.19 Summary Statistics (Small vs. Large in Sample for Quality) ............... 74
1.20 Firm Level Outcomes (Dynamic Coefficients) ...................................... 76
1.21 Intensive Margin (Dynamic Coefficients) ............................................. 77
1.22 Firm Level Coverage (Sample for Coverage Quality) ......................... 78
1.23 Firm Level Coverage (Results in logs) ............................................... 79
1.24 Firm Level Coverage (EU Small vs. US Small and EU Large vs. US Large) 80
1.25 Analyst Informativeness (Small vs. Large) ......................................... 81
1.26 Covariates Balance Table ................................................................. 85
1.27 Firm Level Outcomes (Propensity Score Matching) ............................ 86
List of Figures

1.1 Illustration of Bundling and Unbundling ........................................... 42
1.2 Average Analyst Coverage Over Time (EU vs. US) ................................. 43
1.3 Average Analyst Coverage Over Time between Small Firms and Large Firms in the EU ................................................................. 43
1.4 Empirical Cumulative Distribution of Forecast Error at the Firm Level ........ 44
1.5 Firm Level Outcomes (Dynamic Coefficients) ........................................ 45
1.6 Analyst Level Forecast Error (Dynamic Coefficients) ............................... 46
1.7 Geographic Composition of Stocks in Analyst Portfolio ............................ 67
1.8 Average Analyst Coverage (EU vs. US) ............................................... 68
1.9 Firm Level Forecast Error Distribution ................................................. 69
1.10 Average Number of Analysts per Brokerage House Over Time in the EU, Small vs. Large ............................................................. 70
1.11 Distributions of Propensity Scores after Matching .................................. 84

2.1 Timing of the Model .................................................................................. 96
2.2 Production Technology Choice ................................................................. 113
2.3 Expected Payoff in Public Market ............................................................ 114
2.4 Expected Payoff in Private Market ............................................................ 115
2.5 Equilibrium When $\omega = 0.3 > \frac{L}{L+V_0} = 0.2$ ................................. 116
2.6 Equilibrium with Different Levels of $\omega$ ........................................ 117

3.1 Timeline of the Model ................................................................. 128

3.2 Equilibrium Proportion of Different Investors .............................. 145
Acknowledgements

I am grateful to my advisers, Olivier Darmouni, Harrison Hong, and Laura Veldkamp. They help me to transform my raw ideas into well-crafted research. Laura always stays positive and excited about my progress, cheering me up with positive energy when I feel upset. She helps me to see the big pictures and guides me through when I face tough decisions. She is always there when I need help. Olivier never misses a single line of documents I send to him. He gives detailed plans for my projects. At some point, he probably becomes more familiar with my papers’ introductions than I do. Harrison has critical eyes. He always points out things that I do not but should pay attention to.

I want to thank my coauthors, Milo Bianchi, Joshua Mitts, Lira Mota, and Sophie Moinas. Cooperating with you helps me to learn and grow. I want to thank Lira for her generous support, both in research and in life. My Ph.D. life would have been different (worse) without her. Thank Sophie and Milo for continuing to advise me after I left Toulouse School of Economics.

I am also thankful to other faculty members in Columbia, Simona Abis, Patrick Bolton, Wei Jiang, Giorgia Piacentino, Suresh Sundaresan, Neng Wang, and Kairong Xiao, for their help and advice. Thank you for having your doors open to me all the time.

To my friends, in particular, Yi Cheng, Yu Ding, Tong Geng, Sharon Huang, Yang Jiao, Tie Li, Danqing Mei, Daheng Yang, and Cheng Zheng: I cannot make through the ups and downs without your companion. I am lucky to have you all by my side.

I am indebted to my families, especially my parents. They have seen the best and the worst of
me, but they always feel proud of me and support me in all the ways they can. To my father, Qiang Guo, for teaching me to be responsible, organized, and always plan ahead. To my mother, Jie Fan, for sharing her enthusiasm for her career and showing me that hard work always pays off.

Every journey has an end, but every end is followed by a new beginning.
To my families.
Chapter 1

Should Information be Sold Separately? Evidence from MiFID II

1.1 Introduction

A key to efficient outcomes in financial markets is efficient information production. The production of information, both in terms of quantity and quality, depends on how it is compensated. We study how changes in the compensation schedules for information affect its production through the lens of sell-side research.

The compensation schedule for sell-side research has recently changed dramatically. Sell-side research has traditionally been *bundled* with transactions: rather than being sold separately, it was cross-subsidized by trading commissions. However, a new regulation in Europe (MiFID II, the second Markets in Financial Instrument Directive) *unbundles* research from transactions. Implemented on January 3rd, 2018, MiFID II forces asset managers to separate payments for sell-side research from trading commissions. This change allows us to study how the compensation schedules for sell-side research affect its production, a question difficult for previous empirical research to address due to the lack of economically significant variation in the business model of the sell-side research industry.

Specifically, we ask: What happens to firms’ analyst research quantity and quality? What are

---

1This chapter is based on Guo and Mota (2019). We thank Charles Angelucci, Simona Abis, Olivier Darmouni, Xavier Giroud, Larry Glosten, Juan-Pedro Gómez (the discussant), Harrison Hong, Shiyang Huang (the discussant), Wei Jiang, Luca Mertens, Sophie Moinas, Tano Santos, Paul Tetlock, Tuomas Tomunen, Laura Veldkamp, Kairong Xiao, conference participants at EFA 2019, SFS Cavalcade Asia-Pacific 2019 and seminar participants at AQR Capital Management, City University of Hong Kong, Columbia Business School, Shanghai Advanced Institute of Finance, and Singapore Management University for their insightful and helpful suggestions. We acknowledge support by the Chazen Institute for Global Business at Columbia Business School and Deming Center Doctoral Fellowship.
the underlying mechanisms driving the changes? Answers to these questions are crucial for determining the optimal regulation of financial analysts, and more broadly, for designing the optimal way to pay for information.

The main goal of MiFID II is to protect end investors by improving market transparency. Before MiFID II, trading commissions were opaque and it was difficult for end clients to fully assess the services they were paying for. This generates incentives to assets manager to over charge clients for transaction fees, potentially under-reporting their own management fees (Di Maggio et al. (2019)). Therefore, one provision of MiFID II is to require asset managers to bill research services clearly, which requires unbundling the payments for research from trading.\(^2\)

As a result, unbundling fundamentally changes how research is compensated and, \textit{a priori}, its consequences are unclear. On the one hand, unbundling potentially results in less analyst research covering fewer companies.\(^3\) This is due to the nonrival nature of information. In general, producing research is costly, but once it is produced and revealed to one party, it is difficult to exclude others from using the same information. Therefore, directly selling research could result in market failures (see e.g. Grossman and Stiglitz (1980), Admati and Pfleiderer (1986) and Romer (1990)). Traditionally, brokerage houses overcome this issue by cross-subsidizing the high fixed cost of research with transactions fee and marketing to investor that research was given for “free”. Unbundling prevents cross-subsidizing, potentially causes a marked decrease in research production. Furthermore, a reduction in the coverage quantity could then cause a decrease in the analyst peer pressure, which impairs coverage quality (e.g., Hong and Kacperczyk (2010)).

On the other hand, unbundling could improve analyst incentives to produce better research. Previous studies have shown that when research is bundled with transactions, research analysts may strategically devote more effort to researching firms that generate more trading commissions (Harford et al. (2018)). Analysts may even have incentives to produce inaccurate and biased forecasts to lure trading business in house (e.g., Hong and Kubik (2003), Fang and Yasuda (2009), Hong and Kacperczyk (2010) and Karmaziene (2019)). By making research a standalone product,

\(^2\)We discuss in detail why EU regulators enforce MiFID II in the next section.

\(^3\)See “Mifid II leads to exodus of sellside analysts” Financial Times, June 23rd, 2018.
unbundling could restore analyst incentives and improve research quality.

We find that the number of unique sell-side analysts covering a firm decreases after unbundling. We perform our analyses in a difference-in-difference setting, exploring the different exposures to unbundling between EU public firms and US public firms. We define EU public firms as the treatment group and US public firms as the control group. In our main specification, analyst coverage of EU firms changes by $-0.651$ analysts, which translates into a decrease of $7.45\%$ relative to the average coverage of these firms prior to the regulation. Contrary to the media and industry concerns, this average drop does not come from small- or mid-cap firms but is concentrated in large firms.$^4$ Small firms’ coverage remains almost unchanged and large firms’ coverage drops by $10.53\%$ on average.

Strikingly, we find that the unbundling causes coverage quality to improve. For example, measured by forecast error, coverage quality of affected firms on average increases by $19.19\%$ relative to the average coverage quality of these firms prior to the regulation. This result differs from previous literature (e.g., Hong and Kacperczyk (2010) and Merkley et al. (2017)), which shows a decrease in the coverage quantity implies a reduction in the coverage quality.$^5$ One plausible explanation for the difference is that unbundling strengthens analyst competition. Analysts who sell research as a stand-alone product are likely to compete more directly in the quality domain. Inferior analysts being competed out of business could account for the simultaneous decrease in the coverage quantity and increase in the coverage quality.

To be more specific, at least two economic forces foster competition. First, unbundling forbids asset managers to pass research costs to end investors through trading commissions. Most asset managers have decided to charge research costs against their own profit and loss. Internalizing research cost makes asset managers be much more selective and opt for better research. Second, unbundling puts an explicit price on research. The evaluation scheme prior to the regulation

---


$^5$In their setting (i.e., under bundling), many analysts have other objectives (e.g., generating trading commissions) in addition to quality. The quality of research relies on a few independent and impartial analysts who put quality in the first place. The decrease in the quantity implies both a reduction in the peer pressure and a high propensity to lose these impartial analysts. As a result, research quality worsens.
presents a problem similar to moral hazard in teams. An individual’s evaluation is not directly tied to an individual’s output, but to a bundled output consisting of research and transactions. However, under unbundling, analysts are evaluated directly by the services they provide. They have stronger incentives to provide better research. As a result of those two forces, analysts compete in the quality domain to attract and maintain clients.

We find evidence of two specific predictions of the competition channel: 1) analysts who produce worse research are more likely to be forced out of the analyst market (extensive margin) and 2) analysts who stay produce better research (intensive margin). For example, using ranks generated by analyst average forecast error (Hong and Kubik (2003)), we show that analysts who provide worse forecasts (low-ranked analysts) are more likely to stop working. Conditional on staying, affected analysts’ average forecast error also decreases. Although in a different setting, the beneficial effects of unbundling in our paper could be related to Edelen et al. (2012). They study the effects of unbundling distribution fees from commissions and find that unbundling imposes transparency, mitigates agency conflicts and improves fund performance.

As extensions, we show that individual analyst forecast revision generates a more substantial absolute market-adjusted abnormal return. This evidence suggests that individual forecasts become more informative, which echos our findings on the coverage quality improvement. Interestingly, when we turn to the sum of abnormal returns over all the individual forecast revisions, we find that aggregate abnormal returns decrease after unbundling. This result suggests that the total amount of new information in all the analyst forecast revisions decreases. We interpret these results via the trade-off between forecast quantity and quality. While individual forecasts become more informative, the total amount of valuable forecasts decreases.

We also find that the effects of unbundling may not be homogeneous across brokerage houses. We document that large brokerage houses are likely to be more affected and lay off more analysts. To the extent that brokerage houses affiliated with large banks are larger in size, the result implies that these types of houses, whose research featured substantial bundling before the regulation, are more likely to be affected. This evidence suggests that certain types of brokerage houses could find
it more difficult to survive under unbundling and that the market structure of the sell-side industry may begin to change.

On the aggregate level, we find no evidence that firms’ earnings announcements convey more information, nor do we find evidence on the widening of firms’ bid-ask spread. Although these results do not contradict our findings on coverage quality improving, the aggregate market effect of MiFID II does not seem to manifest itself immediately. Since the reform is relatively recent, we expect that more time is needed to assess its overall impact on the capital markets. As a final note, due to data availability, systematic analyses on the welfare impact of unbundling falls out of the scope of our paper and is left for future studies.\(^6\)

Our paper makes three contributions. To begin with, we are the first to investigate both the impact and the underlying mechanism of unbundling on analyst research production. The output of this project is crucial to resolve the vivid debate among academics, policymakers and practitioners about the consequences of unbundling imposed by MiFID II. Furthermore, investigating unbundling provides valuable information to policy makers and helps to shape future policy design in other jurisdictions in addition to Europe. As US Security Exchange Commission chairman Jay Clayton pointed out:\(^7\)

“It is important to have data and other information about how MiFID II’s research provisions are affecting broker-dealers, investors and small, medium, and large issuers, including whether research availability has been adversely affected.”

Second, our paper contributes to the literature on analyst behavior. It has been long established that analyst forecasts are on average optimistically biased due to incentive problems (e.g., Dreman and Berry (1995), McNichols and O’Brien (1997), and Chopra (1998) and Hong and Kubik (2003)). Fang and Yasuda (2009), Hong and Kacperczyk (2010) and Kempf (2019) convincingly show that implicit mechanisms such as career concerns and peer competition discipline analyst beh-

---

\(^6\) As an example, unbundling may also change the pricing of transactions, which plays an important role in analyzing investors’ welfare. However, systematic information on commission fees is hard to obtain, especially in Europe. Though some surveys suggest commission fees are dropping, a comprehensive analysis of commissions and investor’s welfare falls out of the scope of our paper.

behavior. Our analyses depart from, but also complement, these prior studies by showing that direct market competition helps to discipline analysts even further. A unique finding in our paper is that a decrease in the coverage quantity can be accompanied by an increase in the coverage quality due to the strengthening of analyst competition.

Third, unbundling offers a unique setting through which we can study the optimal way to pay for information (e.g., Veldkamp (2006b), [2006a], Van Nieuwerburgh and Veldkamp (2010a)). Information generally has positive externalities, and its production features high fixed costs. Bundling, though creating incentive problems, covers the high fixed cost and allows for more information production. It potentially enables easier and wider access to information. Unbundling, on the other hand, restores the incentive problems but may result in the market unraveling and the under-provision of certain information. Our results suggest that the improvement in the information quality by selling information separately could come at a cost of reducing information quantity.

In the accounting literature, more recent works study MiFID II, mainly focusing on its effects at the analyst level or the firm level (Fang et al. (2019), Lang et al. (2019)). What differs in our work is that we construct a sample based on analyst-firm pairs, which allows us to identify the mechanism for the drop in coverage quantity and the improvement in coverage quality. Our unique contribution is to show that after the regulation, analyst competition is enhanced. We present clear evidence that inaccurate analysts drop out, and that the ones who stay produce better research. Di Maggio et al. (2019) also studies unbundling using detailed transaction-level data in the US. They emphasize the conflicts of interests between asset managers and end clients. They show that bundling allows asset managers to under-report management fees. Our paper has a different focus. We study how unbundling affects sell-side research provisions. Via investigating unbundling, we aim to understand the effects of changing compensation schedule on information provision.

The remaining of the paper proceeds as follows. In Section 1.2, we provide regulation background. In Section 1.3, we develop our testable hypotheses. In Section 1.4, we describe our

---

8Another difference from Lang et al. (2019) is that we use annual data instead of quarterly data. Since we emphasize causality and mechanism, annual data allows us to mitigate concerns specific to the use of quarterly data that may confound our identification. These include: forecast seasonality, poor data coverage, selection bias, and staleness. Detailed explanation can be found in Appendix 1.11.2.
empirical design and the data we use. We present the firm level analyses in Section 1.5, analyst analyses in Section 1.6, conduct complementary analyses Section 1.7 and conclude in Section 1.8.

1.2 Regulation background

MiFID II attempts to ensure a high degree of harmonized protection for investors in financial instruments by improving market transparency and competitiveness. It is considered as one of the most influential regulatory changes in the EU in response to 2008 financial crisis.

The unbundling provision in MiFID II prohibits asset managers from accepting “fees, commissions or any monetary or non-monetary benefits paid or provided by a third party”.\(^9\) However, it provides an exception for research services provided by third parties as long as the asset management firm pays for it directly. Asset managers can either pay for research directly (against the firm’s profit and loss) or pass costs onto clients by setting up a research payment account (“RPA”) under the consent of clients.\(^10\) Whichever method asset managers choose, payments for research needs to be separated, or unbundled, from payments for transactions (Figure 1.1b). According to a survey by CFA Institute, the majority of asset managers pay for research directly, mainly due to competition pressure from the whole asset management industry.\(^11\) With management fees declining and asset managers absorbing research costs internally, CFA institute expects a greater focus on profitability and efficiency concerning research procurement.

Before MiFID II, analyst research was provided “for free” in exchange for trading or investment banking business. The cost of research was cross-subsidized by trading commissions. Trading commissions were usually passed onto asset managers’ clients who were unaware of the embedded research cost (Figure 1.1a). This practice is known as bundling or payment via “soft dollars”,

Bundling traces its origin to 1950’s. At that time, commission fees used to be fixed at a high level (around 1% per share in the US before the May Day 1975) and brokerage houses compete

\(^10\) Even with RPA, clients are fully aware of the research cost. They are able to opt out purchasing research if they choose to.
\(^11\) “MiFID II, One Year On” CAF Institute, 2019.
for clients by providing “free” research services (Jones (2002)). Gradually, bundling also becomes demand-driven (Blume (1993) and Egan (2018)): Asset managers benefit from bundling. 12 “Soft dollars” are borne ultimately by end-investors. Payments for research in the form of trading commissions are not reported in the expense ratio and are not disclosed in the management fees. Hence, “soft dollars” are hidden cost for end clients. On top of that, the provision of “free” supplementary research services by the executing brokers – such as analyst research reports, corporate access, or other non-monetary benefits – creates incentives for asset managers to route the trades to that broker and the potential to trade more often than is appropriate for the clients. It may also preclude the use of other brokers who may provide better execution services.

Although the practice could hurt end investors, in 1975, the soft-dollar industry lobbied the US Congress to amend the 1934 Securities Exchange Act by adding Section 28 (e). This section allows asset managers in the US to pay for research with soft dollars from client commissions without breaking their fiduciary duty. Member states of the EU essentially followed the same practice before the onslaught of MiFID II.

EU regulators are increasingly concerned that bundling practice hurts end investors and overly favors asset managers. To protect end investors and promote transparency in the fee structure, they mandate unbundling as part of MiFID II. A subsequent effect of unbundling becomes manifested in the sell-side research industry. Unbundling completely changed the way sell-side research is compensated. It provides us a valid and timely setting where we can study how the compensation schedules for sell-side research affect its production.

The MiFID II regime, on its face, applies only to asset managers with a physical presence or domicile in the European Economic Area (EEA). This includes US asset managers with an authorized European subsidiary providing investment services to clients in the EEA. But given the regime’s complexity and wide-ranging reach, the key provision extends its influences globally. For example, US broker-dealers are indirectly affected by MiFID II if they provide investment research.

12 This is different from bundling in the traditional Industrial Organization literature. In many IO studies, bundling is a way for the supply side to price-discriminate consumers and extract consumer surplus. See a summary in Tirole (1988).
services to EEA firms or EEA clients, or to US asset managers that provide services to EEA clients. Receiving payments for research directly forces them to be registered as investment advisers subject to more strict regulation. Foreseeing the important consequences to US asset management business, in October 2017, the SEC issued a temporary no-action relief to “reduce confusion and operational difficulties that might arise in the transition to MiFID II’s research provisions” valid until 2020.\textsuperscript{13} The relief allows US broker-dealers to receive direct payments from EEA clients without being registered as financial advisers.

European Parliament began to discuss MiFID II in May 2014 and plan to apply it on January 3rd, 2017. Due to widespread concerns that the infrastructure would not be ready, the EU parliament decided to allow for a one-year delay for the launch of the regulation. On June 30, 2016, the Official Journal of EU announced that MiFID II was supposed to be translated into national law by July 3rd, 2017 and become fully applicable as from January 3rd, 2018.

It is worth noting that unbundling is only one of many important provisions of MiFID II. For example, under MiFID II, post-trade transparency is extended to non-equity instruments and post-trade information needs to be made public close to real time. High frequency trading firms are required to set order limits and use circuit breakers to limit or temporarily halt trading to avoid erroneous orders.\textsuperscript{14} However, unbundling is the most relevant and direct provision affecting analyst research. It is possible that other provisions in MiFID II also affect some part of analyst research, but as will be clear in the latter part of the paper, it would very difficult for other provisions to account for all the findings in our paper both at the firm level and at the analyst level. The impact of MiFID II on analyst research production arguably comes first and mostly from unbundling.

\textsuperscript{13}See SEC no-action relief, October 26th, 2017.
\textsuperscript{14}See a summary by Bank of America Merrill Lynch: https://www.bofaml.com/en-us/content/mifid-ii-regulation-summary-requirements.html
1.3 Hypotheses development

We begin by examining the effects of unbundling on research quantity. Research is a cost center and is difficult to price (e.g., Romer (1990)). Being unable to cross-subsidize the cost of research from transactions could force brokerage houses to scale back the amount of analyst research they produce. More importantly, if investors, who attach little value to research (e.g., passive investors), opt out from purchasing it, the total amount of wealth that can be allocated to research payments will further decrease. Consequently, using analyst coverage as a measure for research quantity, we anticipate that:

**H1:** *Sell-side analyst coverage of firms in affected regions will decrease compared with the coverage of firms in unaffected regions.*

We further test whether the coverage of small firms and large firms are differentially affected, which is the main concern of unbundling raised by market participants. We argue that large firms are more likely to be affected:

**H2:** *In affected regions, coverage of large firms will decrease MORE than of that of small firms.*

To understand the logic, notice that large firms usually have a larger amount of coverage to begin with. The large number of research analysts covering a certain firm probably implies that a lot of work was of low quality. Under unbundling, asset managers, who internalize research costs in most cases, will have to be much more selective to purchase research. Under such circumstances, we expect that asset managers opt for better research and discard low quality services. This indicates that large firms will lose more coverage.

The above hypothesis goes directly against a widespread concern in the market. Practitioners believe that small- and medium-sized firms will suffer more because they benefit the most from cross-subsidization. Trading commissions collected from trading firms with larger demand (large firms) are used to cover the research cost of small firms. This is no longer allowed under

---

15See “The EU’s Unbundling Directive is Reinforcing the Power of Scale”. The Economist, January 5th, 2019.
unbundling. Robert Ophèle, chairman of the Autorité des Marchés Financiers of France, commented that MiFID II has very detrimental effects on research, especially for mid-caps and that Europe was engaged in a “dangerous game” as research capacity was being pared back for many smaller companies. The answer to this apparent conflict in the priors should be an empirical one and is worth investigating.

We then move to study the impact of unbundling on research quality. We argue that one of the main drivers for the change in the quality is analyst competition. Unbundling makes analyst compete more directly in the quality domain, and the overall research quality will improve:

**H3:** Coverage quality for firms in affected regions will increase compared with unaffected regions.

There are at least two economic forces fostering competition. To begin with, asset managers can no longer blindly pass research costs to end clients. Paying for research out of their own pockets in most cases, they will be selective and opt for the best research. On top of that, unbundling changes the evaluation schemes for research analysts. Before the regulation, research was bundled with transactions. Research analysts used to have other incentives in addition to accuracy and quality (e.g., generating fees). Under unbundling, analysts are evaluated and compensated directly by the services they provide. Their incentives to provide high-quality research are strengthened. The two forces imply that analysts will compete more directly in the quality domain as they try to attract and maintain clients. It is then natural to anticipate that the overall research quality will improve.

It is worth noting that **H1** and **H3** altogether deviate from previous literature. It has been shown that coverage quality decreases together with coverage quantity (e.g., Hong and Kacperczyk (2010) and Merkley et al. (2017)). According to this view, the potential drop in the analyst coverage after MiFID II will imply the deterioration of the analyst coverage quality. We believe that the difference lies precisely in whether research is bundled or not. Under bundling, analysts are disciplined by

---

17 Even with RPA, asset managers will still be careful about quality because end clients can veto the purchase of unnecessary research.
their independent and impartial peers. A decrease in the coverage quantity implies not only a reduction in the peer competition but also an increase in the propensity to lose these impartial analysts. Consequently, coverage quality deteriorates. The unbundling requirement of MiFID II fundamentally changes how sell-side research is evaluated. For the first time, research analysts have to work out the value of their research, which ties closely to the quality. Fighting for the limited resources from the buy-side, they compete more directly in the quality domain and their incentives to provide better research are enhanced, regardless of the total number of analysts. The urge for independent and impartial peers as a discipline device is weakened.

If analysts compete more directly in the quality domain, we expect that the improvement in the coverage quality comes from two margins: 1) conditional on not exiting the analyst market, analysts produce better research (intensive margin); 2) everything else being equal, inaccurate analysts are more likely to exist the analyst market (extensive margin). “The analyst community will [split] into those really good analysts — who will be able to earn more than they currently do — and the average ones who will lose out.”18 These arguments lead to our fourth hypothesis:

**H4:** *Everything else being equal, analysts who stay provide better research than before and analysts who produce worse research are more likely to drop out.*

Importantly, we are not trying to argue that the competition among research analysts was non-existent before the regulation. In fact, competition has always existed. For example, the ranking provided by Institutional Investors on an annual poll is a critical metric to access analysts’ prestige. Career concerns and peer pressure have also served to discipline analysts (Hong and Kubik (2003), Hong and Kacperczyk (2010)). But compared with unbundling, these are all indirect mechanisms disciplining analysts. A direct mechanism such as unbundling, which puts an explicit price on the research services, strengthens the competition and enhances analyst incentives to produce better research.

18 Laurence Hollingworth, former vice-chairman of EMEA investment banking at JPMorgan. “MiFid II and the return of the ‘star’ analysts” Financial Times, February 26th, 2018.
1.4 Empirical design and data

1.4.1 Empirical design

We employ a difference-in-difference strategy. The strategy calls for proper treatment and control groups, which is not straightforward in the case of MiFID II. Unbundling in MiFID II puts constraints on asset managers, not on analysts nor publicly traded firms. It affects the trading relationship between asset managers and brokerage houses. However, systematic information on the trading relationship between the two parties is not available to researchers, especially in Europe. We have instead systematic information on analysts and publicly traded firms they cover.

To investigate the impact of unbundling on analyst research and test the hypotheses with the data we have, we leverage a well-established fact: local investors portfolios mostly consist of local securities (Home Bias).\(^{19}\) This implies that analysts covering EU firms are more likely to serve EU clients who are subject to MiFID II. In other words, these analysts are more likely to be affected. This further implies that coverage outcomes of EU firms domiciled and listed in the EU, the main targets of these analysts, are more likely to be affected as well.

We define the treatment group to be analysts covering mostly EU firms and firms domiciled and listed in Europe, and the control group to be analysts covering mostly US firms and firms domiciled and listed in the US. In our sample, analysts’ portfolios are extremely concentrated: in most cases, they focus either exclusively on EU firms or US firms (See Figure 1.7 in 1.12). We then define EU analysts to be the analysts whose portfolios consist of at least 70% of EU firms and US analysts to be analysts whose portfolios consist of at most 30% of EU firms. The heterogeneous impact of the regulation on analysts and firms between the two regions is the main variation through which we identify the impact of unbundling.\(^{20}\)


\(^{20}\)In theory, if an EU fund manager trades a US stock through a brokerage house, the US stock would be indirectly affected by the MiFID II. This leads to the absence of a perfect control group. What we explore here is the difference in the exposure of the treatment: EU analysts and EU firms are more exposed to the treatment than their US counterparts.
The identifying assumption is that, in the absence of MiFID II, coverage outcomes of the treatment group and the control group would have maintained parallel trend. The good news is that unbundling aims to impose transparency and protect investors. It is not a response to prior events regarding analyst research. This lends us plausible exogenous variations. Nevertheless, a cross-country comparison still calls for careful examinations of endogeneity issues that may violate the identification assumption. We consider a myriad of robustness checks to allow for the possibility of time-varying heterogeneity across differentially treated groups, including splitting post dummy into year dummies, adding timing-varying fixed effects and conducting several placebo tests. All the evidence suggests that the parallel-trend assumption holds and our empirical results are causal.

1.4.2 Data

Our data on analyst forecasts comes from I/B/E/S detail files. We focus on forecasts of annual earnings per share (EPS) in the current fiscal year since it is the most commonly issued forecasts in the I/B/E/S dataset and has the widest coverage. For each firm, we take the most recent forecast by each analyst to account for staleness issues. As a result, we have one forecast issued by one analyst for one firm in a given fiscal year. The sample selection is consistent with a large number of papers in the literature (e.g., Hong and Kacperczyk (2010) and Giroud and Mueller (2011)).

Annual balance sheet information comes from Worldscope. Stock exchange and price information come from Datastream. Since all three datasets are provided by Thomson Reuters, we are able to merge them using the unique I/B/E/S identifier. MiFID II went into formal discussion in the European Parliament in May 2014. To assess its impact, we choose our sample period from fiscal year 2014 to fiscal year 2018.\textsuperscript{21} Details of the construction of firm level observations can be found in Appendix 1.11.1.

To account for IPOs and delistings which mechanically affect analyst coverage, we construct a sample in which all firms have valid accounting and price information from 2014 to 2018. In other

\textsuperscript{21}Firms with fiscal-year end in December 2018 usually have valid financial information in early 2019. Our data runs until the calendar year May 2019.
words, all firms in our sample survived during the entire period. Moreover, we only consider firms that appeared at least once in the I/B/E/S dataset during the relevant time period. We put zero’s for analyst coverage when a firm is not covered by any analysts in a given year, hence not present in the I/B/E/S data set. Our final sample has 21,960 firm × year observations with 4,392 firms in each year. Out of all the firm × year observations, 486 of them do not have return on assets (ROA). Since all the missing values spread out evenly across the years and most of them appear only once for one firm, we fill in these missing ROAs using the industry level (Worldscope Item 06010) cross-sectional median in a given year.22

We construct three measures for firm level analyses: analyst coverage \((\text{Coverage}_{jt})\), forecast error \((\text{ForecastError}_{jt})\) and forecast dispersion \((\text{ForecastDispersion}_{jt})\). \(\text{Coverage}_{jt}\) is defined as the number of unique analysts covering firm \(j\) in fiscal year \(t\). \(\text{ForecastError}_{jt}\) is defined as the absolute difference between the firm’s actual EPS and the mean of the analyst forecasts. \(\text{ForecastDispersion}_{jt}\) is defined as the standard deviation of all the forecasts across all the analysts following the same firm in a given fiscal year. Following previous literature, we scale forecast error and forecast dispersion by the firm’s previous year-end stock price to mitigate heteroskedasticity concerns.23 To ensure the reliability of these two measures, we require at least 2 different analysts providing forecasts for the firm during the fiscal year.24

Since forecast error and dispersion are not defined for firms with coverage less than 2, we construct two samples for firm level analyses. The first one is the sample for coverage quantity, which contains all the firms satisfying our selection criterion (21,960 firm × year observations). This sample captures the change in the analyst coverage quantity even when firms have coverage less than 2, possibly zero (not being covered at all). The second one is the sample for coverage quality in which forecast error and forecast dispersion are properly defined. To limit the effects of outliers on our results, we remove observations in the sample for coverage quality for which

---

22 We obtain very similar results if we simply delete these observations. We also construct a balanced sample by deleting all the firm-year observations if one ROAs in one year is missing. We lose about 9% of the total sample but the results go through.


forecast error and forecast dispersion are larger than 10% of the firm’s previous year-end price (around 2% of the sample).\textsuperscript{25} Finally, to alleviate the concerns for the composition effect, we remove all observations of a firm from this sample if that firm’s forecast error or dispersion is not defined in any given year.\textsuperscript{26}

We present summary statistics for both samples in Table 1.1 and Table 1.2. In general, US firms are larger, have lower book-to-market ratios and lower ROAs.\textsuperscript{27} Looking at the sample for coverage quantity, we observe that on average US firms have more coverage than EU firms. While US firms have on average 11.145 analysts producing earning forecasts each year, EU firms have on average 8.558. Looking at the sample for coverage quality, we observe that both forecast error and dispersion are larger in the EU. The average forecast error is 0.724% for EU firms and 0.442% for US firms. Similarly, the average forecast dispersion is 0.848% for EU firms and 0.539% for US firms. The distribution of analyst coverage is quite left skewed in both regions, and more so in the EU (see Figure 1.8 in Appendix 1.12). Notice that for coverage less than 2, quality measures are not defined and those firms are excluded from the sample for quality. Even after the sample selection, we remain with a large number of unique firms in both regions, 1,111 for the EU and 1,693 to the US.

For analyst level analyses in the latter part of the paper, we focus on analyst-firm pairs. For each pair, we can calculate the forecast error by taking the absolute distance between the actual annual earnings and the analyst forecast, scaled by the firm’s previous year-end price. Similar to firm level data, we remove analyst-firm pairs for which the forecast error is larger than 10% of the firm’s previous year-end price (around 3% of the sample). Details of the construction of other important variables can be found in the Appendix 1.11.3.

\textsuperscript{25}See Gu and Hackbarth (2013) and Giroud and Mueller (2011).
\textsuperscript{26}For example, if firm A does not have a valid forecast error in one year, we remove all the observations of firm A from the sample. Our results hold if we only remove observations of A in that specific year.
\textsuperscript{27}In Appendix 1.15, we check that all our results are robust to a propensity score matching procedure in which we discard observations that are too different.
1.5 Investigating the impact of unbundling: firm level analyses

We first conduct firm level analyses and examine the impact of unbundling on firms’ coverage outcomes. We investigate the following empirical model:

\[ Y_{jt} = \beta_1 \text{Treat}_j \times \text{Post}_t + \beta_2 X_{jt} + \alpha_j + \alpha_t + \epsilon_{jt} \]  

(1.1)

\( Y_{jt} \) measures either analyst coverage \( \text{Coverage}_{jt} \), defined as the number of unique analysts covering firm \( j \) in fiscal year \( t \) or coverage quality \( \text{ForecastError}_{jt} \), defined as the absolute distance between the firm’s actual EPS and the mean of the analyst forecasts, scaled by the firm’s previous year-end price.\(^{28}\) \( \text{Treat}_j \) is a dummy equal to 1 if a firm is domiciled and listed in Europe. \( \text{Post}_t \) is a dummy equal to 1 if the fiscal year is equal to 2018. \( X_{jt} \) is a set of control variables including: log of market capitalization (LN SIZE), log of book to market ratio (LN BM), return on equity (ROA), total investment return in the current year (RET), return volatility (RETVOL), GDP growth rate (GDP GROWTH) and unemployment rate (UNEMPLOYMENT RATE) in the country where the firm is domiciled. When using \( \text{ForecastError}_{jt} \) as the dependent variable, we also control for the log of the average of the days between the analyst forecast date and the actual earnings report date (LN DISTANCE). This control is constructed from the analyst forecasts by taking the average over all the analysts following the same firm. It is likely to be an outcome of unbundling and suffers from the bad control problems (Angrist and Pischke (2009)) if we measure it year by year. Hence, we measure this variable in 2014 and interact it with \( \text{Post} \) (see for example Barrot (2016)).\(^{29}\) Since unbundling does not affect public firms directly, other firm level controls are not likely to be bad controls. Hence, we measure them year by year. We obtain very similar results if we measure them in 2014 and interact them with \( \text{Post} \). Finally, we also include firm and time fixed effects. \( H1 \) and \( H3 \) both imply that \( \beta_1 < 0 \).

To explore the heterogeneous response, we split firms into small firms and large firms. We

\(^{28}\)In the paper, analyst coverage will sometimes be referred as coverage quantity.

\(^{29}\)We obtain very similar results if we exclude this control.
first calculate the average market capitalization in the pre-regulation years for each firm. Small firms are defined as firms whose average market capitalization falls below the median. To maintain the proportion of EU and US firms fixed, we calculate the median cut-offs separately in both regions.\(^{30}\) We first run a difference-in-difference regression within the EU between small firms and large firms:

\[
Y_{jt} = \beta_3 \text{Small}_j \times \text{Post}_t + \beta_4 X_{jt} + \alpha_j + \alpha_t + \epsilon_{jt}
\]  

(1.2)

We then perform the following triple-difference regression:

\[
Y_{jt} = \beta_5 \text{Treat}_j \times \text{Post}_t + \beta_6 \text{Treat}_j \times \text{Post}_t \times \text{Small}_j + \\
+ \beta_7 \text{Small}_j \times \text{Post}_t + \beta_8 X_{jt} + \alpha_j + \alpha_t + \epsilon_{jt}
\]  

(1.3)

When \(Y_{jt}\) measures \(\text{Coverage}_{jt}\), \(H2\) implies both that \(\beta_3 > 0\) and \(\beta_6 > 0\).

To assess the validity of the parallel trends assumption and provide evidence on the dynamic impact of unbundling, we set up a Granger causality test, as suggested in Angrist and Pischke (2009). The goal is to see whether causes happen before consequences and not the other way around. To do this, we split the \(\text{Post}_t\) dummy into \(D_t\) year dummies, where \(t \in \{2015, ..., 2018\}\) and run the following specification:

\[
Y_{jt} = \sum_t \eta_t (D_t \times \text{Treat}_j) + \beta_9 X_{jt} + \alpha_t + \alpha_j + \epsilon_{jt}
\]  

(1.4)

We choose 2014, the year in which MiFID II went into discussion as our reference year. Standard errors in all the analyzes are clustered at the country level. Although we already control for the firm level characteristics, we repeat the analyses using a propensity score matching procedure to mitigate the concerns that our results may be driven by the observable differences at the firm level. All the results are robust to the matching procedure and details can be found in Appendix 1.15.

\(^{30}\)To account for the possibility that firm size changes over the years and large firms may become small firms after the regulation, we also define the \(\text{Small}\) dummy using the market capitalization in each year. This way, the variable \(\text{Small}\) varies over the years. We obtain virtually the same results.
1.5.1 The impact of unbundling on analyst coverage

We first look at the overall effect of unbundling on analyst coverage. Figure 1.2 shows the time series trends in the average analyst coverage of EU and US firms. A sharp decrease in the analyst coverage of EU firms can be observed after unbundling. In fact, Panel A of Table 1.5 shows that the average analyst coverage of EU firms changes from 8.74 in the pre-period to 7.831 in the post-period. Even if there seems to be a small decrease in the average coverage of US firms as well, the magnitude is much smaller and is not statistically significant.

We formally test HI by estimating Model (1.1). Column (1) of Table 1.6 reports the result. After unbundling, analyst coverage of EU firms changes by $-0.651$. This translates into a decrease of 7.45% relative to the average analyst coverage of these firms prior to the regulation.\footnote{We perform the same analyzes in the sample for coverage quality. Results are qualitatively the same and can be found in Table 1.22 of Appendix 1.14.2.}

Model (1.1) is suitable to recover the parameter of interest $\beta_1$ only if the parallel trend assumption holds. We test this assumption by estimating the dynamic coefficients described in Model (1.4). Figure 1.5a shows the plot of dynamic coefficients with a 95% confidence interval.\footnote{Point estimates are reported in Table 1.20 in the Appendix 1.14.1. In the Appendix the readers can find point estimates for all other dynamic coefficient plots presented in the following sections.} The dynamic coefficients are significant only for 2018. This is strong support for the parallel trend assumption. We further conduct a placebo test in which we focus on the pre-regulation years and define the Post dummies as if the regulation occurred in one of these pre-regulation years. If the parallel trend assumption holds, the coefficient in front of the interaction term $Treat_j \times Post_t$ should not be statistically significant and if it is, the magnitude should be much smaller. Column (1) to (3) in Panel A of Table 1.10 shows the results. None of the coefficients are statistically significant. We take these results as evidence that parallel trend assumption is very likely to hold in our setting. It is reasonable to believe that our control group establishes a valid counterfactual of what would have happened to the treatment group in the absence of the reform.

We conclude that unbundling has had a negative effect on analyst coverage. Analyst coverage has been widely used as a measure of research production quantity. We interpret the decrease
in coverage as evidence that there is a causal impact of unbundling on the quantity of research produced by analysts.

We now turn to explore the heterogeneous effects of the regulation on analyst coverage (\textbf{H2}). Both Figure 1.3 and Panel A of Table 1.5 show clear evidence that the analyst coverage of small firms remains relatively unchanged and the analyst coverage of large firms decreases significantly. These results are also confirmed in the formal difference-in-difference estimation. Column (2) of Table 1.6 shows that the analyst coverage of small firms in the EU decreases much less than that of large firms in the EU. The result is further strengthened by the triple-difference analyses in column (3). The coefficient in front of the triple-difference term is 1.891 and statically significant. For large firms, the analyst coverage change by $-1.594$, which translates into a 10.53\% decrease in the average analyst coverage of large firms prior to the regulation. In Table 1.23 of Appendix 1.14.2, we present the regression results in logs. All the results are qualitatively the same. These results support \textbf{H2}: analyst coverage of large firms decreases more than that of small firms.\footnote{In Table 1.24 of Appendix 1.14.2, we also perform difference-in-difference analyses between EU firms and US firms but separately for small and large firms (e.g., small EU firms vs. small US firms; large EU firms vs. large US firms). We obtain very similar results: the decrease in the coverage is concentrated in large firms.}

In summary, we have shown that unbundling causes an aggregate decrease in the analyst coverage of EU firms compared with US ones and that this decrease is more pronounced for large firms. These are evidences in support for \textbf{H1} and \textbf{H2}. The results are surprising: it contradicts the common view in the media that small firms would be most affected. We interpret these results through lenses of competition in the market for analyst research. When it becomes mandatory for investors to pay separately for research, they are likely to stop buying low-quality research. Large firms on average have more coverage than small firms and the probability that low-quality research is produced is higher. If investors opt out of inferior research, large firms are more affected.

\subsection*{1.5.2 The impact of unbundling on analysts forecast quality}

After establishing the negative effect of unbundling on analyst coverage (coverage quantity), we turn to explore the impact on coverage quality. Panel B of Table 1.5 presents the results of a
simple one-difference analysis. It shows that the forecast error decreases for EU firms. In Figure 1.4, we plot the empirical cumulative distribution of the forecast errors. We can see that for EU firms, the distribution of the post-year is “smaller than” the distribution of the pre-years since the post-year distribution lies mostly above that of the pre-years. We do not observe a similar trend among US firms. We also plot the histogram of the forecast errors in the pre- and post-regulation years in Figure 1.9 of Appendix 1.12 and observe a similar pattern: the histogram in the post-year shifts towards 0 for EU firms, indicating an overall decrease in the forecast error. Again, we do not observe a clear similar pattern for US firms. All the evidence suggests that the forecast error of EU firms decreases after unbundling.

Table 1.6 reports the formal difference-in-difference results. Notice that the number of observations changed from the previous analyses on coverage quantity, because we now focus on the sample for coverage quality. Column (4) of Table 1.6 shows that overall, the change in the forecast error of EU firms is $-0.142\%$. This translates into a decrease of 19.19\% relative to the average forecast errors of EU firms prior to the regulation. The statistically and economically significant decrease shows that unbundling has caused analyst research to be more precise, an indication of quality improvement. Column (2) shows that the change in the forecast errors is comparable among both small and large firms. In fact, by checking the triple interaction term in column (3) of the same table, we find no evidence that the coverage quality of small and large firms are affected differentially. Finally, we perform the dynamic coefficient test described in Model (1.4). The results can be found in Figure 1.5b. We do not find evident violations of the parallel trend assumption. Placebo tests in column (4) to (6) of Table (1.10) show reassuring evidence. Hence, we are confident in claiming that unbundling causes an increase in analyst research quality in the EU.

The results presented in this subsection are in line with hypothesis \(H3\). Even though we document a decrease in the coverage quantity due to unbundling, we find that the coverage quality improves. This provides supportive evidence that unbundling disciplines analyst forecast behavior. Being evaluated directly and fighting for limited resources from the buy-side, analysts compete more intensively among the quality domain and produce better forecasts.
1.6 Investigating the channel: analyst level analyses

To reconcile the simultaneous decrease in the coverage quantity and the increase in the coverage quality, we argue that analyst competition is enhanced. Quality becomes the main concern of analysts due to unbundling. The enhancement of competition 1) motivates existing analysts to produce better forecasts (intensive margin) and 2) makes analysts who produce worse forecasts more likely to drop out of the analyst market (extensive margin). To find evidence of the two margins, we now zoom in to the analyst level and test ($H4$).

1.6.1 Intensive margin

We begin with the first part of $H4$. To compare forecast accuracy measured by forecast error within the same analyst over different years, we focus on analyst-firm pairs that exist throughout the whole sample period (from 2014 to 2018). We conduct analyses similar to the ones done at the firm level:

\[ Y_{ijbt} = \gamma_1 \text{Treat}_i \times \text{Post}_t + \gamma_2 X_i \times \text{Post}_t + \gamma_3 Z_{ijt} + \alpha_i + \alpha_j + \alpha_b + \alpha_t + \epsilon_{ijbt} \]  

(1.5)

$Y_{ijbt}$ is the forecast error analyst $i$ working in brokerage house $b$ incurs on firm $j$ in fiscal year $t$ (the absolute difference between firm $j$’s actual earnings and the analyst $i$’s forecast, scaled by firm $j$’s previous year-end price). $\text{Treat}_i$ is a dummy equal to 1 if the analyst is an EU analyst. $\text{Post}_t$ is a dummy equal to 1 if the fiscal year equal to 2018. Analyst fixed effect $\alpha_i$ ensures that we compare forecast error within the same analyst across different years. Ideally, we should include firm $\times$ year fixed effect ($\alpha_{jt}$) to subsume any firm-level, time-varying heterogeneity that may confound our causal interpretation. However, the geographic concentration of analyst portfolios as shown in Figure 1.7 implies that $\alpha_{jt}$ will almost subsume $\text{Treat}_i \times \text{Post}_t$. Instead, we can only include firm fixed effect and year fixed effect. We also include brokerage fixed effect $\alpha_b$ to control for brokerage firms the analysts work for.
To control for time-varying heterogeneity, we include both analyst-level controls and firm-level controls. $X_i$ is a vector of standard analyst-level controls that are shown to affect analyst forecast errors, including the log of the number of firms the analyst follows (LN FIRMS COVERED), the log of the average coverage of the portfolio firms that the analyst follows (LN AVERAGE COVERAGE), the log of the number of years one analyst worked (LN TENURE) and the log of the days between the forecast date and the earnings report date (LN DISTANCE). These variables are likely to be the outcomes of unbundling and serve as bad controls if we measure them year by year. Hence, we measure them in 2014 and interact them with Post. We obtain similar results if we simply exclude these variables. $Z_{ijt}$ include firm level controls as defined in Section 1.5. Since unbundling does not affect public firms directly, these controls are not likely to be bad controls and we measure them year by year. We obtain similar results if we measure them in 2014 and interact them with Post. Standard errors are clustered at the firm level.\textsuperscript{34} $H4$ implies that $\gamma_1 < 0$.

To mitigate the concern for idiosyncratic noise in analyst-firm level analyses, we aggregate our measures at the analyst level. To be more specific, we estimate the following model:

$$Y_{ibt} = \gamma_4 TREAT_i \times Post_t + \gamma_5 X_i \times Post_t + \gamma_6 Z_{it} + \alpha_i + \alpha_b + \alpha_t + \epsilon_{ibt}$$ (1.6)

$Y_{ibt}$ is the average of all the forecast errors analyst $i$ working in brokerage house $b$ incurs in year $t$. $TREAT_i$ and $Post_t$ are defined as before. In the baseline regression, we include brokerage house + year fixed effect ($\alpha_b + \alpha_t$). We also investigate the specification where we include brokerage house $\times$ year fixed effect ($\alpha_{bt}$), which kills all the time-varying brokerage house level variation that may confound our causal interpretation. This strong fixed effect may also kill some meaningful variation. For example, if unbundling disciplines analysts through the brokerage house they work for (e.g., brokerage houses decide to put more pressure on analysts to produce better research after MiFID II), $\alpha_{bt}$ will prevent us from capturing this channel. Hence, we take this specification as a robustness check.

Similar to firm level analyses, we estimate a triple-difference model with respect to analysts

\textsuperscript{34}This is a conservative way of clustering. Clustering at analyst level leads to smaller standard errors.
mostly covering relatively small firms (small analysts) and analysts mostly covering relatively large firms (large analysts). To achieve this, we first calculate the average market capitalization of all the firms within an analyst portfolio in a given year. We then average the average market capitalization across all the pre-regulation years. Small analysts are defined as analysts whose average of the portfolio average market capitalization falls below the median.\textsuperscript{35} As before, we calculate the median cutoff separately in both regions. To aggregate firm level controls to the analyst level, we simply average the firm level variables within an analyst’s portfolio. Standard errors are clustered at the analyst level. \textit{H4} implies that $\gamma_4 < 0$. Table 1.3 reports some summary statistics of the sample we use. EU analysts, in general, follow fewer firms and have larger forecast errors.

Table 1.7 reports the regression results. Column (1) and Column (2) present results at the analyst-firm pair level. Column (3) and (4) present the results at the analyst level with brokerage house + year fixed effects. Column (5) and (6) present the results at the analyst level with brokerage house $\times$ year fixed effect. The results are consistent over all these specifications: analysts who stay after the regulation produce lower forecast errors. Using column (2) as the main specification, the average forecast error of EU analysts changes by $-0.131\%$, which translate into a $16.01\%$ decrease in the average forecast errors of EU analysts prior to the regulation. The triple interaction term further shows that the magnitude in the decrease is comparable among analysts who focus on small firms and analysts who focus on large firms. Notice that the total number of observations drops a lot in the brokerage house $\times$ year fixed effect specification. If we allow for brokerage house $\times$ year fixed effect, all the meaningful variations come from brokerage houses that covering both EU firms and US firms. Hence we focus on analysts who work for multi-continent brokerage houses in this specification. Contrary to the analyst level, we do have a sufficient number of brokerage houses that cover firms in both continents. Even with a smaller number of observations, we still observe a statistically significant decrease in the average forecasting error.

\textsuperscript{35}In principle, one can define small analysts to be analysts focusing on small firms. But analyst portfolio is a mix of both small and large firms. In most cases, large firms are the majority. Focusing on analysts covering only small firms leads to limited observations.
Figure 1.6 presents the dynamic coefficients plots for all three specifications with a 95% confidence interval. Unbundling takes effect only after the implementation date. Placebo tests in Panel B of Table 1.10 show similar evidence: none of the coefficients are statistically significant in the pre-regulation years. The results all suggest that the parallel trend assumption is very likely to hold in our setting.

The results presented in this subsection are in line with the first part of hypothesis \( H4 \): affected analysts who continue to work after the regulation produces better forecasts.

### 1.6.2 Extensive margin

*Baseline comparison*

We now turn to the second part of \( H4 \). We want to compare forecast quality between analysts who drop out due to unbundling and analysts who do not. To achieve this, we need a quality measure that is comparable across different analysts. The simplest measure is the average forecast errors incurred by the analyst as defined above. However, different analysts tend to cover different firms. Some firms are more difficult to forecast than others. An analyst with a high average forecast error may either be 1) analyst who does not perform well or 2) analyst who follows firms more difficult to analyze. Therefore, the average forecast error is problematic when we want to compare across different analysts.

To account for all these issues, we follow Hong and Kubik (2003) to construct a relative accuracy measure comparable across different analysts. First, we sort the analysts covering a specific firm in a year based on their forecast errors. Then, we assign a rank based on the sorting: the best analyst receives the first rank. In case of a tie, we assign all those analysts the midpoint value of the ranks they take up. Notice that the maximum rank an analyst can get depends on the total number of analysts covering the firm. Analysts covering thinly followed firms will on average have lower ranks and are deemed to be better. To take this into account, we scale an analyst’s rank for a firm
by the number of analysts that cover that firm. Formally, we use the following score measure:

$$\text{Score}_{ijt} = 100 - \frac{\text{Rank}_{ijt} - 1}{\text{Number of Analysts}_{jt} - 1} \times 100$$ (1.7)

where $\text{Number of Analysts}_{jt}$ is the number of analysts who cover firm $j$ in year $t$. The more accurate the forecast is, the lower the rank and the higher the score. We compute the scores for every firm $j$ in year $t$. To reduce idiosyncratic noise in the specific analyst-firm pairs and get the analyst level score in the pre-regulation years, we follow Hong and Kubik (2003) and define the relative accuracy measure ($\text{RelativeAccuracy}_i$) to be:

$$\text{RelativeAccuracy}_i = \frac{1}{T} \sum_t \left( \frac{1}{J} \sum_j \text{Score}_{ijt} \right)$$

where $J$ is the set of firms analyst $i$ covers in year $t$ and $T$ includes all the years prior to 2018. To increase the likelihood that the analyst is affected by MiFID II, we focus on analyst-firm pairs that exist consecutively in 2015, 2016, 2017. If the analyst covers the stock for three consecutive years and suddenly stops covering it after the regulation, it is more likely that the stop is due to MiFID II.

Table 1.4 presents summary statistics for the sample we use for the extensive margin analyses in this subsection. By construction, the mean and median of the relative accuracy measure is around 50. The analysts we study covers on average 10 firms and have worked for about 10 years.

We focus on EU analyst and run the following regression at the analyst-firm level:\footnote{We also use $\text{RelativeAccuracy}_{ij}$, relatively accuracy defined on analyst-firm pair by averaging across the years, for robustness checks and results are qualitatively the same.}

$$\text{Stop}_{ij} = \pi_1 \text{RelativeAccuracy}_i + \pi_2 X_{ij} + \alpha_j + \epsilon_{ij}$$ (1.8)
and the following regression at the analyst level:\textsuperscript{37}

\[ \text{DropOut}_i = \pi_3 \text{RelativeAccuracy}_i + \pi_4 X_i + \epsilon_i \]  

(1.9)

On the analyst-firm level, we define a dummy variable \((\text{Stop}_{ij})\) equal to 1 if the analyst \(i\) stops covering the firm \(j\) after the regulation. At the analyst level, we define a dummy variable \((\text{DropOut}_i)\) equal to 1 if the analyst stops covering all the firms he/she used to cover before the regulation. The firm fixed effect \(\alpha_j\) ensures that we compare different analysts within the same firm. Controls \(X_{ij}\) in Model (1.8) include the log of the number of unique firms the analyst follows prior to 2018 (LN FIRMS COVERED), the log of the average of the average coverage of the portfolio firms that the analyst follows across all the years prior to 2018 (LN AVERAGE COVERAGE) and the log of the number of years one analyst has worked prior to 2018 (LN TENURE).\textsuperscript{38} We do not have firm level controls due to the fixed effect \(\alpha_j\). In addition to these variables, controls \(X_i\) in Model (1.9) further include standard firm level controls aggregated at the analyst level as defined in the previous sections. Standard errors are clustered at the analyst level for Model (1.8) and the brokerage house level for Model (1.9). \textbf{H4} implies that \(\pi_1 < 0\) and \(\pi_3 < 0\).

Column (1) in Table 1.8 shows the result of Model (1.8) at the analyst-firm level with firm fixed effect. The result confirms our hypothesis: higher relative accuracy (better forecast quality) makes it less likely for an analyst to stop covering a firm. Within the same firm, if the analyst’s relative accuracy improves by 1 point, he/she is 83.9 basis points less likely to stop covering a firm. This translates into a 3.93\% drop in the unconditional probability of stopping covering a firm after the regulation.

Column (3) shows the result of Model (1.9) at the analyst level with the brokerage house fixed effect. The dependent variable is a dummy equal to 1 if the analyst stops covering all firms he/she used to cover prior to the regulation. As we can see, a one-point increase in the relative accuracy

\textsuperscript{37}We also conduct logit regression and results are very close. It is more econometrically appealing to cluster standard errors and to include fixed effects in the linear model.

\textsuperscript{38}For LN AVERAGE COVERAGE, we first take the average over the firms the analyst covers and then take the average over all the years prior to 2018. All the variables are measured prior to 2018 and are not likely to be affected by the regulation. Our results are robust when we measure them in 2015 or just simply exclude them from the regression.
results in a 62.2 basis decrease in the probability of dropping out of the sample. This translates into a 2.58% decrease in the unconditional probability of dropping out. Suppose that there are 8 analysts (the average number of analysts covering an EU firm prior to the regulation) covering a firm. If one of them improves his/her ranking from 4th to 3rd, his/her relative accuracy will improve by 14.3 points. This makes him/her 8.89% less likely to drop out and potentially be laid off (14.3 × 0.00622). Moreover, the coefficients in front of analyst experience (LNTENURE) is positive and highly significant. This shows that experienced analysts are less likely to drop out. To the extent that analysts with longer experiences tend to produce better forecasts, this result conveys further supportive evidence that more accurate analysts are more likely to stay.

**Pre and post comparison**

To sharpen our analyses, we perform similar regressions year by year and investigate whether analyst accuracy matters more after unbundling. To achieve this, we define a dummy variable Stop_{ijt+1} equal to 1 if analyst i covering firm j at year t stops covering the same firm at year t + 1. At the analyst level, we define a dummy variable DropOut_{it+1} equal to 1 if the analyst stops covering all the firms he/she used to cover at year t. We then turn (1.8) and (1.9) into:

\[
\text{Stop}_{ijt+1} = \pi_5 \text{RelativeAccuracy}_{it} \times \text{Post}_t + \pi_6 \text{RelativeAccuracy}_{it}
+ \pi_7 X_{ijt} + \alpha_j + \alpha_b + \alpha_t + \epsilon_{ijt} \quad (1.10)
\]

\[
\text{DropOut}_{it+1} = \pi_8 \text{RelativeAccuracy}_{it} \times \text{Post}_t + \pi_9 \text{RelativeAccuracy}_{it}
+ \pi_{10} X_{it} + \alpha_b + \alpha_t + \epsilon_{it} \quad (1.11)
\]

Relative Accuracy_{it} is the average of Score_{ijt} over all firm j analyst i covers at year t. Post_t is a dummy equal to 1 for year equal to 2017.\(^{39}\) Controls X_{ijt} and X_{it} are defined in a similar fashion. Standard errors are clustered at the firm level for (1.10) and at the brokerage house level for (1.11).

\(^{39}\)Following Hong and Kubik (2003), we assume that the employment decision depends on the analyst performance in the previous year.
The question we seek to answer here is that among the analyst population, whether more inaccurate analysts are on average more likely to stop/drop out after unbundling. If the answer is yes, we anticipate $\pi_5 < 0$ and $\pi_8 < 0$ in the EU analyst sample but not in the US analyst sample. Notice that (1.10) and (1.11) are not standard panel regressions. They are different cross-section regressions stacked over time. The way we define $Stop_{ijt+1}$ and $DropOut_{it+1}$ conditions on analyst covering the firm at $t$. If analyst $i$ stops or drops out at $t + 1$, he/she will not be included when we define dummies for year $t + 2$. In fact, it is common that analysts who drop out in one year never reappear in the later years, i.e., many analysts appear only once in the analyses. For this reason, we do not include the analyst fixed effect. We include other important controls as defined in the above subsection. Instead of taking the average over all the pre-regulation years, we need to measure them year by year. Since the latest control variables are measured in 2017, they are determined before the regulation and are not likely to be the outcomes of the regulation.

Table 1.9 presents the results. Column (1) shows the result of the analyst-firm level focusing on EU analysts. As predicted, the coefficient in front of $RelativeAccuracy_{it} \ast Post_{i}$ is negative and statistically significant, which implies that analysts producing better forecasts are more likely to keep covering the same firm after the regulation. Column (2) presents the same analysis on US analysts and we do not observe a similar pattern. This provides further evidence that analysts with worse forecast records are more likely to stop covering a firm after unbundling only when they are affected. We perform a triple-difference-type analysis in column (3), and the coefficient in front of the triple interaction terms again confirms our hypothesis: forecast accuracy matters more after unbundling for affected analysts. The results using $DropOut_{it+1}$ as the dependent variable can be found in column (4) to (6). The magnitude of all the coefficients is comparable with the ones obtained at the analyst-firm level. Notice that when an analyst drops out, he/she stops covering all the firms. This is a strong requirement because many analysts stop covering a set of firms without completely dropping out. Thus, the variation at the analyst level may not be enough to allow for a triple-difference-type analysis. This could explain why the coefficient in front of the triple interaction term is not statistically significant.
In summary, the results presented in Table (1.8) and (1.9) are all in line with the second part of \textit{H4}. In other words, analysts who have lower relatively accuracy (analysts who produce worse forecasts) are more likely to stop covering a firm and even drop out. Forecast accuracy matters more after unbundling only among affected analysts. This section provides additional evidence on the competition channel: analysts compete for research quality due to unbundling. The ones with worse forecast records are more likely to be cast out.

1.7 Robustness check and extensions

Previous sections focus on analyzing the impact of unbundling on sell-side research production and the underlying channel driving the results. We now conduct a few additional analyses to show that our findings are robust. We also study how unbundling affects brokerage house employment and what the capital market effects of MiFID II are.

1.7.1 Firm’s brokerage house coverage

In the previous analyses, we show that firms’ analyst coverage drops. One concern is that this drop comes from brokerage houses laying off analysts providing duplicated research services. For example, if two analysts cover the same firm in the same brokerage house, the one providing duplicated or worse research may be laid off. If this is the case, a drop in the analyst coverage does not imply a drop in the brokerage house coverage. Investors are still able to get access to the same amount research from these brokerage houses and potentially from better analysts. In other words, the quantity of research does not decline.

Although one brokerage house rarely hires two analysts covering the same public firm due to redundancy, we check firms’ brokerage house overage directly.\footnote{This is a well-established empirical fact (e.g., Hong and Kacperczyk (2010)). In our sample, among all the brokerage house-firm pairs, 86\% of them has only one analyst coverage.} We perform difference-in-difference analyses similar to Model (1.1) but replace the dependent variable with the number of unique brokerage houses covering the firm. Under the condition that the relationship between
investors and brokerage houses are sticky (as shown, for example, in Di Maggio et al. (2019)), a reduction in firms’ brokerage house coverage implies that investors lose access to research on these firms from these brokerage houses, which further suggest a decrease in the quantity of research supplied to these investors.

Table 1.11 shows the results. As we can see, firms’ brokerage house coverage drops after unbundling. Similar to previous analyses, this decrease is driven by large firms. Large firms lose more than 1 brokerage house after this regulation. The results here suggest that investors’ access to firms’ research through brokerage houses may reduce. From these investors’ point of view, the quantity of research decreases.41

1.7.2 Other quality measures

Analyst forecast quality is hard to measure. To show that our results are robust, we repeat our analyzes using other commonly used measures of forecast quality.

Non-market-response based measures

First, we follow previous research and use forecast dispersion, defined as the standard deviation of all the forecasts across all the analysts following the same firm in a given fiscal year, as a proxy for firm level forecast quality (e.g., Gu and Hackbarth (2013), Behn et al. (2008)). One intuition behind this measure is the following: if accuracy becomes analysts’ main concern, dispersion will decrease when all of their forecast quality improves and the forecast values converge to the true value.

One-difference analyses in Panel C of Table 1.5 reveal that the dispersion of EU firms decreases after the regulation. Difference-in-difference analyses in Panel A of Table 1.12 further confirm the result: unbundling causes a decrease in the forecast dispersion of the EU firms compared with the US firms. The magnitude is comparable between small firms and large firms. These results are consistent with the analyses using forecast error. Dynamic coefficient plot in Figure 1.5c and

41In untabulated test, we also find that the number of unique firms EU brokerage houses cover decreases.
placebo tests in column (7) to (9) of Table 1.10 both show no evident violations of the parallel trend assumption. We conclude that the coverage quality of EU firms, measured by forecast dispersion, indeed improves after the regulation. The analyses also show that the beliefs of different analysts are converging after this regulation.

Another measure for quality is the number of firms each analyst covers. It is reasonable to believe that if one analyst covers fewer firms, due to specialization, he/she will devote more time to each firm and provide better forecasts. Panel B of Table 1.12 shows the results of the difference-in-difference analyses similar to Model (1.6). The dependent variable is the number of firms each analyst covers. We focus on analysts who cover at least one firm throughout the sample in Column (1). As we can see, EU analysts cover less firms after the regulation and the decrease is statically significant. We find similar evidence in column (2) in which we focus on all the analysts in our sample. These results further support that the quality of analysts research in EU improves after unbundling.

*Market-response based measures*

We can also assess analyst forecast quality using market responses to analyst forecasts. Intuitively, if analyst forecast quality improves and contains more information, stock prices should move more substantially on the analyst forecast revision dates. Hence, we focus on the absolute market-adjusted abnormal returns on the analyst forecast revision dates. Improvement in the forecast quality suggests more significant absolute market-adjusted abnormal returns on the revision dates.

We define three measures to capture this effect. At the analyst-firm level, we calculate the absolute market-adjusted abnormal return on the forecast revision date:

$$ABRet_{j,i,d} = |Ret_{j,i,d} - Ret_{m,d}|$$  \hfill (1.12)

\footnote{In this section, we focus on the revision dates, i.e., dates on which analysts revise his/her forecasts. Our results, both at the analyst-firm level and at the firm level, are robust if we focus on all dates on which analysts issue forecasts.}

32
\( Ret_{j,i,d} \) denotes the daily return of firm \( j \) when analyst \( i \) revises forecast for firm \( j \) at date \( d \).\(^{42}\) To mitigate data errors in Datastream, we winsorize \( Ret_{j,i,d} \) at the 1% level. \( Ret_{m,d} \) denotes the daily return of the stock market. Here, we define the market return for EU and US separately by calculating the value-weighted stock returns of all the firms in each continent. If multiple revisions occur on the same day for the same firm, we assign each analyst covering that firm the same abnormal return. In an extreme case, we could attribute the abnormal returns generated by more informative analysts to analysts who simply herds with the others. To mitigate this concern, we keep only the first three analyst revisions by the exact revision time. \( ABRet_{j,i,d} \) then helps us to capture the new information in the forecast revisions at the analyst-firm level.

At the firm level, we aggregate analyst revisions. We first compute a yearly measure of aggregate analyst informativeness (AGAI). Following previous literature (e.g., Frankel et al. (2006) and Lehavy et al. (2011)), we define AGAI as follows: For a given firm in a given fiscal year, we first sum the absolute market-adjusted abnormal returns on all the forecast revisions dates across all the analysts. We then divide it by the sum of absolute market-adjusted abnormal returns of all trading days for the firm in the given fiscal period. Specifically,

\[
AGAI_{jt} = \frac{\sum_{d=1}^{NREVS} |Ret_{j,d} - Ret_{m,d}|}{\sum_{d=1}^{TDAYS} |Ret_{j,d} - Ret_{m,d}|} \tag{1.13}
\]

where \( Ret_{j,d} \) denotes the daily return of firm \( j \) and \( Ret_{m,d} \) denotes the daily return of the stock market. \( NREVS \) is the set of unique forecast revision dates over all the analysts for the given firm in the given fiscal year. We exclude forecast revision dates that coincide with earnings announcements. This mitigates the concern that analyst forecast revisions respond to publicly available earnings information. Intuitively, AGAI measures the aggregate abnormal returns on all forecast revision dates. It allows us to capture the total amount of new information in the analyst forecast revisions.\(^{43}\) Furthermore, AGAI incorporates changes in the total number of forecasts (the quantity of information) caused by both changes in the number of analysts covering a firm and the number

\(^{43}\)AGAI treats multiple revisions by different analysts on the same firm as one aggregate revision. It measures the aggregate information generated by all analysts on a given revision date.
of forecasts issued by one analyst. All else being equal, a decrease in the total number of forecast revisions causes a decrease in AGAI.

To account for the changes in the total number of forecast revisions, we normalize AGAI by the number of forecast revision dates of a given firm in the given fiscal year. Formally, we define the average analyst informativeness as:

\[
AVGAI_{jt} = \frac{\sum_{d=1}^{NREVS} |Ret_{j,d} - Ret_{m,d}|}{\sum_{d=1}^{TDAYS} |Ret_{j,d} - Ret_{m,d}|} \times \frac{1}{NREVS_{j}}
\]

where \( NREVS_{j} \) is the number of unique forecast revision dates. AVGAI then measures the average informativeness of one analyst forecast revision date.

We perform similar difference-in-difference analyses (Model 1.1) using the three new measures. Standard firm level controls are included. At the analyst-firm level, we include the log of the number of days between two revision dates as an additional control. Standard errors are clustered at the firm level.

We report the results in Table 1.13. Column (1) shows the results at the analyst-firm level. We include analyst × firm fixed effect to capture changes in the abnormal return within the same analyst-firm pair across different times. As we can see, analyst forecasts generate larger abnormal returns after unbundling. This is evidence that individual analyst forecasts become more informative. Column (2) aggregates different analyst revisions at the firm level. It shows that the average informativeness of one revision date improves after the regulation. These results provide further evidence that individual analyst quality increases. In column (3), we show the results of aggregate informativeness without normalizing the number of forecast revision dates. Interestingly, the coefficient in front of \( TREAT \times POST \) becomes statistically negative, suggesting a likely decrease in the aggregate informativeness of analyst forecasts.

We interpret the results as a trade-off between quantity and quality. While individual forecasts become more informative, the total number of valuable forecasts decreases after the regulation. Hence, even though average analyst informativeness increases, the aggregate informativeness of

\footnote{Results of small firms and large firms can be found in the Appendix 1.14.3.}
analyst forecasts decreases.\textsuperscript{44}

As a final note, we are aware of potential concerns with abnormal returns on the revision dates. For example, analysts could respond to prior public disclosure events. In this case, their revisions could appear to provide more information than they actually do. We try to address these concerns by focusing on a short window (1-day) and by excluding revision dates coinciding with earnings announcement dates. But still, abnormal returns on the forecast revision dates are not the cleanest measure. Hence, we view our results here suggestive rather than conclusive.

1.7.3 Alternatives

In our paper, we argue that the enhancement of competition among analysts in the quality domain is one driver of our results. One assumption we rely on is that asset managers internalize research costs and put more emphasis on research quality. What if these asset managers just claim to internalize the cost but in effect still make hidden research payments through inflated commissions? If that is the case, they probably would not focus a lot on research quality because end investors bear the costs. Although it is hard to preclude this possibility based on the data we currently observe, for this argument to hold, commissions fees should at least stay more or less the same, if not increase. However, recent surveys suggest that commissions fees drop by around 30\% due to unbundling.\textsuperscript{45} A study by Financial Conduct Authority (FCA) also finds no evidence of asset managers passing research cost through trading commissions and break the new rule. According to this study, expensing research costs out of asset managers’ own profits already saved end investors in the UK around £180 million in 2018. The FCA expects that the overall savings from MiFID II rules could reach £1 billion over the next five years.\textsuperscript{46}

Moreover, unbundling makes many end investors become aware of the hidden cost. End investors such as pensions, insurance companies, and sovereign funds have always been proponents of unbundling due to opaque fee structures and inflated trading costs. As a result, asset managers

\textsuperscript{45}“Bank and brokers suffer ‘dramatic’ fall in commissions” Financial Times, June 2nd, 2018.

\textsuperscript{46}Andrew Bailey, Chief Executive of the FCA, keynote speech on MiFID II at the European Independent Research Providers Association.
face greater pressure to reduce commission fees after unbundling. Hiding research cost by inflating fees not only exposes asset managers to severe legal risks but put them at a disadvantage amid the increased market competition in the asset management industry. All these arguments suggest that it is unlikely for asset managers to hide research costs and continue the “soft dollar” practice as under bundling.

Another common worry is that other provisions in MiFID II may also affect sell-side research. For example, post-trading transparency may make more information available to the general public and less analysts research is needed. That is why we observe a drop in the analyst coverage. However, post-trading transparency reveals mainly order flow information. Translating order flow information into firms’ fundamental information which analysts specialize in is difficult. In this case, analyst research should still be needed. Even if we assume that post-trading transparency reveals fundamental information, the transparency story implies that small firms will be more affected: the trading process of thinly traded stocks, which in most cases are small stocks, is more likely to feature opaqueness. The improvement in transparency should benefit small firms more than large firms. This further implies that fewer analysts on small firms are needed because more information on small firms is now available. But we observe the opposite: small firms’ analyst coverage remains unchanged and large firms’ coverage drops. To sum up, it could be the case that other provisions may confound our results but it is unlikely for a provision other than unbundling to explain all our findings.

An alternative story that can explain our findings is learning. Maybe EU analysts are becoming better at learning over the years and begin to produce better forecast results regardless of unbundling. While this alternative is plausible and harder to preclude entirely, it does not fully justify the sudden improvement in the overall forecast quality right after unbundling. First, if it were not for unbundling, the learning process is likely to take more time and the effect should manifest itself gradually over the years. If that were the case, our placebo test should have picked up the learning effect, but we do not observe changes in the forecast quality in our placebo test.  

\[47\] If analysts become better at learning due to unbundling, the story does not contradict our argument: unbundling incentivizes analysts to produce better forecasts. One way to improve their forecasts is to become better at learning.
Second, the change in the cumulative distributions of forecast errors (i.e., Figure 1.4) shows that quality improvement is a general phenomenon. It could be the case that a few analysts are better at learning right after the regulation for reasons unrelated to unbundling, but it is be difficult to argue that most of the analysts happen to become better at learning at the same time, right after the regulation, and for reasons unrelated to unbundling.

We further investigate the learning story in Table 1.14. We introduce the lagged forecast error in our regressions (similar to Model (1.5) and Model (1.6)). The intuition behind is the following: analysts producing larger forecast errors in the previous period could face greater career pressure and become better at learning to improve their performance in the current period, regardless of unbundling. Thus, controlling for lagged forecast error help us to capture the change in the analyst learning process unrelated to unbundling. We then check whether our results are robust to the inclusion of the new control. Column (1) presents the results at the analyst-firm pair level. Column (2) presents the results at the analyst level with brokerage house + year fixed effects. Column (3) presents the results at the analyst level with brokerage house × year fixed effects. After adding the new control, our results still hold and the coefficient of interest remains statistically significant and economically meaningful.\textsuperscript{48} The results also show that larger previous forecast errors do predict smaller current forecast errors, which explains why the magnitudes of the coefficients of $EU \times POST$ generally become slightly smaller. We conclude that learning may contribute but does not explain the whole story: on top of learning, the competition introduced by unbundling still plays an important role in driving the results.

\textbf{1.7.4 Brokerage house level employment}

In addition to research quantity and quality, unbundling may also affect different types of brokerage houses differentially. The differential impact has important implications for the market structure of the sell-side research industry (e.g., maybe the industry will begin to consolidate be-

\textsuperscript{48}For example, in column (2), 0.122\% translates into a 14.91\% decrease in the average forecast error of EU analysts prior to the regulation.
cause certain types of brokerage houses are more affected and could not survive). To explore this heterogeneity, we attempt to compare the employment results of independent research firms (we will call them boutique houses) with the results of brokerage houses affiliated with large investment banks. Boutique houses normally do not provide trading services, while the business model of research in brokerage houses affiliated with large banks features heavy bundling. It is likely that the latter will be more affected by the regulation. However, our data only provides us with an anonymous ID for the brokerage house analysts work for. We are unable to recover the real name of these brokerage houses. To overcome this data limitation, we focus on the size of the brokerage houses, i.e., the number of analysts each brokerage house employs. It is reasonable to assume that boutique houses are smaller in size than houses affiliated with large banks.

To understand which type of houses are more affected by the regulation, we then perform the following analysis within EU brokerage houses:

\[
Y_{bt} = \phi_1 \text{Small}_b \times Post_t + \phi_2 X_{bt} + \alpha_b + \alpha_t + \epsilon_{bt}
\]  

(1.15)

\(Y_{bt}\) is the number of unique analysts working in brokerage house \(b\) in year \(t\). EU brokerage houses are defined to be the brokerage houses that only hire EU analysts. \(\text{Small}_b\) is a dummy equal to 1 if the average number of analysts a brokerage house hires in the pre-regulation years falls below the median. \(Post_t\) is a dummy equal to 1 if the fiscal year \(t\) is equal to 2018. The control \(X_{bt}\) includes standard firm level controls as defined previously. We average the firm level variables within the brokerage house portfolio to get the brokerage house level firm controls. Standard errors are clustered at the brokerage house level.

The results are presented in Table 1.15. Column (1) shows the results within EU brokerage houses: after the regulation, large brokerage houses on average lose around 3 more analysts than small houses. This can also be seen from the plot in Figure 1.10 of Appendix 1.12. Column (2) shows the results of a triple-difference specification similar to Model (1.3). The coefficient in front of the triple interaction term is not significantly different from 0: it seems that US large brokerage houses also lose more analyst during this period.
The results in this section seem to be inconclusive. It is true that within the EU, large brokerage houses are more affected. However, within the US, a region unaffected by the regulation directly, large brokerage houses are also more affected. It is possible that small brokerage houses may not be a good proxy for independent research firms and for this reason, brokerage houses in the US are not valid counterfactuals in the triple-difference analyses. In addition, the number of observations at the brokerage house level is small and the results thus may suffer from a low-power issue. With all these limitations in mind, we leave careful analyses on how MiFID II have changed the landscape of brokerage houses for future research.

1.7.5 Capital market effects

So far, our analyses focus on sell-side research. We now explore the capital market effects of unbundling, or more generally, the capital market effects of MiFID II. Since capital market outcomes are potentially affected by many other factors, the goal in this section is to show what has happened, rather than to apply tight identification strategies and claim causality. To be coherent with our previous analyses, we only focus on firms with valid I/B/E/S coverage in our sample.

Following the previous literature, we focus on two important variables at the firm level: earnings announcement information content ($EAinfo_{jt}$) and bid-ask spread ($Bid\ Ask\ Spread_{jt}$) (e.g., Cumming et al. (2011), Christensen et al. (2016) and Merkley et al. (2017) ). We perform exercises similar to Model (1.1) but replace the dependent variables to be either $EAinfo_{jt}$ or $Bid\ Ask\ Spread_{jt}$.

$EAinfo_{jt}$ is defined as:

$$EAinfo_{jt} = \sum_{d=-1}^{1} |Ret_{j,d} - Ret_{m,d}|$$

(1.16)

where $d$ denotes days around a firm’s earnings announcement date $t$, $j$ denotes the firm, $Ret_{j,d}$

\[49\] Here, we define the market return for EU and US separately by calculating the value-weighted returns of all the firms in a given continent. Our results are qualitatively the same if we use all stocks in each country to calculate the stock market return.
denotes the daily return of firm \( j \) and \( \text{Ret}_{m,d} \) denotes the daily return of the stock market.\(^{49}\) To mitigate data errors in Datastream, we winsorize \( \text{Ret}_{j,d} \) at the 1\% level. \( EAinfo \) captures the informativeness earnings announcements: its value should be close to 0 if analyst forecasts are informative enough and no additional information is conveyed at the actual earnings report date. The daily bid-ask spread is computed as the difference between the two prices divided by the midpoint. We winsorize the spread at the 1\% level. We then take the mean of the daily spread over the year for a given firm and obtain \( \text{Bid Ask Spread}_{jt} \). Bid-ask spread is a common measure of market liquidity. If the overall informativeness of analyst forecast and market liquidity situation do not deteriorate, we anticipate \( \beta \leq 0 \). Results are shown in Table 1.16.

Column (1) and (2) are results for earnings announcement information content. Column (1) shows the standard difference-in-difference result, while column (2) presents the dynamic coefficient estimates. Although the coefficient of interest is positive in column (1), it is not statistically significant. In fact, column (2) reveals that all the coefficients are negative, and the parallel trend assumption may not hold in this regression. The evidence suggests that firms’ earnings announcements do not convey more information after unbundling. Similarly, for bid-ask spread, the overall impact is not statistically different from zero. It seems that bid-ask spread decreased in 2017, one year prior to the regulation. Overall, the results in this table imply that capital market situations do not deteriorate after MiFID II.

These findings do not contradict with the previous studies by Christensen et al. (2016) and Cumming et al. (2011) who show that capital market regulation aiming to improve transparency and competition has overall positive effects on capital market outcomes. The findings are also consistent with our story. Competition among analysts is strengthened after the regulation, resulting in the improvement of forecast quality. This further improves the informativeness of analyst reports and has non-negative effects on the liquidity of the market. However, we are aware that both \( EAinfo_{jd} \) and \( Bid Ask Spread_{jt} \) are noisy aggregate outcomes that can be affected by many other factors in addition to analyst forecasts. Moreover, the reform is fairly recent, and its aggregate impact may take some more time to manifest. The aggregate impact of MiFID II on capital markets...
is an important question and is worth investigating more carefully in future research projects.

1.8 Conclusion

In this paper, we investigate the impact of unbundling on analyst research production. Firms affected by unbundling experienced a more substantial drop in analyst coverage. Such a drop does not come from small- or mid-cap firms but is concentrated in large firms. Furthermore, firms’ coverage quality improves. We argue that the enhancement of analyst competition could explain the results. Unbundling puts an explicit price on analyst research. On the one hand, asset managers become more selective in purchasing research because they need to internalize research costs. On the other hand, analysts are evaluated directly by the services they provide and have stronger incentives to produce better research. As a result, analysts compete more directly in the quality domain to maintain and attract clients. We find evidence of the two specific predictions of the competition channel: 1) analysts who stay produce better research (intensive margin); 2) analysts who produce worse research are more likely to stop working (extensive margin). Our results highlight a potential trade-off between quantity and quality: improvement in the quality of research by unbundling comes at a cost of reducing the quantity of research. Overall, our paper provides a timely investigation of an important and controversial regulation. It helps to shed light on key questions and debates, including how research should be paid for and how the market for analysts should be regulated.

Our paper focuses on one important aspect of information production – the sell-side research production but is silent on other dimensions of information production. For example, the rising cost of sell-side research may encourage more buy-side players to establish their in-house research team. A larger proportion of information production may then migrate from the sell-side to the buy-side. Instead of offering conclusive answers, we view our paper as an important starting point to understand how the compensation schedule for information affects the overall information production. There are many interesting follow-up questions worth exploring: Will the sell-side
research markets begin to consolidate? Will research migrate from the sell-side to the buy-side? What are the welfare implications for investors? We leave these questions for future research.

1.9 Figures

This figure illustrates heuristically the two regimes in which sell-side research is compensated. Under bundling, research is paid through trading commissions and the commissions are passed on to end clients. Under unbundling, research is required to be paid separately. Asset managers can either pay for research directly (against the firm’s profit and loss) or pass the costs to clients by setting up a research payment account (“RPA”) under the consent of clients. Most asset managers opt for purchasing researching directly. The figure also shows that data we use for this project comes from the sell-side and the public firms it covers.
This figure shows the average analyst coverage of EU firms and US firms over time. Analyst coverage is the number of unique analysts covering a specific firm in fiscal year $t$. We report the averages of the period from 2014 to 2018 and adjust the coverage value so that both lines start at 0 in 2014.

This figure shows the average analyst coverage of small firms and large firms in the EU. Analyst coverage is the number of unique analysts covering a specific firm in fiscal year $t$. Small firms are firms whose average fiscal year-end market capitalization before 2018 falls below the median. We report the averages of the period during 2014 to 2018 and adjust the coverage value so that both lines start at 0 in 2014.
This figure plots the empirical cumulative distribution of the forecast errors in the pre- and post-regulation years. To get the forecast errors in the pre-regulation years, we simply average firms’ forecast errors overall all the pre-years.
This figure shows the dynamic effect of unbundling. We split the post dummies into year dummies (Model (1.4)) and plot the coefficient estimates along with a 95% confidence interval. We choose 2014, the year in which MiFID II went into discussion, as our reference year. Standard errors are clustered at the country level.
This figure shows the dynamic effect of unbundling in the intensive margin analyses at the analyst level. We split the post dummies into year dummies and plot coefficient estimates along with a 95% confidence interval. We choose 2014, the year in which MiFID II went into discussion, as our reference year. Standard errors are clustered at the analyst level.
1.10 Tables

Table 1.1: Firm Level Summary Statistics (Sample for Coverage Quantity)

EU Firms: 2133 firms in total

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Mean</th>
<th>Median</th>
<th>Pctl(25)</th>
<th>Pctl(75)</th>
<th>St. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coverage</td>
<td>8.558</td>
<td>4</td>
<td>1</td>
<td>13</td>
<td>9.761</td>
</tr>
<tr>
<td>Book to Market</td>
<td>2.448</td>
<td>1.046</td>
<td>0.595</td>
<td>1.855</td>
<td>5.799</td>
</tr>
<tr>
<td>Investment Return (%)</td>
<td>11.597</td>
<td>7.429</td>
<td>−13.127</td>
<td>28.959</td>
<td>43.462</td>
</tr>
<tr>
<td>Market Capitalization (Bil $)</td>
<td>5.324</td>
<td>0.547</td>
<td>0.114</td>
<td>2.751</td>
<td>17.385</td>
</tr>
<tr>
<td>Return on Assets (%)</td>
<td>2.919</td>
<td>4.574</td>
<td>0.992</td>
<td>8.020</td>
<td>17.484</td>
</tr>
<tr>
<td>Return Volatility (%)</td>
<td>33.330</td>
<td>28.996</td>
<td>22.628</td>
<td>38.431</td>
<td>17.041</td>
</tr>
</tbody>
</table>

US Firms: 2259 firms in total

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Mean</th>
<th>Median</th>
<th>Pctl(25)</th>
<th>Pctl(75)</th>
<th>St. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coverage</td>
<td>11.145</td>
<td>8</td>
<td>4</td>
<td>16</td>
<td>10.162</td>
</tr>
<tr>
<td>Book to Market</td>
<td>1.792</td>
<td>0.814</td>
<td>0.450</td>
<td>1.556</td>
<td>3.228</td>
</tr>
<tr>
<td>Market Capitalization (Bil $)</td>
<td>10.235</td>
<td>1.531</td>
<td>0.388</td>
<td>5.608</td>
<td>38.128</td>
</tr>
<tr>
<td>Return on Assets (%)</td>
<td>2.919</td>
<td>3.954</td>
<td>0.934</td>
<td>7.698</td>
<td>23.676</td>
</tr>
<tr>
<td>Return Volatility (%)</td>
<td>35.025</td>
<td>29.326</td>
<td>22.355</td>
<td>41.698</td>
<td>19.025</td>
</tr>
</tbody>
</table>

This table provides the summary statistics of important firm level variables in the sample for coverage quantity. Coverage is the number of unique analysts covering a certain firm. Book to Market is defined as total asset (Worldscope item 02999) minus long-term debt (Worldscope item 03251) over market value (Worldscope item 08002). Investment return is Worldscope item 08801, measured as (market price at year $t$ + dividends−market price at year $t−1$)/market price at year $t−1$. Market capitalization is Worldscope Item 08002 measured in billion dollars. Return on Assets is Worldscope Item 08326. Return volatility is the annualized standard deviation of daily returns over a year for a given firm.
Table 1.2: Firm Level Summary Statistics (Sample for Coverage Quality)

EU Firms: 1111 firms in total

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Mean</th>
<th>Median</th>
<th>Pctl(25)</th>
<th>Pctl(75)</th>
<th>St. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coverage</td>
<td>13.887</td>
<td>11</td>
<td>5</td>
<td>21</td>
<td>10.163</td>
</tr>
<tr>
<td>Forecast Error (%)</td>
<td>0.724</td>
<td>0.340</td>
<td>0.126</td>
<td>0.848</td>
<td>1.085</td>
</tr>
<tr>
<td>Dispersion (%)</td>
<td>0.848</td>
<td>0.456</td>
<td>0.224</td>
<td>0.960</td>
<td>1.163</td>
</tr>
<tr>
<td>Book to Market</td>
<td>2.415</td>
<td>0.922</td>
<td>0.537</td>
<td>1.564</td>
<td>5.701</td>
</tr>
<tr>
<td>Distance</td>
<td>120.245</td>
<td>115.029</td>
<td>94.026</td>
<td>141.500</td>
<td>43.478</td>
</tr>
<tr>
<td>Investment Return (%)</td>
<td>11.775</td>
<td>9.067</td>
<td>−8.107</td>
<td>27.482</td>
<td>33.974</td>
</tr>
<tr>
<td>Market Capitalization (Bil $)</td>
<td>9.595</td>
<td>2.174</td>
<td>0.692</td>
<td>7.529</td>
<td>23.104</td>
</tr>
<tr>
<td>Return on Assets (%)</td>
<td>6.046</td>
<td>5.374</td>
<td>2.661</td>
<td>8.866</td>
<td>12.240</td>
</tr>
</tbody>
</table>

US Firms: 1693 firms in total

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Mean</th>
<th>Median</th>
<th>Pctl(25)</th>
<th>Pctl(75)</th>
<th>St. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coverage</td>
<td>13.679</td>
<td>11</td>
<td>6</td>
<td>19</td>
<td>10.071</td>
</tr>
<tr>
<td>Forecast Error (%)</td>
<td>0.442</td>
<td>0.176</td>
<td>0.064</td>
<td>0.458</td>
<td>0.801</td>
</tr>
<tr>
<td>Dispersion (%)</td>
<td>0.539</td>
<td>0.218</td>
<td>0.091</td>
<td>0.558</td>
<td>0.930</td>
</tr>
<tr>
<td>Book to Market</td>
<td>1.575</td>
<td>0.732</td>
<td>0.416</td>
<td>1.362</td>
<td>2.398</td>
</tr>
<tr>
<td>Distance</td>
<td>114.431</td>
<td>111.444</td>
<td>90.545</td>
<td>134.000</td>
<td>35.818</td>
</tr>
<tr>
<td>Investment Return (%)</td>
<td>7.821</td>
<td>5.300</td>
<td>−12.515</td>
<td>24.561</td>
<td>35.532</td>
</tr>
<tr>
<td>Market Capitalization (Bil $)</td>
<td>13.356</td>
<td>2.598</td>
<td>0.892</td>
<td>8.674</td>
<td>43.561</td>
</tr>
<tr>
<td>Return on Assets (%)</td>
<td>3.818</td>
<td>4.605</td>
<td>1.289</td>
<td>8.276</td>
<td>12.466</td>
</tr>
<tr>
<td>Return Volatility (%)</td>
<td>31.033</td>
<td>27.152</td>
<td>21.528</td>
<td>36.711</td>
<td>14.279</td>
</tr>
</tbody>
</table>

This table provides the summary statistics of important firm level variables in the sample for coverage quality. To be included in this sample, firms need to be covered by at least 2 analysts each year. Coverage is the number of unique analysts covering a certain firm. Forecast error is defined as the absolute distance between the firm’s actual annual earnings per share and the mean of the analyst forecasts, scaled by the firm’s previous year-end price. Forecast dispersion is defined as the standard deviation of all the forecasts across all the analysts following the same firm in the same year, scaled by the firm’s previous year-end price. Book to Market is defined as total asset (Worldscope item 02999) minus long-term debt (Worldscope item 03251) over market value (Worldscope item 08002). Distance is the average of the days between the analyst forecast date and the firm’s actual EPS report date. We take the average over all the analysts following the same firm and measure this variable in 2014. Investment return is Worldscope item 08801, measured as (market price at year t + dividends – market price at year t – 1)/market price at year t − 1). Market capitalization is Worldscope Item 08002 measured in billion dollars. Return on Assets is Worldscope Item 08326. Return volatility is the annualized standard deviation of daily returns over a year for a given firm. We remove observations for which forecast error and forecast dispersion is larger than 10% of the firm’s share price at the end of the previous year.
Table 1.3: Analyst Level Summary Statistics (Sample for Intensive Margin)

EU Analysts: 1211 in total

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Mean</th>
<th>Median</th>
<th>Pctl(25)</th>
<th>Pctl(75)</th>
<th>St. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Forecast Error (%)</td>
<td>0.812</td>
<td>0.656</td>
<td>0.376</td>
<td>1.091</td>
<td>0.623</td>
</tr>
<tr>
<td>Average Distance</td>
<td>78.751</td>
<td>70.500</td>
<td>41.000</td>
<td>105.000</td>
<td>52.712</td>
</tr>
<tr>
<td>Number of Firms Follows</td>
<td>7.557</td>
<td>7</td>
<td>5</td>
<td>10</td>
<td>3.817</td>
</tr>
<tr>
<td>Tenure (years)</td>
<td>8.915</td>
<td>8</td>
<td>4</td>
<td>14</td>
<td>5.797</td>
</tr>
</tbody>
</table>

US Analysts: 1387 in total

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Mean</th>
<th>Median</th>
<th>Pctl(25)</th>
<th>Pctl(75)</th>
<th>St. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Forecast Error (%)</td>
<td>0.403</td>
<td>0.272</td>
<td>0.156</td>
<td>0.504</td>
<td>0.396</td>
</tr>
<tr>
<td>Average Distance</td>
<td>81.877</td>
<td>84.235</td>
<td>56.845</td>
<td>101.500</td>
<td>38.765</td>
</tr>
<tr>
<td>Average Coverage of Portfolio Firms</td>
<td>21.367</td>
<td>20.250</td>
<td>14.229</td>
<td>27.279</td>
<td>10.047</td>
</tr>
<tr>
<td>Number of Firms Follows</td>
<td>10.547</td>
<td>10</td>
<td>7</td>
<td>14</td>
<td>5.972</td>
</tr>
<tr>
<td>Tenure (years)</td>
<td>8.740</td>
<td>8</td>
<td>4</td>
<td>13</td>
<td>5.432</td>
</tr>
</tbody>
</table>

This table provides the summary statistics of important variables used in the intensive margin analyses at the analyst level. Average Forecast Error is the average of forecast errors (scaled by the firm’s previous year price) across all the firms the analyst covers in a given year. Average Distance is the average of the days between the analyst forecast date and the firm’s actual EPS report date. We take the average across all the firms within an analyst portfolio and measure this variable in 2014. Average Coverage of Portfolio Firms is the average of the analyst coverage of all the firms the analyst follows in 2014. We take the average across all the firms within an analyst portfolio. N of Firms Follows is the number of firms the analyst follows in 2014. Tenure is the total number of years the analyst appeared in I/B/E/S, from 1995 to 2014. EU analysts are the analysts whose portfolios consist of at least 70% of EU stock and US analysts are the analysts whose portfolios consist of at most 30% of EU stocks. We remove analyst-firm pairs for which the forecast error incurred by the analyst are larger than 10% of the firm’s previous year-end price.
Table 1.4: Analyst Level Summary Statistics (Sample for Extensive Margin)

EU Analysts: 1841 in total

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Mean</th>
<th>Median</th>
<th>Pctl(25)</th>
<th>Pctl(75)</th>
<th>St. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Coverage of Portfolio Firms</td>
<td>20.069</td>
<td>21.000</td>
<td>13.562</td>
<td>26.781</td>
<td>8.509</td>
</tr>
<tr>
<td>Number of Firms Follows</td>
<td>9.402</td>
<td>9</td>
<td>6</td>
<td>12</td>
<td>4.799</td>
</tr>
<tr>
<td>Relative Accuracy</td>
<td>51.141</td>
<td>51.619</td>
<td>45.160</td>
<td>57.690</td>
<td>10.616</td>
</tr>
<tr>
<td>Tenure (years)</td>
<td>10.299</td>
<td>9</td>
<td>5</td>
<td>14</td>
<td>5.938</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Unconditional Probability</th>
<th>Stop Covering a Firm</th>
<th>Drop Out</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unconditional Probability (%)</td>
<td>24.156</td>
<td>21.347</td>
</tr>
</tbody>
</table>

This table shows the summary statistics of important variables in the extensive margin analyses at the analyst level. Relative accuracy is a measure capturing the analyst forecast quality and is comparable across different analysts. The higher the number, the more accurate the analyst forecast is. N of Firms Follows is the number of unique firms the analyst follows prior to 2018. Average Coverage of Portfolio Firms is the average of the analyst coverage of all the portfolio firms which the analyst follows. We take the average across all the years prior to 2018. Stop is a dummy variable equal to 1 if the analyst stops covering a firm after the regulation. Drop Out is a dummy equal to 1 if the analyst stops covering all the firms he/she used to cover prior the regulation. Unconditional probability is the simple average of the two dummy variables.
### Table 1.5: One Difference (Pre and Post)

#### Panel A: Coverage (Sample for Quantity)

<table>
<thead>
<tr>
<th></th>
<th>Pre</th>
<th>Post</th>
<th>Diff</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>EU Full</td>
<td>8.740</td>
<td>7.831</td>
<td>-0.909</td>
<td>0</td>
</tr>
<tr>
<td>EU Small</td>
<td>2.350</td>
<td>2.292</td>
<td>-0.058</td>
<td>0.436</td>
</tr>
<tr>
<td>EU Large</td>
<td>15.136</td>
<td>13.375</td>
<td>-1.761</td>
<td>0.090</td>
</tr>
<tr>
<td>US Full</td>
<td>11.219</td>
<td>10.846</td>
<td>-0.373</td>
<td>0.117</td>
</tr>
<tr>
<td>US Small</td>
<td>5.003</td>
<td>4.447</td>
<td>-0.556</td>
<td>0</td>
</tr>
<tr>
<td>US Large</td>
<td>17.441</td>
<td>17.252</td>
<td>-0.189</td>
<td>0.588</td>
</tr>
</tbody>
</table>

#### Panel B: Forecast Error (%)

<table>
<thead>
<tr>
<th></th>
<th>Pre</th>
<th>Post</th>
<th>Diff</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>EU Full</td>
<td>0.742</td>
<td>0.652</td>
<td>-0.090</td>
<td>0.010</td>
</tr>
<tr>
<td>EU Small</td>
<td>0.880</td>
<td>0.800</td>
<td>-0.079</td>
<td>0.157</td>
</tr>
<tr>
<td>EU Large</td>
<td>0.604</td>
<td>0.502</td>
<td>-0.102</td>
<td>0.013</td>
</tr>
<tr>
<td>US Full</td>
<td>0.432</td>
<td>0.482</td>
<td>0.050</td>
<td>0.034</td>
</tr>
<tr>
<td>US Small</td>
<td>0.578</td>
<td>0.656</td>
<td>0.078</td>
<td>0.052</td>
</tr>
<tr>
<td>US Large</td>
<td>0.285</td>
<td>0.307</td>
<td>0.022</td>
<td>0.335</td>
</tr>
</tbody>
</table>

#### Panel C: Forecast Dispersion (%)

<table>
<thead>
<tr>
<th></th>
<th>Pre</th>
<th>Post</th>
<th>Diff</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>EU Full</td>
<td>0.870</td>
<td>0.762</td>
<td>-0.108</td>
<td>0.003</td>
</tr>
<tr>
<td>EU Small</td>
<td>0.920</td>
<td>0.831</td>
<td>-0.089</td>
<td>0.121</td>
</tr>
<tr>
<td>EU Large</td>
<td>0.819</td>
<td>0.693</td>
<td>-0.126</td>
<td>0.004</td>
</tr>
<tr>
<td>US Full</td>
<td>0.527</td>
<td>0.585</td>
<td>0.057</td>
<td>0.034</td>
</tr>
<tr>
<td>US Small</td>
<td>0.677</td>
<td>0.740</td>
<td>0.063</td>
<td>0.157</td>
</tr>
<tr>
<td>US Large</td>
<td>0.378</td>
<td>0.429</td>
<td>0.052</td>
<td>0.084</td>
</tr>
</tbody>
</table>

#### Panel D: Analyst Forecast Error (%)

<table>
<thead>
<tr>
<th></th>
<th>Pre</th>
<th>Post</th>
<th>Diff</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>EU Full</td>
<td>0.818</td>
<td>0.790</td>
<td>-0.028</td>
<td>0.320</td>
</tr>
<tr>
<td>EU Small</td>
<td>0.869</td>
<td>0.878</td>
<td>0.009</td>
<td>0.848</td>
</tr>
<tr>
<td>EU Large</td>
<td>0.767</td>
<td>0.701</td>
<td>-0.065</td>
<td>0.028</td>
</tr>
<tr>
<td>US Full</td>
<td>0.383</td>
<td>0.485</td>
<td>0.102</td>
<td>0</td>
</tr>
<tr>
<td>US Small</td>
<td>0.477</td>
<td>0.627</td>
<td>0.149</td>
<td>0</td>
</tr>
<tr>
<td>US Large</td>
<td>0.288</td>
<td>0.343</td>
<td>0.055</td>
<td>0.001</td>
</tr>
</tbody>
</table>

This table provides the average values of Coverage, Forecast Error and Forecast Dispersion and Average Forecast Error (Analyst Level) in the pre- and post-years (one difference). Pre-years include year 2014, 2015, 2016, 2017 while post-year is 2018. Panel (a) to (c) show the firm level outcomes. “Full” denotes the results for all the firms. “Small” denotes the results for small firms. “Large” denotes the results for large firms. Small firms are firms whose average fiscal year-end market capitalization over the pre-regulation years falls below the median. To maintain the proportion of EU and US firms fixed, we calculate the cutoff separately in both regions. Panel (d) shows the analyst level outcomes in the sample for intensive margin analyses. “Full” denotes the results for all the analysts. “Small” denotes the results for small analysts. “Large” denotes the results for large analysts. To define small analysts, we first calculate the average market capitalization of all the firms within an analyst portfolio in a given year. We then average the average market capitalization across all the pre-regulation years. Small analysts are defined as analysts whose average of the portfolio average market capitalization over the pre-regulation years falls below the median. To maintain the proposition of analysts fixed, we calculate the median cutoff separately for EU analysts and US analysts.
<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>Coverage</th>
<th>Forecast Error (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Full (1)</td>
<td>Small vs Large (2)</td>
</tr>
<tr>
<td>EU × POST</td>
<td>−0.651*** (0.196)</td>
<td>−1.594*** (0.268)</td>
</tr>
<tr>
<td>EU × POST × SMALL</td>
<td>1.891*** (0.200)</td>
<td>0.002</td>
</tr>
<tr>
<td>SMALL × POST</td>
<td>1.797*** (0.212)</td>
<td>1.542*** (0.228)</td>
</tr>
<tr>
<td>LN SIZE</td>
<td>1.797*** (0.072)</td>
<td>0.841*** (0.146)</td>
</tr>
<tr>
<td>LN BM</td>
<td>0.003*** (0.001)</td>
<td>−0.004* (0.002)</td>
</tr>
<tr>
<td>GDP GROWTH</td>
<td>−0.062* (0.033)</td>
<td>−0.049* (0.027)</td>
</tr>
<tr>
<td>GDP GROWTH</td>
<td>0.142 (0.108)</td>
<td>0.124 (0.092)</td>
</tr>
</tbody>
</table>

| Firm FE Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year FE Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 21,960 | 10,665 | 21,960 | 14,020 | 5,555 | 14,020 |
| R²          | 0.967 | 0.971 | 0.967 | 0.489 | 0.497 | 0.489 |
| Adjusted R² | 0.958 | 0.964 | 0.959 | 0.361 | 0.369 | 0.361 |

Note: * p<0.1; ** p<0.05; *** p<0.01

This table shows the results of the panel regressions specified in Model (1.1) and Model (1.3). The dependent variable in column (1) to (3) is Coverage, the number of unique analysts covering a specific firm j in fiscal year t. The dependent variable in column (4) to (6) is Forecast Error, defined as the absolute distance between the firm’s actual EPS and the mean of the analyst forecasts, scaled by the firm’s previous year-end price. EU is a dummy variable equal to one if the firm is domiciled and listed in Europe. POST is a dummy variable equal to one if the fiscal year t is equal to 2018. SMALL is a dummy variable equal to one if firms’ average fiscal year-end market capitalization over the pre-regulation years falls below the median. To maintain the proportion of EU and US firms fixed, we calculate the cutoff separately in both regions. Column (1) and column (4) are the results for the difference-in-difference regression between EU firms and US firms. Column (2) and column (5) are the results for the difference-in-difference regression between small firms and larger firms within the EU. Column (3) and Column (6) are the results for the triple-difference regression. Standard errors are clustered at the country level.
<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>Forecast Error (Analyst) (%)</th>
<th>Average Forecast Error (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Full</td>
<td>Triple Diff</td>
</tr>
<tr>
<td>EU × POST</td>
<td>-0.129***</td>
<td>-0.122***</td>
</tr>
<tr>
<td></td>
<td>(0.032)</td>
<td>(0.040)</td>
</tr>
<tr>
<td>EU × POST × SMALL</td>
<td>-0.017</td>
<td>-0.035</td>
</tr>
<tr>
<td></td>
<td>(0.049)</td>
<td>(0.062)</td>
</tr>
<tr>
<td>SMALL × POST</td>
<td>0.031</td>
<td>0.048</td>
</tr>
<tr>
<td></td>
<td>(0.025)</td>
<td>(0.055)</td>
</tr>
<tr>
<td>LN FIRMS COVERED × POST</td>
<td>-0.014</td>
<td>-0.013</td>
</tr>
<tr>
<td></td>
<td>(0.020)</td>
<td>(0.020)</td>
</tr>
<tr>
<td>LN COVERAGE × POST</td>
<td>-0.023</td>
<td>-0.015</td>
</tr>
<tr>
<td></td>
<td>(0.019)</td>
<td>(0.021)</td>
</tr>
<tr>
<td>LN TENURE × POST</td>
<td>-0.005</td>
<td>-0.004</td>
</tr>
<tr>
<td></td>
<td>(0.012)</td>
<td>(0.012)</td>
</tr>
<tr>
<td>LN DISTANCE × POST</td>
<td>-0.018**</td>
<td>-0.019**</td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td>(0.008)</td>
</tr>
</tbody>
</table>

| Firm Level Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Firm FE            | Yes | Yes | Yes | Yes | Yes | Yes |
| B House FE         | Yes | Yes | Yes | Yes | No  | No  |
| Year FE            | Yes | Yes | Yes | Yes | No  | No  |
| B House FE × Year FE | No  | No  | No  | No  | Yes | Yes |

Observations: 81,445 81,445 12,990 12,990 4,111 4,111

R²: 0.422 0.422 0.559 0.559 0.597 0.598

Adjusted R²: 0.377 0.377 0.431 0.431 0.480 0.480

*Note: Δp<0.1; **p<0.05; ***p<0.01

This table shows the results of the panel regressions specified in Model (1.5) and Model (1.6). We focus on analyst-firm pairs that survive throughout the sample period (2014 to 2018). The dependent variable in column (1) and (2) is Forecast Error, defined as the absolute distance between the firm’s actual EPS and the analyst’s forecast, scaled by the firm’s previous year-end price. The dependent variable in column (3) to (6) is Average Forecast Error, defined as the average of forecast errors across all the firms the analyst covers in a given year. EU is a dummy variable equal to one if the analyst’s portfolio consists of at least 70% of EU stock and zero if the analyst’s portfolio consists of at most 30% of EU stocks. POST is a dummy variable equal to one if the fiscal year t is equal to 2018. SMALL is a dummy variable equal to one for the analyst whose average of the portfolio average market capitalization over the pre-regulation years falls below the median. We take the average first over all the firms the analyst covers each year and then over all the pre-regulation years. To maintain the proposition of analysts fixed, we calculate the median cutoff separately for EU analysts and US analysts. Column (1) and column (2) are the results at the analyst-firm pair level. Column (3) to column (6) are the results at the analyst level. Standard errors are clustered at the firm level in column (1) and column (2), and at the analyst level column (3) to column (6).
This table shows the results of the regressions specified in Model (1.8) and Model (1.9). The dependent variable in column (1) and (2) is *Stop*, a dummy variable equal to 1 if the analyst stops covering a firm after the regulation. The dependent variable in column (3) and (4) is *DropOut*, a dummy variable equal to 1 if the analyst stops covering all the firms he used to cover prior to the regulation. \( \text{RELATIVE ACCURACY} = \frac{1}{T} \sum_{t}^{T} \left( \frac{1}{J} \sum_{j}^{J} \text{Score}_{ijt} \right) \), where \( J \) is the set of firms analyst \( i \) covers in year \( t \) and \( T \) includes all the years prior to 2018. It is a measure capturing the analyst forecast quality and is comparable across different analysts. The higher the number, the more accurate the analyst forecast is. Column (1) and (2) shows the results for the analyst-firm level analyses and column (3) and (4) the analyst level analyses. Standard errors are clustered at the firm level in column (1) and (2), and at the brokerage house level in column (3) and (4).
Table 1.9: Analyst Level Outcomes (Extensive Margin). Probability of Stop or Dropping Out Before and After Unbundling

<table>
<thead>
<tr>
<th></th>
<th>Dependent variable:</th>
<th>EU</th>
<th>US</th>
<th>Pooled</th>
<th>EU</th>
<th>US</th>
<th>Pooled</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Stop Covering a Firm</td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
</tr>
<tr>
<td>RELATIVE ACCURACY × POST</td>
<td>-0.001***</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>-0.001**</td>
<td>-0.0005</td>
<td>-0.0005</td>
</tr>
<tr>
<td></td>
<td>(0.0003)</td>
<td>(0.0002)</td>
<td>(0.0002)</td>
<td>(0.0002)</td>
<td>(0.0005)</td>
<td>(0.001)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>RELATIVE ACCURACY × POST × EU</td>
<td>-0.001***</td>
<td></td>
<td>-0.001</td>
<td></td>
<td>-0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0004)</td>
<td></td>
<td>(0.001)</td>
<td></td>
<td>(0.001)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>POST × EU</td>
<td>0.058***</td>
<td></td>
<td></td>
<td>0.021</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.020)</td>
<td></td>
<td></td>
<td>(0.046)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EU × RELATIVE ACCURACY</td>
<td>0.006***</td>
<td></td>
<td></td>
<td>0.004***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0002)</td>
<td></td>
<td></td>
<td>(0.0005)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RELATIVE ACCURACY</td>
<td>-0.005***</td>
<td>-0.011***</td>
<td>-0.011***</td>
<td>-0.011***</td>
<td>-0.003***</td>
<td>-0.007***</td>
<td>-0.007***</td>
</tr>
<tr>
<td></td>
<td>(0.0002)</td>
<td>(0.0001)</td>
<td>(0.0001)</td>
<td>(0.0001)</td>
<td>(0.0003)</td>
<td>(0.0004)</td>
<td>(0.0003)</td>
</tr>
<tr>
<td>EU</td>
<td>-0.293***</td>
<td></td>
<td></td>
<td>-0.198***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.016)</td>
<td></td>
<td></td>
<td>(0.037)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LN TENURE</td>
<td>-0.023***</td>
<td>0.001</td>
<td>-0.011***</td>
<td>0.018***</td>
<td>0.007</td>
<td>0.013***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.001)</td>
<td>(0.005)</td>
<td>(0.005)</td>
<td>(0.004)</td>
<td></td>
</tr>
</tbody>
</table>

|                          | Drop Out                                  | (4)      | (5)      | (6)      |          |          |          |
| RELATIVE ACCURACY        | -0.005***                                 |          |          |          | -0.005   | -0.0005  | -0.0005  |
| (0.0003)                 | (0.0002)                                  | (0.0002) | (0.0002) | (0.0002) | (0.0005) | (0.001)  | (0.001)  |

| Controls                 | Yes                                       | Yes      | Yes      | Yes      | Yes      | Yes      | Yes      |
| Firm FE                  | Yes                                       | Yes      | Yes      | Yes      | No       | No       | No       |
| B House FE               | No                                        | No       | No       | Yes      | Yes      | Yes      | Yes      |
| Year FE                  | Yes                                       | Yes      | Yes      | Yes      | Yes      | Yes      | Yes      |
| Observations             | 68,373                                    | 96,974   | 165,347  | 11,498   | 10,800   | 22,298   |
| R²                       | 0.112                                     | 0.230    | 0.174    | 0.237    | 0.333    | 0.274    |
| Adjusted R²              | 0.083                                     | 0.211    | 0.152    | 0.221    | 0.311    | 0.256    |

Note: *p<0.1; **p<0.05; ***p<0.01

This table shows the results of the regressions specified in Model (1.10) and Model (1.11). The dependent variable in column (1) to (3) is \( Stop_{ijt+1} \), a dummy variable equal to 1 if the analyst \( i \) covering firm \( j \) in year \( t \) stops covering the same firm in year \( t + 1 \). The dependent variable in column (3) to (6) is \( DropOut_{ijt+1} \), a dummy variable equal to 1 if the analyst stops covering all the firms in year \( t + 1 \). RELATIVE ACCURACY = \( \frac{1}{J} \sum_{J} Score_{ijt} \), where \( J \) is the set of firms analyst \( i \) covers in year \( t \). It is a measure capturing the analyst forecast quality and is comparable across different analysts. The higher the number, the more accurate the analyst forecast is. POST is a dummy equal to 1 for the year 2017. Column (1) to (3) show the results for the analyst-firm level analyses. Column (1) presents the results for EU analysts; column (2) for US analysts and column (3) for a triple-difference analysis. Column (4) to (6) show similar results for the analyst level analyses. Column (4) presents the results for EU analysts; column (5) for the US analysts and column (6) for a triple-difference analysis. Standard errors are clustered at the firm level in column (1) to (3), and at the brokerage house level in column (4) to (6).
Panel A: Firm Level

<table>
<thead>
<tr>
<th>Coverage</th>
<th>Forecast Error (%)</th>
<th>Dispersion (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>EU × POST</td>
<td>−0.071</td>
<td>−0.177</td>
</tr>
<tr>
<td></td>
<td>(0.113)</td>
<td>(0.138)</td>
</tr>
<tr>
<td>Controls</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Firm FE</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Year FE</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>17,568</td>
<td>17,568</td>
</tr>
<tr>
<td>R²</td>
<td>0.974</td>
<td>0.974</td>
</tr>
<tr>
<td>Adjusted R²</td>
<td>0.965</td>
<td>0.965</td>
</tr>
</tbody>
</table>

Note: *p<0.1; **p<0.05; ***p<0.01

Panel B: Analyst Level

<table>
<thead>
<tr>
<th>Average Forecast Error (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2015</td>
</tr>
<tr>
<td>(1)</td>
</tr>
<tr>
<td>EU × POST</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Firm Level Controls</td>
</tr>
<tr>
<td>Analyst Level Controls × Post</td>
</tr>
<tr>
<td>Analyst FE</td>
</tr>
<tr>
<td>B House FE</td>
</tr>
<tr>
<td>Year FE</td>
</tr>
<tr>
<td>Observations</td>
</tr>
<tr>
<td>R²</td>
</tr>
<tr>
<td>Adjusted R²</td>
</tr>
</tbody>
</table>

Note: *p<0.1; **p<0.05; ***p<0.01

Panel A of this table shows the results of the panel regressions similar to Model (1.1) but only covers the pre-regulation years. The dependent variable in column (1) to (3) is Coverage, the number of unique analysts covering a specific firm j in fiscal year t. The dependent variable in column (4) to (6) is Forecast Error, defined as the absolute distance between the firm’s actual EPS and the mean of the analyst forecasts, scaled by the firm’s previous year-end price. The dependent variable in column (4) to (6) is Forecast Dispersion, defined as the standard deviation of all the forecasts across all the analysts following the same firm in the same year, scaled by the firm’s previous year-end price. EU is a dummy variable equal to one if the firm is domiciled and listed in Europe. From column (1) to column (3), POST is a dummy variable defined as if the regulation occurred in 2015, 2016, 2017, respectively. Similar definition applies to column (4) to (6) and column (7) to (9). Column (1) to (3) are the results for Coverage. Column (4) to (6) are the results for Forecast Error. Column (7) to (9) are the results for Forecast Dispersion. Standard errors are clustered at the country level. Panel B of this table shows the results of the panel regressions similar to Model (1.6). The dependent variable in column (3) to (6) is Average Forecast Error, defined as the average of forecast errors over all the firms the analyst covers in a given year. EU is a dummy variable equal to one if the analyst’s portfolio consists of at least 70% of EU stock and zero if the analyst’s portfolio consists of at most 30% of EU stocks. From column (1) to column (3), POST is a dummy variable defined as if the regulation occurred in 2015, 2016, 2017, respectively. Standard errors are clustered at the analyst level.


<table>
<thead>
<tr>
<th></th>
<th>Dependent variable:</th>
<th>Brokerage Houses per Firm</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Full vs Large</td>
<td>Triple Diff</td>
</tr>
<tr>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>EU × POST</td>
<td>−0.403***</td>
<td>−1.081***</td>
</tr>
<tr>
<td></td>
<td>(0.151)</td>
<td>(0.207)</td>
</tr>
<tr>
<td>EU × POST × SMALL</td>
<td>1.359***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.154)</td>
<td></td>
</tr>
<tr>
<td>SMALL × POST</td>
<td>1.275***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.157)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>−0.099***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.019)</td>
<td></td>
</tr>
</tbody>
</table>

Firm FE Yes Yes Yes
Year FE Yes Yes Yes
Observations 21,960 10,665 21,960
R² 0.972 0.978 0.972
Adjusted R² 0.965 0.972 0.965

Note: *p<0.1; **p<0.05; ***p<0.01

This table shows the results of the panel regressions similar Model (1.1) on the sample for coverage quality. The dependent variable is Brokerage Houses per Firm, the number of unique brokerage houses covering a specific firm in fiscal year $t$. EU is a dummy variable equal to one if the firm is domiciled and listed in Europe. POST is a dummy variable equal to one if the fiscal year $t$ is equal to 2018. SMALL is a dummy variable equal to one if firms’ average fiscal year-end market capitalization over the pre-regulation years falls below the median. To maintain the proportion of EU and Column (1) is the result of the difference-in-difference regression between EU firms and US firms. Column (2) is the result of the difference-in-difference regression between small firms and larger firms within the EU. Column (3) is the result of the triple-difference regression. Standard errors are clustered at the country level.
Table 1.12: Other Measures of Quality

Panel A: Firm Level Dispersion

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>Dispersion (%)</th>
<th>Full (1)</th>
<th>Small vs Large (2)</th>
<th>Triple Diff (3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EU × POST</td>
<td>-0.166***</td>
<td></td>
<td></td>
<td>-0.195***</td>
</tr>
<tr>
<td></td>
<td>(0.024)</td>
<td></td>
<td></td>
<td>(0.042)</td>
</tr>
<tr>
<td>EU × POST × SMALL</td>
<td>0.058</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SMALL × POST</td>
<td>0.048</td>
<td></td>
<td>-0.015**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.066)</td>
<td></td>
<td>(0.007)</td>
<td></td>
</tr>
</tbody>
</table>

Controls: Yes, Yes, Yes
Firm FE: Yes, Yes, Yes
Year FE: Yes, Yes, Yes
Observations: 14,020, 5,555, 14,020
R²: 0.561, 0.550, 0.561
Adjusted R²: 0.451, 0.436, 0.451

Note: ∗p<0.1; ∗∗p<0.05; ∗∗∗p<0.01

Panel B: Number of Firms per Analyst

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>Firms Covered</th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EU × POST</td>
<td>-0.791***</td>
<td>-0.705***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.130)</td>
<td>(0.110)</td>
<td></td>
</tr>
</tbody>
</table>

Firm Level Controls: Yes, Yes
Analyst Level Controls × Post: Yes, Yes
Analyst FE: Yes, Yes
Brokerage House FE: Yes, Yes
Year FE: Yes, Yes
Observations: 14,435, 26,555
R²: 0.875, 0.865
Adjusted R²: 0.839, 0.805

Note: ∗p<0.1; ∗∗p<0.05; ∗∗∗p<0.01

Panel A of this table shows the results of the panel regressions specified in Model (1.1) and Model (1.3). The dependent variable is Forecast Dispersion, defined as the standard deviation of all the forecasts across all the analysts following the same firm in the same year, scaled by the firm’s previous year-end price. EU is a dummy variable equal to one if the firm is domiciled and listed in Europe. POST is a dummy variable equal to one if the fiscal year t is equal to 2018. SMALL is a dummy variable equal to one if firms’ average fiscal year-end market capitalization over the pre-regulation years falls below the median. To maintain the proportion of EU and US firms fixed, we calculate the cutoff separately in both regions. Column (1) is the result for the difference-in-difference regression between EU firms and US firms. Column (2) is the result for the difference-in-difference regression between small firms and larger firms within the EU. Column (3) is the result for the triple-difference regression. Standard errors are clustered at the country level. Panel B of this table shows the results of the panel regressions specified in Model (1.6). The dependent variable is Firms Covered, the number of firms one analyst covers in fiscal year t. EU is a dummy variable equal to one if the analyst’s portfolio consists of at least 70% of EU stock and zero if the analyst’s portfolio consists of at most 30% of EU stocks. POST is a dummy variable equal to one if the fiscal year t is equal to 2018. Column (1) is results for analysts who cover at least one firm throughout the sample period. Column (2) is the result of all the analysts in our sample. Standard errors are clustered at the analyst level.
Table 1.13: Abnormal Return and Analyst Informativeness

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>Absolute Abnormal Return</th>
<th>Average Analyst Informativeness</th>
<th>Aggregate Analyst Informativeness</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>I(EU *post)</td>
<td>0.001***</td>
<td>0.0003***</td>
<td>−0.009***</td>
</tr>
<tr>
<td></td>
<td>(0.0002)</td>
<td>(0.0001)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>Controls</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Firm FE</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Analyst × Firm FE</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Year FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>550,693</td>
<td>19,391</td>
<td>19,391</td>
</tr>
<tr>
<td>R²</td>
<td>0.319</td>
<td>0.496</td>
<td>0.917</td>
</tr>
<tr>
<td>Adjusted R²</td>
<td>0.231</td>
<td>0.353</td>
<td>0.894</td>
</tr>
</tbody>
</table>

Note: *p<0.1; **p<0.05; ***p<0.01

This table shows the results of the panel regressions similar to Model (1.1). The dependent variable in column (1) is \( ABRet \) (Absolute Abnormal Return), the absolute market-adjusted abnormal return on the forecast revision date for each analyst-firm pair. The dependent variable in column (2) is \( AVGAI \) (Average Analyst Informativeness). It captures the average informativeness of one forecast revision date. The dependent variable in column (3) is \( AGAI \) (Aggregate Analyst Informativeness). It captures the aggregate informativeness of all revision dates. \( EU \) is a dummy variable equal to one if the firm is domiciled and listed in Europe. \( POST \) is a dummy variable equal to one if the fiscal year \( t \) is equal to 2018. Standard errors are clustered at the firm level.
Table 1.14: Learning Effect (Lagged Forecast Error as Control)

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>Forecast Error (Analyst) (%)</th>
<th>Average Forecast Error (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>EU × POST</td>
<td>−0.134***</td>
<td>−0.122***</td>
</tr>
<tr>
<td></td>
<td>(0.033)</td>
<td>(0.035)</td>
</tr>
<tr>
<td>LAGGED ERROR (PAIR)</td>
<td>−0.071***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.013)</td>
<td></td>
</tr>
<tr>
<td>LAGGED ERROR</td>
<td>−0.142***</td>
<td>−0.085</td>
</tr>
<tr>
<td></td>
<td>(0.039)</td>
<td>(0.097)</td>
</tr>
<tr>
<td>Firm Level Controls</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Analyst Level Controls × Post</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Firm FE</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Analyst FE</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>B House FE</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Year FE</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>B House FE × Year FE</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Observations</td>
<td>65,156</td>
<td>10,392</td>
</tr>
<tr>
<td>R²</td>
<td>0.447</td>
<td>0.600</td>
</tr>
<tr>
<td>Adjusted R²</td>
<td>0.393</td>
<td>0.446</td>
</tr>
</tbody>
</table>

Note: *p<0.1; **p<0.05; ***p<0.01

This table shows the results of the panel regressions specified in Model (1.5) and Model (1.6). The dependent variable in column (1) is Forecast Error, defined as the absolute distance between the firm’s actual EPS and the analyst’s forecast, scaled by the firm’s previous year-end price. The dependent variable in column (2) to (3) is Average Forecast Error, defined as the average of forecast errors across all the firms the analyst covers in a given year. EU is a dummy variable equal to one if the analyst’s portfolio consists of at least 70% of EU stock and zero if the analyst’s portfolio consists of at most 30% of EU stocks. POST is a dummy variable equal to one if the fiscal year t is equal to 2018. We include lagged forecast error as a control for the analyst learning ability. LAGGED ERROR (PAIR) is the lagged forecast error at the analyst-firm pair level. LAGGED ERROR is the average lagged forecast error at the analyst level. We take the average over all the firms the analyst covers. Column (1) presents the results on the analyst-firm pair level. Column (2) presents the results on the analyst level with brokerage house plus year fixed effects. Column (3) presents the results on the analyst level with brokerage house times year fixed effects. Standard errors are clustered at the firm level in column (1) and at the analyst level in column (2) and (3).
Table 1.15: Brokerage House Employment

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>Number of Analysts</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Small vs Large</td>
</tr>
<tr>
<td></td>
<td>(1)</td>
</tr>
<tr>
<td>SMALL $\times$ POST</td>
<td>3.490$^{**}$</td>
</tr>
<tr>
<td></td>
<td>(1.467)</td>
</tr>
<tr>
<td>EU $\times$ POST $\times$ SMALL</td>
<td>$-0.543$</td>
</tr>
<tr>
<td></td>
<td>(1.587)</td>
</tr>
</tbody>
</table>

Controls Yes Yes
B House FE Yes Yes
Year FE Yes Yes
Observations 545 1,185
R$^2$ 0.953 0.953
Adjusted R$^2$ 0.940 0.941

Note: *$p<0.1$; **$p<0.05$; ***$p<0.01$

This table shows the results of the panel regressions specified in Model (1.15). The dependent variable is *Number of Analysts Hired*, defined as the number of unique analysts each brokerage house hires in each year. EU is a dummy variable equal to one if the brokerage house only hires EU analysts in the years prior to the regulation. POST is a dummy variable equal to one if the fiscal year $t$ is equal to 2018. SMALL is a dummy variable equal to one if the average of the number of analysts a brokerage house hires over the pre-regulation years falls below the median. To maintain the proportion of EU and US brokerage houses fixed, we calculate the cutoff separately in both regions. Column (1) is the results for a difference-in-difference regression within the EU between small brokerage houses and large brokerage houses. Column (2) is the result of the triple-difference regression. Standard errors are clustered at the brokerage house level.
Table 1.16: Capital Market Effects

<table>
<thead>
<tr>
<th></th>
<th>Dependent variable:</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EAinfo</td>
<td>Bid Ask Spread</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>EU × POST</td>
<td>0.002</td>
<td>0.0001</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
<td></td>
</tr>
<tr>
<td>EU × Y15</td>
<td>−0.005***</td>
<td></td>
<td>0.0005</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td></td>
<td>(0.001)</td>
</tr>
<tr>
<td>EU × Y16</td>
<td>−0.004**</td>
<td></td>
<td>0.0002</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td></td>
<td>(0.003)</td>
</tr>
<tr>
<td>EU × Y17</td>
<td>−0.004**</td>
<td></td>
<td>−0.001**</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td></td>
<td>(0.005)</td>
</tr>
<tr>
<td>EU × Y18</td>
<td>−0.001</td>
<td></td>
<td>−0.00005</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td></td>
<td>(0.001)</td>
</tr>
</tbody>
</table>

|                  | Controls Yes | Yes | Yes | Yes |
|                  | Firm FE Yes | Yes | Yes | Yes |
|                  | Year FE Yes | Yes | Yes | Yes |
| Observations     | 20,322      | 20,322| 20,322| 20,322|
| R²               | 0.453       | 0.454| 0.925| 0.925|
| Adjusted R²      | 0.304       | 0.305| 0.904| 0.904|

Note: *p<0.1; **p<0.05; ***p<0.01

This table shows the results of the panel regressions specified in Model (1.1). The dependent variable in column (1) and column (2) is $E_{Ainfo_{jt}}$. $E_{Ainfo_{jt}}$ is defined as $E_{Ainfo_{jt}} = \sum_{d=-1}^{1} |Ret_{j,d} - Ret_{m,d}|$, where $d$ denotes days around a firm’s earnings announcement date $t$, $j$ denotes the firm, $Ret_{j,d}$ denotes the daily return of firm $j$ and $Ret_{m,d}$ denotes the daily return of the stock market. To mitigate data errors in Datastream, we winsorize $Ret_{j,d}$ at the 1% level. The dependent variable in column (3) and column (4) is $Bid Ask Spread_{jt}$. The daily bid-ask spread is computed as the difference between the two prices divided by the midpoint. We winsorize the spread at the 1% level. We then take the mean of the daily spread over the year for a given firm and obtain $Bid Ask Spread_{jt}$. $EU$ is a dummy variable equal to one if the firm is domiciled and listed in Europe. $POST$ is a dummy variable equal to one if the calendar year $t$ is equal to 2018. $Y_{15}$, $Y_{16}$, $Y_{17}$, $Y_{18}$ are year dummy variables that are ones if the calendar year $t$ is equal to 2015, 2016, 2017, 2018. Column (1) to column (2) are results for $E_{Ainfo_{jt}}$. Column (3) to column (4) are results for $Bid Ask Spread_{jt}$. Standard errors are clustered at the country level.
Appendix of Chapter 1

1.11 Data construction

1.11.1 Firm level observations

We restrict our attention to publicly traded firms in Europe and the US. To achieve this, we require firms’ primary quotes of major shares to be listed in one of the major stock exchanges in Belgium, Denmark, Finland, France, Germany, Iceland, Ireland, Italy, the Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, UK, and US.

To be more specific, from Datastream we select stocks with type Equity whose primary quotes of major shares are listed in the following stock exchange markets: Euronext Brussel, OMX Nordic Exchange Copenhagen, Helsinki, Euronext Paris, Berlin, Deutsche Boerse, Munich, NASDAQ OMX Iceland, Dublin, Milan, Euronext Amsterdam, Oslo Bors, Lisbon, Madrid, Mercado Continuo, Stockholm, Six Swiss, London Stock Exchange, NYSE, NASDAQ (US) and Amex. We further delete firms whose main listing place is the OTC market. The European stock exchange markets we select consist of the 18 largest stock exchange markets in Europe. Notice that Switzerland is not part of EEA but given the integrity of the European financial market and the fact that Switzerland will have to amend its current legislation to ensure that Swiss financial institutions to have unfettered access to the EU market, we include Switzerland in our sample. Our results are robust if we exclude Switzerland. We also require firms listed in Europe to be domiciled in one of these European countries and firms listed in the US to be domiciled in the US. We further select firms whose financial documents are presented in one of these countries’ local currencies: Denmark Krone, Euro, Great Britain Pound, Iceland Krona, Swedish Krona, Swiss Franc and USD.

To be included in our final sample, firms need to have non-empty total assets and positive

\[ \text{For example, Ambac financial group went into bankruptcy in 2010 and emerged back in 2013. Its total investment return in 2013 is 170455.55\%. This is simply because Worldscope assigned the price of this firm from OTC the market in 2012 to be } \$0.014 \text{ and the publicly traded price in 2013 is } \$24.56. \]
book-to-market ratio. In some cases, Worldscope records information of firms that went into re-organization and stopped trading. In these cases, extreme values appear in total investment return (RET, Worldscope Item 08801).\textsuperscript{50} We delete observations with total investment return above 99\% or below 1\% of the total distribution of this item. Other variables used in our study are: the log of total market capitalization at fiscal year $t$ in dollars (LNSIZE), the log of book to market ratio (LNBM), return on assets (ROA) and price volatility (PRICEVOL). The precise construction of all these variables can be found in Appendix 1.11.3. Table 1.17 of Appendix 1.13 reports the number of firms, average coverage, average GDP growth rate and average unemployment rate in each country in our sample.

1.11.2 Choice of data frequency

Following literature, we use annual forecast instead of quarterly forecast data for several reasons (e.g., Hong and Kacperczyk (2010) and Giroud and Mueller (2011)). First, analyst forecasts exhibit strong frequency seasonality. For example, in the US, most of the new forecasts are clustered at the beginning of the year. Such seasonality varies across countries and even across firms. It is difficult to attribute changes in quarterly forecast outcomes to MiFID II rather than to seasonality. Second, quarterly financial reports are not mandatory. Though quarterly filing is common in the US, it is uncommon in Europe. Data coverage and quality on a quarterly basis are poor, especially for EU firms. More importantly, since quarterly filings are not mandatory and uncommon in Europe, focusing on firms with quarterly financial reports and earnings per share (EPS) data suffers from severe selection bias: firms who report quarterly inherently differs from firms who do not. Finally, since financial reports normally release with a certain lag, analyst information set will not be updated timely when they issue quarterly forecasts, i.e., it is likely that they do not observe firms’ last quarter true earnings when they make quarterly forecasts. Consequently, staleness is going to be more severe in quarterly forecasts.
### 1.11.3 Variable definitions

<table>
<thead>
<tr>
<th>Variable Definition</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Coverage of Portfolio Firms</td>
<td>The average of the analyst coverage of all the portfolio firms that the analyst covers in a given year.</td>
</tr>
<tr>
<td>Average Forecast Error</td>
<td>The average of forecast errors over all the firms the analyst covers in a given year.</td>
</tr>
<tr>
<td>Book to Market (BM)</td>
<td>Book to Market is defined as total asset (Worldscope item 02999) minus long-term debt (Worldscope item 03251) over market value (Worldscope item 08002) (e.g., Fama and French (1995)).</td>
</tr>
<tr>
<td>Coverage</td>
<td>Coverage is the number of unique analysts who produce forecasts for a certain firm during its fiscal year. Analyst forecast information is obtained from I/B/E/S data set.</td>
</tr>
<tr>
<td>Distance (DISTANCE)</td>
<td>Number of days between the analyst forecast date and the firm’s actual earnings report date.</td>
</tr>
<tr>
<td>Drop Out</td>
<td>A dummy variable equal to 1 if the analyst stops covering all the firms he used to cover prior to the regulation.</td>
</tr>
<tr>
<td>Forecast Dispersion</td>
<td>Forecast dispersion is defined as the standard deviation of all the forecasts over all the analysts following the same firm in the same year, scaled by the firm’s previous year-end price. Analyst forecast information is obtained from the I/B/E/S data set.</td>
</tr>
<tr>
<td>Forecast Error</td>
<td>Forecast error is defined as the absolute distance between the firm’s actual annual earnings per share and the mean of the analyst forecasts, scaled by the firm’s previous year-end price. Analyst forecast information is obtained from I/B/E/S data set.</td>
</tr>
<tr>
<td>Variable</td>
<td>Description</td>
</tr>
<tr>
<td>----------</td>
<td>-------------</td>
</tr>
<tr>
<td>Market Capitalization (ME)</td>
<td>Worldscope item 08002. It is firm’s market capitalization at the fiscal year end. We convert it to USD using the exchange rate at firm’s fiscal year end date.</td>
</tr>
<tr>
<td>N of Firms Follows</td>
<td>The number of firms the analyst follows in a given fiscal year.</td>
</tr>
<tr>
<td>Relative Accuracy</td>
<td>The average of all the scores across all the firms the analyst covers and across all the years prior to the regulation, as defined in Section 1.4.</td>
</tr>
<tr>
<td>Return on Assets (ROA)</td>
<td>Worldscope item 08326. Firms’ return on assets.</td>
</tr>
<tr>
<td>Return Volatility (RETVOL)</td>
<td>Annualized standard deviation of daily returns over a year for a given firm. It is constructed from Datastream item RI (Return Index).</td>
</tr>
<tr>
<td>Stop</td>
<td>A dummy variable equal to 1 if the analyst stops covering a firm after the regulation.</td>
</tr>
<tr>
<td>Tenure</td>
<td>The total number of years the analyst appeared in I/B/E/S, starting from 1995.</td>
</tr>
<tr>
<td>Total Investment Return (RET)</td>
<td>Worldscope item 08801. It is measured as ( \frac{P_t + \text{Dividends}<em>{t} - P</em>{t-1}}{ME_{t-1}} ), ( P ) is the firm’s year-end price.</td>
</tr>
</tbody>
</table>
1.12 Additional figures

Figure 1.7: Geographic Composition of Stocks in Analyst Portfolio.

This figure shows the histogram of the geographic composition of the stocks in the analyst portfolio prior to the regulation. 1 means that all the firms the analyst covers are EU firms. 0 means that all the firms the analyst covers are US firms.
This figure shows the histogram of the average coverage of EU firms and US firms. Analyst coverage is the number of unique analysts covering a specific firm in fiscal year $t$. We report the averages of the period from 2014 to 2018.
This figure plots the histograms of the forecast errors in the pre- and post-regulation years. To get the forecast errors in the pre-regulation years, we simply average firms’ forecast errors overall these years.
This figure shows the average number of unique analysts one brokerage house hires over the years in the EU between small brokerage houses and large brokerage houses. Small brokerage houses are brokerage houses whose average number of employees (analysts they hire) prior to the regulation falls below the median. We adjusted the number so that both lines start at 0 in 2014.
# 1.13 Other summary statistics

Table 1.17: Summary Statistics (Country Information)

<table>
<thead>
<tr>
<th>Country</th>
<th>N of Firms</th>
<th>Coverage</th>
<th>Average GDP Growth Rate</th>
<th>Average Unemployment Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Belgium</td>
<td>70</td>
<td>5.851</td>
<td>1.490</td>
<td>7.648</td>
</tr>
<tr>
<td>Denmark</td>
<td>52</td>
<td>10.292</td>
<td>2.064</td>
<td>5.930</td>
</tr>
<tr>
<td>Finland</td>
<td>83</td>
<td>7.899</td>
<td>1.438</td>
<td>8.652</td>
</tr>
<tr>
<td>France</td>
<td>292</td>
<td>8.384</td>
<td>1.340</td>
<td>9.858</td>
</tr>
<tr>
<td>Germany</td>
<td>284</td>
<td>9.487</td>
<td>1.824</td>
<td>4.178</td>
</tr>
<tr>
<td>Iceland</td>
<td>3</td>
<td>1.333</td>
<td>3.924</td>
<td>3.504</td>
</tr>
<tr>
<td>Ireland</td>
<td>27</td>
<td>9.719</td>
<td>10.046</td>
<td>8.506</td>
</tr>
<tr>
<td>Italy</td>
<td>132</td>
<td>7.120</td>
<td>0.770</td>
<td>11.536</td>
</tr>
<tr>
<td>Netherlands</td>
<td>60</td>
<td>12.997</td>
<td>2.048</td>
<td>5.804</td>
</tr>
<tr>
<td>Norway</td>
<td>95</td>
<td>8.105</td>
<td>1.824</td>
<td>4.108</td>
</tr>
<tr>
<td>Portugal</td>
<td>19</td>
<td>7.526</td>
<td>1.828</td>
<td>10.626</td>
</tr>
<tr>
<td>Spain</td>
<td>75</td>
<td>15.107</td>
<td>2.654</td>
<td>19.770</td>
</tr>
<tr>
<td>Sweden</td>
<td>195</td>
<td>6.591</td>
<td>2.608</td>
<td>7.106</td>
</tr>
<tr>
<td>Switzerland</td>
<td>144</td>
<td>9.025</td>
<td>1.620</td>
<td>4.846</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>602</td>
<td>8.131</td>
<td>2.022</td>
<td>4.902</td>
</tr>
<tr>
<td>United States</td>
<td>2,259</td>
<td>11.145</td>
<td>2.284</td>
<td>4.922</td>
</tr>
</tbody>
</table>

This table provides information for the number of firms, the average coverage, the average GDP growth rate and the average unemployment rate in each country. Coverage is the number of unique analysts covering a certain firm. We report the averages of the period from 2014 to 2018 within each country.
Table 1.18: Summary Statistics (Small vs. Large in Sample for Quantity)

EU Firms (small): 1067 firms in total

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Mean</th>
<th>Median</th>
<th>Pctl(25)</th>
<th>Pctl(75)</th>
<th>St. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coverage</td>
<td>2.339</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>2.373</td>
</tr>
<tr>
<td>Book to Market</td>
<td>2.133</td>
<td>1.117</td>
<td>0.650</td>
<td>1.948</td>
<td>4.802</td>
</tr>
<tr>
<td>Investment Return (%)</td>
<td>11.957</td>
<td>4.938</td>
<td>−18.651</td>
<td>32.701</td>
<td>51.324</td>
</tr>
<tr>
<td>Market Capitalization (Bil $)</td>
<td>0.175</td>
<td>0.114</td>
<td>0.044</td>
<td>0.257</td>
<td>0.176</td>
</tr>
<tr>
<td>Return on Assets (%)</td>
<td>−0.456</td>
<td>3.639</td>
<td>−1.042</td>
<td>7.087</td>
<td>21.521</td>
</tr>
<tr>
<td>Return Volatility (%)</td>
<td>39.285</td>
<td>34.366</td>
<td>26.288</td>
<td>46.162</td>
<td>20.008</td>
</tr>
</tbody>
</table>

US Firms (small): 1130 firms in total

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Mean</th>
<th>Median</th>
<th>Pctl(25)</th>
<th>Pctl(75)</th>
<th>St. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coverage</td>
<td>4.892</td>
<td>4</td>
<td>2</td>
<td>7</td>
<td>4.092</td>
</tr>
<tr>
<td>Book to Market</td>
<td>2.130</td>
<td>1.000</td>
<td>0.569</td>
<td>2.103</td>
<td>3.358</td>
</tr>
<tr>
<td>Investment Return (%)</td>
<td>3.669</td>
<td>0.029</td>
<td>−23.309</td>
<td>23.484</td>
<td>46.361</td>
</tr>
<tr>
<td>Market Capitalization (Bil $)</td>
<td>0.552</td>
<td>0.391</td>
<td>0.149</td>
<td>0.829</td>
<td>0.516</td>
</tr>
<tr>
<td>Return on Assets (%)</td>
<td>−3.578</td>
<td>1.762</td>
<td>−2.570</td>
<td>5.939</td>
<td>31.654</td>
</tr>
<tr>
<td>Return Volatility (%)</td>
<td>42.377</td>
<td>36.543</td>
<td>26.659</td>
<td>51.664</td>
<td>21.888</td>
</tr>
</tbody>
</table>

This table provides the summary statistics for important variables in the sample for coverage quantity. We split the sample into small firms and large firms. Small firms are firms whose average fiscal year-end market capitalization falls below the median. We define the median cutoff separately in the EU and the US. Coverage is the number of unique analysts covering a certain firm. Book to Market is defined as total asset (Worldscope item 02999) minus long-term debt (Worldscope item 03251) over market value (Worldscope item 08002). Investment return is Worldscope item 08801, measured as (market price at year $t+1$ dividends−market price at year $t−1$)/market price at year $t−1$. Market capitalization is Worldscope Item 08002 measured in billion dollars. Return on Assets is Worldscope Item 08326. Return volatility is the annualized daily standard deviation of returns over a year for a given firm.
Table 1.18 Continued: Summary Statistics (Small vs. Large in Sample for Quality)

EU Firms (large): 1066 firms in total

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Mean</th>
<th>Median</th>
<th>Pctl(25)</th>
<th>Pctl(75)</th>
<th>St. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coverage</td>
<td>14.784</td>
<td>13</td>
<td>6</td>
<td>22</td>
<td>10.369</td>
</tr>
<tr>
<td>Book to Market</td>
<td>2.764</td>
<td>0.972</td>
<td>0.553</td>
<td>1.741</td>
<td>6.634</td>
</tr>
<tr>
<td>Investment Return (%)</td>
<td>11.237</td>
<td>8.954</td>
<td>-8.126</td>
<td>26.709</td>
<td>33.810</td>
</tr>
<tr>
<td>Market Capitalization (Bil $)</td>
<td>10.477</td>
<td>2.753</td>
<td>1.211</td>
<td>8.424</td>
<td>23.488</td>
</tr>
<tr>
<td>Return on Assets (%)</td>
<td>6.296</td>
<td>5.237</td>
<td>2.476</td>
<td>8.838</td>
<td>11.196</td>
</tr>
<tr>
<td>Return Volatility (%)</td>
<td>27.370</td>
<td>25.408</td>
<td>20.765</td>
<td>31.475</td>
<td>10.458</td>
</tr>
</tbody>
</table>

US Firms (large): 1129 firms in total

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Mean</th>
<th>Median</th>
<th>Pctl(25)</th>
<th>Pctl(75)</th>
<th>St. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coverage</td>
<td>17.403</td>
<td>16</td>
<td>9</td>
<td>23</td>
<td>10.563</td>
</tr>
<tr>
<td>Book to Market</td>
<td>1.453</td>
<td>0.658</td>
<td>0.375</td>
<td>1.240</td>
<td>3.057</td>
</tr>
<tr>
<td>Investment Return (%)</td>
<td>8.035</td>
<td>6.647</td>
<td>-10.257</td>
<td>24.342</td>
<td>31.781</td>
</tr>
<tr>
<td>Market Capitalization (Bil $)</td>
<td>19.927</td>
<td>5.608</td>
<td>2.813</td>
<td>15.068</td>
<td>52.163</td>
</tr>
<tr>
<td>Return on Assets (%)</td>
<td>5.745</td>
<td>5.350</td>
<td>2.315</td>
<td>9.224</td>
<td>8.684</td>
</tr>
<tr>
<td>Return Volatility (%)</td>
<td>27.667</td>
<td>24.843</td>
<td>19.983</td>
<td>31.756</td>
<td>11.686</td>
</tr>
</tbody>
</table>
Table 1.19: Summary Statistics (Small vs. Large in Sample for Quality)

EU Firms (small): 556 firms in total

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Mean</th>
<th>Median</th>
<th>Pctl(25)</th>
<th>Pctl(75)</th>
<th>St. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coverage</td>
<td>6.819</td>
<td>6</td>
<td>4</td>
<td>9</td>
<td>4.397</td>
</tr>
<tr>
<td>Forecast Error (%)</td>
<td>0.864</td>
<td>0.424</td>
<td>0.149</td>
<td>1.043</td>
<td>1.226</td>
</tr>
<tr>
<td>Dispersion (%)</td>
<td>0.902</td>
<td>0.463</td>
<td>0.205</td>
<td>1.013</td>
<td>1.274</td>
</tr>
<tr>
<td>Book to Market</td>
<td>1.799</td>
<td>0.921</td>
<td>0.547</td>
<td>1.494</td>
<td>4.330</td>
</tr>
<tr>
<td>Distance</td>
<td>127.678</td>
<td>123.683</td>
<td>94.475</td>
<td>156.745</td>
<td>52.578</td>
</tr>
<tr>
<td>Investment Return (%)</td>
<td>13.388</td>
<td>8.986</td>
<td>−10.783</td>
<td>31.634</td>
<td>40.040</td>
</tr>
<tr>
<td>Market Capitalization (Bil $)</td>
<td>0.874</td>
<td>0.694</td>
<td>0.328</td>
<td>1.289</td>
<td>0.691</td>
</tr>
<tr>
<td>Return on Assets (%)</td>
<td>4.973</td>
<td>5.395</td>
<td>2.651</td>
<td>8.696</td>
<td>11.331</td>
</tr>
<tr>
<td>Return Volatility (%)</td>
<td>30.715</td>
<td>28.603</td>
<td>23.283</td>
<td>35.474</td>
<td>11.319</td>
</tr>
</tbody>
</table>

US Firms (small): 847 firms in total

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Mean</th>
<th>Median</th>
<th>Pctl(25)</th>
<th>Pctl(75)</th>
<th>St. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coverage</td>
<td>7.821</td>
<td>6</td>
<td>4</td>
<td>9</td>
<td>5.195</td>
</tr>
<tr>
<td>Forecast Error (%)</td>
<td>0.594</td>
<td>0.266</td>
<td>0.099</td>
<td>0.656</td>
<td>0.974</td>
</tr>
<tr>
<td>Dispersion (%)</td>
<td>0.690</td>
<td>0.290</td>
<td>0.117</td>
<td>0.767</td>
<td>1.105</td>
</tr>
<tr>
<td>Book to Market</td>
<td>1.809</td>
<td>0.868</td>
<td>0.515</td>
<td>1.577</td>
<td>2.562</td>
</tr>
<tr>
<td>Distance</td>
<td>121.377</td>
<td>118.4</td>
<td>96.3</td>
<td>144.6</td>
<td>38.789</td>
</tr>
<tr>
<td>Investment Return (%)</td>
<td>6.578</td>
<td>2.473</td>
<td>−16.726</td>
<td>24.887</td>
<td>40.533</td>
</tr>
<tr>
<td>Market Capitalization (Bil $)</td>
<td>1.098</td>
<td>0.894</td>
<td>0.430</td>
<td>1.577</td>
<td>0.845</td>
</tr>
<tr>
<td>Return on Assets (%)</td>
<td>1.269</td>
<td>3.461</td>
<td>0.896</td>
<td>6.793</td>
<td>15.282</td>
</tr>
<tr>
<td>Return Volatility (%)</td>
<td>36.119</td>
<td>32.136</td>
<td>24.824</td>
<td>43.103</td>
<td>15.930</td>
</tr>
</tbody>
</table>

This table provides the summary statistics for important variables in the Sample for coverage quality. To be included in this sample, firms need to be covered by at least 2 analysts each year. We split the sample into small firms and large firms. Small firms are firms whose average fiscal year-end market capitalization fall below the median. We define the median cutoff separately in the EU and the US. Coverage is the number of unique analysts covering a certain firm. Forecast error is defined as the absolute distance between the firm’s actual annual earnings per share and the mean of the analyst forecasts, scaled by the firm’s previous year-end price. Forecast dispersion is defined as the standard deviation of all the forecasts across all the analysts following the same firm in the same year, scaled by the firm’s previous year-end price. Book to Market is defined as total asset (Worldscope item 02999) minus long-term debt (Worldscope item 03251) over market value (Worldscope item 08002). Distance is the average of the days between the analyst forecast date and the firm’s actual EPS report date. We take the average over all the analysts following the same firm and measure this variable in 2014. Investment return is Worldscope item 08801, measured as (market price at year \( t + 1 \) − dividends−market price at year \( t − 1 \))/market price at year \( t − 1 \). Market capitalization is Worldscope Item 08002 measured in billion dollars. Return on Assets is Worldscope Item 08326. Return volatility is the annualized standard deviation of daily returns over a year for a given firm. We remove observations for which forecast error and forecast dispersion is larger than 10% of the firm’s share price at the end of the previous year.
Table 1.19 Continued: Summary Statistics Continued (Small vs. Large in Sample for Quality)

EU Firms (large): 555 firms in total

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Mean</th>
<th>Median</th>
<th>Pctl(25)</th>
<th>Pctl(75)</th>
<th>St. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forecast Error (%)</td>
<td>0.584</td>
<td>0.281</td>
<td>0.102</td>
<td>0.667</td>
<td>0.901</td>
</tr>
<tr>
<td>Dispersion (%)</td>
<td>0.794</td>
<td>0.448</td>
<td>0.236</td>
<td>0.896</td>
<td>1.037</td>
</tr>
<tr>
<td>Book to Market</td>
<td>3.032</td>
<td>0.925</td>
<td>0.526</td>
<td>1.662</td>
<td>6.748</td>
</tr>
<tr>
<td>Distance</td>
<td>112.798</td>
<td>110.250</td>
<td>94.000</td>
<td>129.087</td>
<td>30.075</td>
</tr>
<tr>
<td>Market Capitalization (Bil $)</td>
<td>18.331</td>
<td>7.534</td>
<td>3.731</td>
<td>18.190</td>
<td>30.261</td>
</tr>
<tr>
<td>Return on Assets (%)</td>
<td>7.121</td>
<td>5.337</td>
<td>2.674</td>
<td>8.983</td>
<td>13.000</td>
</tr>
<tr>
<td>Return Volatility (%)</td>
<td>25.015</td>
<td>23.768</td>
<td>19.848</td>
<td>28.728</td>
<td>7.888</td>
</tr>
</tbody>
</table>

US Firms (large): 846 firms in total

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Mean</th>
<th>Median</th>
<th>Pctl(25)</th>
<th>Pctl(75)</th>
<th>St. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coverage</td>
<td>19.545</td>
<td>18</td>
<td>12</td>
<td>25</td>
<td>10.354</td>
</tr>
<tr>
<td>Forecast Error (%)</td>
<td>0.289</td>
<td>0.121</td>
<td>0.045</td>
<td>0.298</td>
<td>0.537</td>
</tr>
<tr>
<td>Dispersion (%)</td>
<td>0.388</td>
<td>0.174</td>
<td>0.075</td>
<td>0.405</td>
<td>0.681</td>
</tr>
<tr>
<td>Book to Market</td>
<td>1.340</td>
<td>0.618</td>
<td>0.349</td>
<td>1.168</td>
<td>2.198</td>
</tr>
<tr>
<td>Distance</td>
<td>107.477</td>
<td>105.570</td>
<td>86.808</td>
<td>125.500</td>
<td>31.060</td>
</tr>
<tr>
<td>Investment Return (%)</td>
<td>9.065</td>
<td>7.647</td>
<td>−7.965</td>
<td>24.327</td>
<td>29.647</td>
</tr>
<tr>
<td>Market Capitalization (Bil $)</td>
<td>25.628</td>
<td>8.691</td>
<td>4.335</td>
<td>21.100</td>
<td>59.127</td>
</tr>
<tr>
<td>Return on Assets (%)</td>
<td>6.369</td>
<td>5.718</td>
<td>2.841</td>
<td>9.663</td>
<td>8.011</td>
</tr>
</tbody>
</table>
1.14 Additional regression results

1.14.1 Dynamic coefficients point estimates

Table 1.20: Firm Level Outcomes (Dynamic Coefficients)

<table>
<thead>
<tr>
<th></th>
<th>Coverage (1)</th>
<th>Forecast Error (%) (2)</th>
<th>Dispersion (%) (3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EU × Y15</td>
<td>0.050</td>
<td>−0.020</td>
<td>−0.026</td>
</tr>
<tr>
<td></td>
<td>(0.094)</td>
<td>(0.052)</td>
<td>(0.045)</td>
</tr>
<tr>
<td>EU × Y16</td>
<td>−0.134</td>
<td>−0.018</td>
<td>−0.080***</td>
</tr>
<tr>
<td></td>
<td>(0.121)</td>
<td>(0.051)</td>
<td>(0.029)</td>
</tr>
<tr>
<td>EU × Y17</td>
<td>−0.201</td>
<td>0.033</td>
<td>−0.036</td>
</tr>
<tr>
<td></td>
<td>(0.193)</td>
<td>(0.045)</td>
<td>(0.073)</td>
</tr>
<tr>
<td>EU × Y18</td>
<td>−0.715***</td>
<td>−0.143***</td>
<td>−0.199***</td>
</tr>
<tr>
<td></td>
<td>(0.269)</td>
<td>(0.044)</td>
<td>(0.038)</td>
</tr>
</tbody>
</table>

Controls: Yes
Firm FE: Yes
Year FE: Yes
Observations: 21,960
R²: 0.967
Adjusted R²: 0.958

Note: *p<0.1; **p<0.05; ***p<0.01

This table shows the results of the dynamic coefficient regressions specified in Model (1.4). The dependent variable in column (1), is Coverage, the number of analysts covering a specific firm in fiscal year \( t \). The dependent variable in column (2), is Forecast Error, defined as the absolute distance between the firm’s actual annual earnings per share and the mean of the analyst forecasts, scaled by the firm’s previous year-end price. The dependent variable in column (3) is Forecast Dispersion, defined as the standard deviation of all the forecasts across all the analysts following the same firm in the same year, scaled by the firm’s previous year-end price. EU is a dummy variable equal to one if the firm is domiciled and listed in Europe. \( Y_{15}, Y_{16}, Y_{17}, Y_{18} \) are year dummy variables equal to one if year \( t \) is equal to 2015, 2016, 2017, 2018. We include other important firm level controls whose definitions and constructions can be found in the Appendix 1.11.3. Standard errors are clustered at the country level.
Table 1.21: Intensive Margin (Dynamic Coefficients)

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>Forecast Error (Analyst) (%)</th>
<th>Average Forecast Error (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>EU × Y15</td>
<td>0.031</td>
<td>0.010</td>
</tr>
<tr>
<td>EU × Y16</td>
<td>−0.044</td>
<td>−0.075*</td>
</tr>
<tr>
<td>EU × Y17</td>
<td>0.004</td>
<td>−0.008</td>
</tr>
<tr>
<td>EU × Y18</td>
<td>−0.132***</td>
<td>−0.150***</td>
</tr>
</tbody>
</table>

| Frim level Controls | Yes                         | Yes                        | Yes                        |
| Analyst level Controls × Post | Yes                       | Yes                        | Yes                        |
| Analyst FE          | Yes                         | Yes                        | Yes                        |
| Firm FE             | Yes                         | No                         | No                         |
| Year FE             | Yes                         | Yes                        | No                         |
| B House FE          | No                          | Yes                        | No                         |
| B House FE × Year FE| No                          | No                         | Yes                        |

| Observations | 81,445 | 12,990 | 4,111 |
| R^2           | 0.422  | 0.559  | 0.598 |
| Adjusted R^2  | 0.378  | 0.431  | 0.480 |

Note: *p<0.1; **p<0.05; ***p<0.01

This table shows the results of the dynamic coefficient regressions specified in Model (1.4). The dependent variable in column (1), is *Forecast Error (Analyst)*, defined as the absolute distance between the actual annual earnings per share and the analyst’s forecast, scaled by the firm’s previous year-end price. The dependent variable in column (2) and (3) is *Average Forecast Error*, defined as the average of forecast errors across all the firms the analyst covers in a given year. *EU* is a dummy variable equal to one if the analyst’s portfolio consists of at least 70% of EU stock and zero if the analyst’s portfolio consists of at most 30% of EU stocks. *Y*15, *Y*16, *Y*17, *Y*18 are year dummy variables equal to one if year *t* is equal to 2015, 2016, 2017, 2018. We include all the analysts (small and large) used in the intensive margin analyses. Standard errors are clustered at the firms level in column (1) and at the analyst level in column (2) and (3).
## 1.14.2 Additional results for firm level coverage

Table 1.22: Firm Level Coverage (Sample for Coverage Quality).

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>Coverage</th>
<th>Full</th>
<th>Small vs Large</th>
<th>Triple Diff</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EU × POST</td>
<td>-1.219***</td>
<td>-2.496***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.322)</td>
<td>(0.391)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>EU × POST × SMALL</td>
<td>2.551***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.242)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SMALL × POST</td>
<td>1.968***</td>
<td>-0.564***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.229)</td>
<td>(0.009)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Controls</th>
<th>Yes</th>
<th>Yes</th>
<th>Yes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Firm FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Year FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>14,020</td>
<td>5,555</td>
<td>14,020</td>
</tr>
<tr>
<td>R²</td>
<td>0.964</td>
<td>0.965</td>
<td>0.964</td>
</tr>
<tr>
<td>Adjusted R²</td>
<td>0.954</td>
<td>0.956</td>
<td>0.955</td>
</tr>
</tbody>
</table>

Note: *p<0.1; **p<0.05; ***p<0.01

This table shows the results of the panel regressions specified in Model (1.1) on the sample for coverage quality. The dependent variable is Coverage, the number of unique analysts covering a specific firm in fiscal year \( t \). EU is a dummy variable equal to one if the firm is domiciled and listed in Europe. POST is a dummy variable equal to one if the fiscal year \( t \) is equal to 2018. SMALL is a dummy variable equal to one if firms’ average fiscal year-end market capitalization over the pre-regulation years falls below the median. To maintain the proportion of EU and US firms fixed, we calculate the cutoff separately in both regions. Column (1) is the result of the difference-in-difference regression between EU firms and US firms. Column (2) is the result of the difference-in-difference regression between small firms and larger firms within the EU. Column (3) is the result of the triple-difference regression. Standard errors are clustered at the country level.
This table shows the results of panel regressions specified in Model (1.1). The dependent variable in column (1) to column (3) is $\text{Log} (1 + \text{Coverage})$, the log of $1 +$ the number of unique analysts covering a specific firm in fiscal year $t$. The dependent variable in column (4) to column (6) is $\text{Log Coverage}$, log of the number of unique analysts covering a specific firm in fiscal year $t$. $\text{EU}$ is a dummy variable equal to one if the firm is domiciled and listed in Europe. $\text{POST}$ is a dummy variable equal to one if the fiscal year $t$ is equal to 2018. $\text{SMALL}$ is a dummy variable equal to one if firms’ average fiscal year-end market capitalization over the pre-regulation years falls below the median. To maintain the proportion of EU and US firms fixed, we calculate the cutoff separately in both regions. Column (1) to column (3) are results obtained in the sample for coverage quantity. Since coverage can be 0 in this sample, we put $1 + \text{Coverage}$ inside the log function. Column (4) to (6) are the results obtained in the sample for coverage quality. Column (1) and column (4) are the results for the difference-in-difference regression between EU firms and US firms. Column (2) and column (5) are the results for the difference-in-difference regression between small firms and larger firms within the EU. Column (3) and Column (6) are the results for the triple-difference regression. Standard errors are clustered at the country level.
Table 1.24: Firm Level Coverage (EU Small vs. US Small and EU Large vs. US Large)

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>Coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Full (1)</td>
</tr>
<tr>
<td>EU × POST</td>
<td>-0.651*** (0.196)</td>
</tr>
<tr>
<td>EU × POST × SMALL</td>
<td>1.891*** (0.200)</td>
</tr>
<tr>
<td>SMALL × POST</td>
<td>-0.228*** (0.018)</td>
</tr>
</tbody>
</table>

Controls: Yes
Firm FE: Yes
Year FE: Yes
Observations: 21,960
R²: 0.958
Adjusted R²: 0.959

Note: *p<0.1; **p<0.05; ***p<0.01

This table shows the results of panel regressions specified in Model (1.1) and Model (1.3). The dependent variable is Coverage, the number of unique analysts covering a specific firm $j$ in fiscal year $t$. EU is a dummy variable equal to one if the firm is domiciled and listed in Europe. POST is a dummy variable equal to one if the fiscal year $t$ is equal to 2018. SMALL is a dummy variable equal to one if firms’ average fiscal year-end market capitalization over the pre-regulation years falls below the median. To maintain the proportion of EU and US firms fixed, we calculate the cutoff separately in both regions. We split the firms into small firms and large firms and perform difference-in-difference analyses separately for them. Column (1) to column (3) are results for all the firms, small firms, and large firms respectively. Column (4) is the result of the triple-difference regression. Standard errors are clustered at the country level.
1.14.3 Additional results for analyst informativeness

Table 1.25: Analyst Informativeness (Small vs. Large).

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>Analyst Informativeness</th>
<th>Aggregate Analyst Informativeness</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Full (1)</td>
<td>Small vs Large (2)</td>
</tr>
<tr>
<td>EU × POST</td>
<td>0.0003***</td>
<td>0.0004***</td>
</tr>
<tr>
<td></td>
<td>(0.0001)</td>
<td>(0.0001)</td>
</tr>
<tr>
<td>EU × POST × SMALL</td>
<td>−0.0002</td>
<td>0.011***</td>
</tr>
<tr>
<td>SMALL × POST</td>
<td>0.0002</td>
<td>0.0004***</td>
</tr>
</tbody>
</table>

Controls: Yes Yes Yes Yes Yes Yes
Firm FE: Yes Yes Yes Yes Yes Yes
Year FE: Yes Yes Yes Yes Yes Yes
Observations: 19,391 8,844 19,391 19,391 8,844 19,391
R²: 0.496 0.398 0.496 0.917 0.924 0.917
Adjusted R²: 0.353 0.218 0.353 0.894 0.902 0.894

*Note:* *p<0.1; **p<0.05; ***p<0.01

This table shows the results of panel regressions similar to Model (1.1). The dependent variable in column (1) to column (3) is Analyst Informativeness. It captures the average informativeness of one forecast revision date. The dependent variable in column (4) to column (6) is Aggregate Analyst Informativeness. It captures the aggregate informativeness of all revision dates. EU is a dummy variable equal to one if the firm is domiciled and listed in Europe. POST is a dummy variable equal to one if the fiscal year \( t \) is equal to 2018. SMALL is a dummy variable equal to one if firms’ average fiscal year-end market capitalization over the pre-regulation years falls below the median. To maintain the proportion of EU and US firms fixed, we calculate the cutoff separately in both regions. Column (1) and column (4) are the results for the difference-in-difference regression between EU firms and US firms. Column (2) and column (5) are the results for the difference-in-difference regression between small firms and larger firms within the EU. Column (3) and Column (6) are the results for the triple-difference regression. Standard errors are clustered at the firm level.
1.15 Propensity score matching

In our difference-in-difference analyses, we use EU firms as the treatment group and US firms as the control group. One concern is that firms in these two continents are different. For example, Table 1.1 shows that US firms are in general larger, have lower book-to-market ratios and lower ROAs. Although we explicitly control for these firm level characteristics, the differences may still confound the effects of unbundling. In this section, we conduct a standard propensity score matching to further check whether our results are driven by the observable differences in the firm level characteristics.

In the sample for coverage quantity, we use 2014 data and match EU firms with US firms according to their book to market ratio, investment return, market capitalization, return on assets and return volatility. We choose the nearest-neighbor matching method and set the caliper parameter to be 0.25 standard deviation of the estimated propensity scores (e.g., Rosenbaum and Rubin (1985)). With this restriction, nearest-neighbor matching will be considered only if $|P_t - P_c| < 0.25\sigma_p$, where $P_t$ and $P_c$ are the propensity scores for the treatment and the control, and $\sigma_p$ is the standard deviation of the estimated propensity score of the sample. We repeat a similar process in the sample for coverage quality but include distance (the average of the days between the analyst forecast date and the firm’s actual EPS report date) as an additional matching covariate. Table 1.26 shows the summary statistics of the covariates before and after the matching. After the propensity score matching, in the sample for coverage quantity, the treatment group and the control group are only different in the investment return and market capitalization. In the sample for coverage quality, the treatment group and the control group are comparable along all the firm level characteristics. Figure 1.11 shows the histograms of the propensity scores after matching. We can observe that the distributions between treatment group and control groups are comparable. In fact, a formal Kolmogoro-Smirnov test on the two samples does not reject the null that the treatment sample and the control sample are drawn from the same distribution. Hence, the propensity score

---

51 We do not have a large number of control firms to select. Requiring each firm in the treatment group to be matched with a firm in the control group will again introduce dissimilarities between the two groups.
matching effectively removes observations that are too different between the two groups based on observables.

We then repeat the difference-in-difference analyses at the firm level. Table 1.27 reports the results. Consistent with what we find before, analyst coverage drops for EU firms and the drop is concentrated in large firms. Similarly, forecast errors of EU firms decrease. The magnitude of these coefficients is comparable with our baseline results without matching. Hence we conclude that observable differences between EU firms and US firms do not drive our results and US firms could serve as a valid counterfactual. In untabulated tests, we perform propensity score matching at the analyst level for the intensive margin analyses and obtain similar results to the baseline model in which we do not match the observations.
Figure 1.11: Distributions of Propensity Scores after Matching.

This figure plots the histogram of the propensity scores in the matched sample between the treatment group and the control group.
Table 1.26: Covariates Balance Table.

Panel A: Before Matching (Sample for Coverage Quantity)

<table>
<thead>
<tr>
<th></th>
<th>Treated</th>
<th>Control</th>
<th>Diff</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>N of Obs</td>
<td>2133</td>
<td>2259</td>
<td>126</td>
<td></td>
</tr>
<tr>
<td>BM (%)</td>
<td>2.53</td>
<td>1.73</td>
<td>0.80</td>
<td>0.00</td>
</tr>
<tr>
<td>Investment Return (%)</td>
<td>13.82</td>
<td>8.27</td>
<td>5.55</td>
<td>0.00</td>
</tr>
<tr>
<td>Market Cap (Bil $)</td>
<td>5.24</td>
<td>9.21</td>
<td>−3.97</td>
<td>0.00</td>
</tr>
<tr>
<td>ROA (%)</td>
<td>2.87</td>
<td>1.53</td>
<td>1.34</td>
<td>0.09</td>
</tr>
<tr>
<td>Return Volatility</td>
<td>32.39</td>
<td>32.06</td>
<td>0.33</td>
<td>0.53</td>
</tr>
</tbody>
</table>

Panel B: After Matching (Sample for Coverage Quantity)

<table>
<thead>
<tr>
<th></th>
<th>Treated</th>
<th>Control</th>
<th>Diff</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>N of Obs</td>
<td>2074</td>
<td>2074</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>BM (%)</td>
<td>1.97</td>
<td>1.78</td>
<td>0.19</td>
<td>0.11</td>
</tr>
<tr>
<td>Investment Return (%)</td>
<td>11.83</td>
<td>9.20</td>
<td>2.63</td>
<td>0.02</td>
</tr>
<tr>
<td>Market Cap (Bil $)</td>
<td>5.25</td>
<td>6.36</td>
<td>1.11</td>
<td>0.04</td>
</tr>
<tr>
<td>ROA (%)</td>
<td>3.01</td>
<td>2.05</td>
<td>0.96</td>
<td>0.11</td>
</tr>
<tr>
<td>Return Volatility</td>
<td>32.11</td>
<td>32.01</td>
<td>0.10</td>
<td>0.86</td>
</tr>
</tbody>
</table>

Panel C: Before Matching (Sample for Coverage Quality)

<table>
<thead>
<tr>
<th></th>
<th>Treated</th>
<th>Control</th>
<th>Diff</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>N of Obs</td>
<td>1111</td>
<td>1693</td>
<td>582</td>
<td></td>
</tr>
<tr>
<td>BM (%)</td>
<td>2.44</td>
<td>1.53</td>
<td>0.91</td>
<td>0.00</td>
</tr>
<tr>
<td>Distance</td>
<td>120.24</td>
<td>114.43</td>
<td>5.81</td>
<td>0.00</td>
</tr>
<tr>
<td>Investment Return (%)</td>
<td>15.18</td>
<td>10.10</td>
<td>5.08</td>
<td>0.00</td>
</tr>
<tr>
<td>Market Cap (Bil $)</td>
<td>9.45</td>
<td>11.96</td>
<td>2.51</td>
<td>0.04</td>
</tr>
<tr>
<td>ROA (%)</td>
<td>6.08</td>
<td>4.26</td>
<td>1.82</td>
<td>0.00</td>
</tr>
<tr>
<td>Return Volatility</td>
<td>26.26</td>
<td>28.35</td>
<td>−2.09</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Panel D: After Matching (Sample for Coverage Quality)

<table>
<thead>
<tr>
<th></th>
<th>Treated</th>
<th>Control</th>
<th>Diff</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>N of Obs</td>
<td>1067</td>
<td>1067</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>BM (%)</td>
<td>1.74</td>
<td>1.74</td>
<td>0.00</td>
<td>0.96</td>
</tr>
<tr>
<td>Distance</td>
<td>118.86</td>
<td>117.37</td>
<td>1.49</td>
<td>0.38</td>
</tr>
<tr>
<td>Investment Return (%)</td>
<td>14.75</td>
<td>13.35</td>
<td>0.408</td>
<td>0.31</td>
</tr>
<tr>
<td>Market Cap (Bil $)</td>
<td>9.29</td>
<td>8.34</td>
<td>0.95</td>
<td>0.31</td>
</tr>
<tr>
<td>ROA (%)</td>
<td>5.91</td>
<td>5.24</td>
<td>0.67</td>
<td>0.10</td>
</tr>
<tr>
<td>Return Volatility</td>
<td>26.34</td>
<td>26.64</td>
<td>−0.30</td>
<td>0.52</td>
</tr>
</tbody>
</table>

This table shows the difference in the average values of matching covariates before and after the propensity score matching. Book to Market is defined as total asset (Worldscope item 02999) minus long-term debt (Worldscope item 03251) over market value (Worldscope item 08002). Distance is the average of the days between the analyst forecast date and the firm’s actual EPS report date. We take the average over all the analysts following the same firm and measure this variable in 2014. Investment return is Worldscope item 08801, measured as (market price at year $t$+ dividends−market price at year $t−1$)/market price at year $t−1$. Market capitalization is Worldscope Item 08002 measured in billion dollars. Return on Assets is Worldscope Item 08326. Return volatility is the annualized standard deviation of daily returns over a year for a given firm. Panel A and Panel show the results of sample for coverage quantity and Panel C and Panel D shows the results of the sample for coverage quality.
Table 1.27: Firm Level Outcomes (Propensity Score Matching).

<table>
<thead>
<tr>
<th></th>
<th>Coverage</th>
<th></th>
<th>Forecast Error</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Full</td>
<td>Small vs Large</td>
<td>Triple Diff</td>
<td>Full</td>
</tr>
<tr>
<td>EU \times POST</td>
<td>-0.694*** (0.193)</td>
<td>-1.568*** (0.266)</td>
<td>-0.141*** (0.036)</td>
<td>-0.120** (0.048)</td>
</tr>
<tr>
<td>EU \times POST \times SMALL</td>
<td>1.753*** (0.198)</td>
<td></td>
<td>0.042 (0.065)</td>
<td></td>
</tr>
<tr>
<td>SMALL \times POST</td>
<td>1.662*** (0.203)</td>
<td>-0.102*** (0.017)</td>
<td>0.061 (0.059)</td>
<td>0.083*** (0.016)</td>
</tr>
</tbody>
</table>

| Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Firm FE  | Yes | Yes | Yes | Yes | Yes | Yes |
| Year FE  | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 20,780 | 10,390 | 20,780 | 10,670 | 5,335 | 10,670 |
| $R^2$    | 0.956 | 0.972 | 0.965 | 0.500 | 0.499 | 0.500 |
| Adjusted $R^2$ | 0.956 | 0.964 | 0.957 | 0.374 | 0.372 | 0.374 |

Note: *p<0.1; **p<0.05; ***p<0.01

This table shows the results of the panel regressions specified in Model (1.1) and Model (1.3). We use the propensity score matching and match EU firms with US firms according to their book to market ratio, investment return, market capitalization, return on assets and return volatility. For the sample for coverage quality, we include distance as an additional matching covariate. All variables are measured in 2014. The dependent variable in column (1) to (3) is Coverage, the number of unique analysts covering a specific firm $j$ in fiscal year $t$. The dependent variable in column (4) to (6) is Forecast Error, defined as the absolute distance between the firm’s actual EPS and the mean of the analyst forecasts, scaled by the firm’s previous year-end price. EU is a dummy variable equal to one if the firm is domiciled and listed in Europe. POST is a dummy variable equal to one if the fiscal year $t$ is equal to 2018. SMALL is a dummy variable equal to one if firms’ average fiscal year-end market capitalization over the pre-regulation years falls below the median. To maintain the proportion of EU and US firms fixed, we calculate the cutoff separately in both regions. Column (1) and column (4) are the results for the difference-in-difference regression between EU firms and US firms. Column (2) and column (5) are the results for the difference-in-difference regression between small firms and larger firms within the EU. Column (3) and Column (6) are the results for the triple-difference regression. Standard errors are clustered at the country level.
Chapter 2

Going Public or Staying Private? The Cost of Mandated Transparency ¹

2.1 Introduction

Once upon a time, entrepreneurs dreamed of raising capital in the public markets through an initial public offering (IPO). But the past decade has seen an explosion in the breadth and magnitude of capital formation in private markets. In August 2018, the SEC concluded that “amounts raised through unregistered securities offerings have outpaced the level of capital formation through registered securities offerings, and totaled more than $3.0 trillion during 2017.”² The growth of private capital formation at the expense of public IPOs has been the subject of a large literature (e.g., Ewens and Farre-Mensa (2019)) and policy debate.³

It is difficult to identify a single driver of the public - private tradeoff. But one defining institutional characteristic of public markets is transparency. Publicly traded companies are subject to a long list of disclosure mandates under the securities laws. In addition, in public markets order flow is transparent: everyone knows, in real time, which shares are being exchanged and at what price. There are strong legal constraints on trading on the basis of material, nonpublic information. These attributes of public markets enhance liquidity and facilitate smooth trading between buyers and sellers.

In this paper, we identify an under-appreciated cost of institutional transparency. The intuition

¹This chapter is based on Guo and Mitts (2019). For valuable comments, we thank Ian Ayres, Patrick Bolton, Anthony Casey, Merritt Fox, Gur Huberman, Robert J. Jackson, Jr., Wei Jiang, Yair Listokin, Ed Morrison, Giorgia Piacentino, Eric Tulley and Neng Wang.
is simple: communication is imperfect. Mandated metrics like earnings can mislead as to the fundamental value of a firm. Disclosure rules may thus force a firm to “speak” when it would prefer to remain silent. Suppose, for example, that a public company embarks on a project that yields a high long-run value but involves a poor short-term signal like low initial revenues. Observing the poor short-term signal, investors may worry that the firm’s project is not very valuable. To exploit the market’s mistaken inference, informed traders can acquire information about the project’s long-run value. If private information is fully incorporated into prices, *i.e.*, markets are strong-form efficient, then the price will still reflect the long-run value of the firm, even if the short-term signal does not. But because the public market reveals order flow, informed traders may not expect to make sufficient profits to engage in costly information acquisition. When this happens, the short-term price deviates from the long-run value and an entrepreneur who cares about the short-term price (e.g., for liquidity reasons) is penalized by the disclosure rule for pursuing a high-value project, and will underinvest as a result.

This not merely a theoretical phenomenon. Public-company managers have testified that the securities laws require the disclosure of information that has little to do with firm value. In the words of one executive, “[i]nvestors find value in biotech companies by understanding scientific milestones and clinical trial progress,” “not financial disclosures that simply show a decade-plus of R&D expenses” yet “small, pre-revenue biotechs are often required to file the same reports as revenue generating, profitable corporate behemoths.” Another example is a “total compensation” figure for executive pay, which reduces complex contractual contingencies to a single number. Or, in a related vein, the “actuarial” value of executive pension benefits, which may rely on dubious assumptions about life expectancy and economic uncertainties. All of these disclosures are mandatory under the securities laws, even when they may mislead as to long-run firm value.

The core contribution of our project is to show how the “double friction” of observable order

---

4It will be clear in our model that the short-term price reflects all available information and the market is semi-strong efficient. From the standpoint of investors, there is only ex post mispricing.


6See, *e.g.*, Rule 7.04, Regulation S-X, 17 C.F.R. § 210.7–04 (requiring disclosure of earnings per share); Item 402(c), Regulation S-K, 17 C.F.R. 229.402, 229.405 (mandating inclusion of summary compensation table and pension benefits table in proxy statement).
flow and imperfect disclosure can render mandated transparency harmful, not just superfluous. To be sure, prior literature has considered each of these separately. The informed trading literature Kyle (1985) has long shown that the expected profit to information acquisition in a market with observable order flow is less than the value of the information itself. And the disclosure literature has shown that accounting metrics are poorly suited for firms with a heavy amount of intangible assets Leuz and Wysocki (2016); Lev and Gu (2016). But the whole is more than the sum of its parts: the inefficiencies of imperfect disclosure are exacerbated when the trading environment is transparent.

We develop a simple model to illustrate this idea. In the first period, the entrepreneur decides to go public or stay private. Depending on her type, which is private information, she will choose the firm’s production technology, which boosts the firm’s long-run value but may also generate a poor short-term signal. The entrepreneur cares about the short-term signal because she may suffer a liquidity shock and need to liquidate the firm in the short run at a price reflecting that signal. In our model, the entrepreneur can choose between two kinds of financial markets: a transparent “public” market and a more opaque “private” market. The former mandates disclosure of the short-run signal with observed order flow while the latter facilitates bargaining with potential investors.

When public markets cannot perfectly separate (a) high-value firms who issue poor short-term signals from (b) low-value firms who also issue poor short-term signals, high-value firms may avoid valuable investment opportunities. There is an inherent tension in mandating disclosure: these signals allow investors to separate good firms from bad ones, but may also deter valuable investments when short-run signals diverge from long-run value. To some extent, arbitrage incentives mitigate the inefficiencies of disclosure, but order-flow transparency limits the expected profits to information acquisition.

Our point is that private markets provide a tempting alternative. Opacity is a virtue: it allows investors to recoup the cost of looking past noisy short-run signals. We show that entrepreneurs will prefer to raise capital in an illiquid private market to the extent that there is a greater divergence between short-run signals and long-run value. In our framework, imperfect communication is what
drives the entrepreneur’s preference for private markets. This result can be generalized to a variety of settings, including a limited number of investors, unobservable liquidity shocks, and financing a project on the primary market in an initial public offering (rather than merely selling the firm on the secondary market in response to a liquidity shock, as in our baseline model).

The key normative implication of our framework is that mandatory disclosure may lead to inefficient investment. Disclosure rules have long been criticized for imposing compliance costs in excess of marginal benefits. But we identify an independent cost of transparency: some firms which, on the margin, had optimally opted to remain silent, will no longer be able to do so. Of course, we recognize that these losses must be balanced against the gains that accrue to society from disclosure in the form of greater liquidity and narrower bid-ask spreads Fox (2009). We acknowledge that it is difficult to aggregate these gains and losses with respect to any given disclosure rule. But our framework can guide policymakers seeking to evaluate which public companies are most likely to suffer from a disclosure rule: namely, those for whom short-run signals diverge from long-run value.

Our conclusion turns on the assumption that it is possible that short-run signals negatively correlate with long-run payoffs. Doidge et al. (2018) suggest this occurs in industries like technology, where accounting metrics understate the value of intangible assets. When that happens, great projects can look more like a waste of money than mediocre ones. The greater the divergence between short-run signals and long-run value, the less useful short-term prices are. Entrepreneurs take into account price informativeness when choosing whether to raise capital on the public or private markets: because going public necessarily implies disclosure, a gap between how a market reacts to information today and the firm’s long-run value adds a kind of “hidden cost” to becoming a public company.

Our project relates to the literature on “cream skimming” in financial markets. Bolton et al. (2016) show how dealers in private markets exploit private information to pick off higher-quality assets originated by uninformed entrepreneurs, lowering the overall quality of assets in public, exchange markets. Unlike Bolton et al. (2016), our framework focuses on the role of mandatory
disclosure in driving the public-private choice. We embrace a key idea in Bolton et al. (2016), namely, that the transparency of order flow in public markets reduces the incentive for potential informed investors to acquire costly private information. Therefore, in equilibrium, informed investors tend to prefer private markets, where they can capture the gains to these investments in information acquisition.

Our work is also related to the real effect of price accuracy: prices help structure managerial incentives and guide managers to make better decisions in the real economy Holmström and Tirole (1993); Bolton et al. (2006); Edmans et al. (2012, 2015). Indeed, Fox (1999) shows that price accuracy is a primary justification for mandatory disclosure. For prices to be informative, investors must acquire information and trade in the public market. But the cost of acquiring information must be compensated by expected trading profits. We link price informativeness to a firm’s decision to raise capital in transparent, public markets and provide a new explanation on why mandated transparency may lead to low price efficiency for high-value firms.

The central policy implication of our model is that mandatory disclosure can be value-destroying. While high-value firms benefit from price accuracy in the short run, disclosure can also induce inefficient investment decisions. Our framework counsels policymakers against a rush to mandate the reporting of metrics that are “grossly simplified.” Disclosure can distort real investment when these metrics are a noisy proxy for fundamental value.

### 2.2 Baseline Model

The basic structure of our model is as follows. An entrepreneur chooses a financial market and a production technology. With some probability, the entrepreneur suffers a liquidity shock and must sell the firm to either informed or uninformed investors. Investors quote prices according to the trading mechanism in a given market. For simplicity, all agents are risk-neutral and there is no discounting. Our model features two financial markets: a competitive and transparent exchange.
market, which we call the “public” market; and an opaque and less competitive market, which we call the “private” market.

2.2.1 Agents

Entrepreneur

At $t = 0$, the entrepreneur must raise initial investment $I$ and choose a market in which to list the firm. In our baseline model, we make the simplifying assumption that the entrepreneur can finance the project herself. We do this to focus attention on the choice between public and private markets, $\phi \in \{Pub, Priv\}$, which matters to the entrepreneur because she may suffer a liquidity shock and will be forced to sell the firm at the prevailing price in the given market $\phi$. We later relax this assumption and consider the case where the entrepreneur sells an equity stake $\alpha$ to outside investors. Regardless of whether the entrepreneur sells equity, we assume throughout the paper that the entrepreneur controls the firm and thus we use the terms “entrepreneur”, “manager” and “firm” interchangeably.

At $t = 1$, the entrepreneur’s type $\tau$ is realized. She can be of two types, $\tau \in \{H, L\}$ with equal probability. Type $H\,(L)$ corresponds to a high (low) quality firm. An $L$ type firm has nothing to decide and yields a final value of $V^L$ at $t = 3$, which we normalize to 0. An $H$ type firm can choose a production technology $\lambda \in [0, 1]$ which generates a final value of $V^H = V_0 + \lambda I$. We can also interpret $\lambda$ as a parameter indexing different types of projects. Project $\lambda$ generates final cash flow $V_0 + \lambda I$ and the high-type firm chooses its project. The type $\tau$ and the production technology $\lambda$ are private information of the firm. We assume that $\tau$ cannot be credibly disclosed. We also make the technical assumption that $V_0 > 2I$, i.e., the project’s ex ante expected value exceeds the requisite investment.

At $t = 2$, if the firm chooses to go public, it must disclose a short-term signal (such as quarterly earnings) $s \in \{G, B\}$ to the public market. We call $s$ hard information that is imperfectly informative of the true fundamental value of the firm: $s$ not only depends on firm’s type $\tau$ but also
the investment intensity $\lambda$ and a parameter $\theta$ measuring the extent to which $s$ reveals the $V^H$ or $V_0$. That is, a low-quality firm always generates a poor short-term signal ($s = B$). A high-quality firm generates a poor short-term signal with probability $\theta\lambda^2$ and a good short-term signal ($s = G$) with probability $1 - \theta\lambda^2$.

This signal structure implies that projects with a higher long-run value are more likely to generate poor short-term signals. This is a critical condition that must be satisfied for our conclusions to hold. We believe this is the case in some industries (e.g., technology). For example, investment in research and development tends to impose substantial costs in the short-run but may boost the value of the firm in the long run Edmans et al. (2016). $\theta$ measures the gap between short-run signals and long-run value. The extreme case of $\theta = 0$ implies that a high-type firm can always separate from a low-type one, because short-run signals are perfectly informative of long-run value. We assume that $\theta$ is publicly known.

For simplicity, we assume that if the entrepreneur chooses to stay private, she does not transmit the signal $s$. That is, disclosure is nonexistent in the private market, and as we will show later, under certain conditions, the firm is strictly better off by committing not to disclose. Moreover, in an extension, we show that our results continue to hold even when voluntary disclosure is allowed in the private market.

Finally, the entrepreneur suffers a liquidity shock with probability $\omega$ that forces her to sell the firm at $t = 1$. In the baseline model, we assume that the liquidity shock is observable but not contractible. Our results remain if we relax this assumption. Letting $P$ denote the price at which the entrepreneur can sell the firm at $t = 1$, her objective function is simply $\omega P + (1 - \omega)V$. $\omega$ can also be interpreted as the short-term stock price concern arising from the myopia literature Stein (1988). We discuss price formation shortly.

---

8We make the technical assumption that acquiring private information entails a nonzero cost. Otherwise, informed investors may still opt for zero profit in the public market.
Informed Investors

At \( t = 0 \), a proportion \( e \) of informed investors can obtain private information \( i \) about the fundamental value of the firm.\(^8\) At \( t = 1 \), \( i \) is realized. For simplicity, we assume that the private information always reveals the true value of the firm (\( i \in \{i_0 = V^H, i_b = V^L\} \)). Recall that in our baseline model, the entrepreneur personally funds the project. Thus, we use the term “investors” to refer to those who offer to purchase the firm when the entrepreneur encounters a liquidity shock at \( t = 2 \).

2.2.2 Markets

Public Market

At \( t = 2 \), the public market is operational. There is a large measure of uninformed investors who make rational inferences about the value of the firm based on the information they possess. Uninformed investors compete among each other and post bids at the expected value of the firm, conditional on all available information.

The public market is transparent in the sense that all bids are publicly observable. We consider the following trading protocol for the public market as in Bolton et al. (2016). First, the firm is put up for sale at some price \( P^T \). Any buyer willing to bid more than \( P^T \) can make a targeted bid for the firm. All bids are public information. Then any other buyer can submit a counterbid using what they inferred from the first round of bidding. Finally, the firm is sold to the highest bidder.

It is straightforward to see that informed investors will not remain in the public market because their private information will be revealed to the market via their bids: any deviation from the uninformed bid will lead to a counterbid and thus informed investors cannot profit from the acquisition of private information. We therefore obtain the following lemma:

**Lemma 2.1** Informed investors do not trade in the public market.

The intuition behind Lemma 2.1 is similar to Grossman and Stiglitz (1980): the acquisition of private information is betrayed by trading behavior. This result holds when we have a more so-
phisticated trading mechanism. Thus, if the entrepreneur goes public, the price she will receive is simply $P^{pub} = P_{\tau}^{s(H|s)} \hat{V}^H$ where $\hat{V}^H$ is the equilibrium value of the high-type firm, as perceived by uninformed investors in the public market.

**Private Market**

The private market is opaque and less competitive in that quoted prices are not observable and there are no uninformed investors who can set the price competitively as in the public market. When the entrepreneur undergoes a liquidity shock she searches for an investor who will buy her firm. Recall that $e$ and $1 - e$ denote the probability that she is matched with an informed investor and an uninformed investor, respectively. In our baseline model, we consider the case in which $e = 1$. Later, we show our results hold when $e \neq 1$.

Once matched, the entrepreneur and the informed investors enter into a bilateral symmetric bargaining game. We model the result of the bargaining game in a reduced-form manner. Bilateral bargaining leads to a surplus sharing rule: entrepreneur gets $\beta$ of the total surplus and informed investors get $1 - \beta$. There are multiple ways to microfound this result. One simple way is to assume a take-it-or-leave-it bargaining game with alternating offers and let $\beta$ denote the probability that an entrepreneur gets the chance to make the offer. When hit by a liquidity shock, the entrepreneur has no outside option because private capital markets are relatively illiquid — it is costly and time-consuming to find another investor. Thus, if the entrepreneur refuses the offer, she obtains nothing, implying that with probability $1 - \beta$ the investor will offer the entrepreneur a price of zero.

Another way to microfound the bargaining game is to employ the classical infinite horizon bargaining game in the literature (e.g., Rubinstein (1982)). The entrepreneur (she) and the investor (he) take turns to make offers about how to divide the firm. At date 1, she proposes a split which he can accept or reject. If he accepts, the game ends. If he rejects, the game continues to date 2 when he proposes a split. The game will continue if no agreement is reached. Let her discounting factor be $\delta_E$ and his discounting factor be $\delta_I$. This game has a unique SPNE in which the pay-off

\footnote{$e < 1$ corresponds to the case in which informed investors are in short supply. In that case, we suppose that entrepreneurs are randomly matched to investors without modeling the matching game explicitly.}
of the entrepreneur is \( \frac{1-\delta}{1-\delta E\delta I} V^H \) and the pay-off of the investor is \( \frac{\delta I-\delta E\delta I}{1-\delta E\delta I} V^H \). We can then set
\[
\beta = \frac{1-\delta_I}{1-\delta E\delta I}
\]
to obtain the result.

In this setting, we see that in the private market, the good firm receives \( \beta V^H \) when a liquidity shock occurs, where \( V^H \) denotes the equilibrium level of the value of the firm in the private market, and the bad firm receives nothing with certainty. (Recall that there are no uninformed investors in the private market and informed investors receive a perfect signal as to the type of the firm.) In addition, because information acquisition in the private market occurs with certainty by paying a fixed cost, disclosure of any imperfectly informative signal has no effect on the entrepreneur’s payoff.\(^{10}\)

The timeline of the model is the following:

\[
\begin{align*}
&\text{\( t = 0 \)} & \text{\( t = 1 \)} & \text{\( t = 2 \)} & \text{\( t = 3 \)} & \text{\( t \)} \\
&\text{Entrepreneur sets up} & \tau \text{ and } i \text{ realize.} & \text{Liquidity shock occurs.} & V \text{ is publicly revealed} \\
&\text{the firm and choose} & \text{Investment decision} & \text{Trade happens and} & \\
&\text{which market to enter} & \text{by } \tau = H \text{ is made} & \text{prices are formed} & \\
& & s \text{ is released if the firm} & & \\
& & \text{goes public} & & \\
\end{align*}
\]

Figure 2.1: Timing of the Model

### 2.2.3 Definition of Equilibrium

The solution concept we adopt is Perfect Bayesian Equilibrium (PBE). A PBE is given by: the entrepreneur’s ownership structure choice \( \phi \), the \( H \) type firm’s production technology choice \( \lambda \), the uninformed investors’ pricing strategy \( P^{Pub} \) and their belief \( \hat{\lambda} \) about the technology choice in the public market, the potential informed investors’ entry decision, their pricing strategies in the public and private markets (\( P^{Pri} \)), such that:

\(^{10}\)For simplicity, we do not allow for an entrepreneur to switch from the private to the public market. One justification is that the IPO process is costly and lengthy, and when a liquidity shock occurs, the entrepreneur cannot file a “fire sale” IPO. After all, the going-public process takes substantial time and is subject to extensive SEC review. Another way to view our assumption is that our model applies only to settings in which the entrepreneur cannot easily switch from the private to public market, which we view as realistic in light of the substantial cost and planning required to go public.
1. given $\phi$, $\hat{\lambda}$ and $P^{Pub}$ (or $P^{Pri}$), $\lambda$ maximizes H type firm’s value within a given market.

2. given $\phi = Pub$ and $\hat{\lambda}$, the uninformed investors set $P^{Pub}$ to be the expected value of the firm given all public information available.

3. given $\phi = Pri$, the informed investors set $P^{Pri}$ to implement the outcome of the bargaining game.

4. $\phi$ is chosen optimally at $t = 0$.

5. $\hat{\lambda} = \lambda$ in equilibrium.

Since the choice of $\phi \in \{Pub, Pri\}$ is essentially a choice between two distinct subgames, we will first analyze each of these two subgames and then consider the decision to go public or private in the last step.

### 2.3 Equilibrium Analysis

In the following analysis, we assume that liquidity shocks are observable but not contractible. The ability to observe a liquidity shock means that a bad firm cannot strategically claim to be selling for liquidity reasons in the public market. We will later consider the case in which liquidity shocks are not observable.

#### 2.3.1 Public Market

If $\omega = 0$, i.e., there is no chance that the entrepreneur undergoes a liquidity shock at $t = 1$, then there is no need to rely on short-term signals. The entrepreneur can simply wait for terminal payoffs to be realized at $t = 3$ (recall there is no discounting). In that case, the high-type firm will set $\lambda = 1$ and the expected value of the firm at $t = 0$ is simply $E(V) = \frac{1}{2}(V_0 + I)$. We now consider the case where a liquidity shock happens with nonzero probability, i.e., $\omega \neq 0$. In that case, the short-term signal matters because it reveals information about the firm’s investment decision.
Price Formation

The trading stage occurs at $t = 2$ when the firm’s investment level $\lambda$ has been undertaken. Let $\hat{\lambda}$ denote the uninformed investors’ equilibrium belief about $\lambda$. With some abuse of notation, we will let $P$ denote the price in the public market.

Recall that if $s = G$, the type of the firm is revealed perfectly and thus the public market price is $P = V_0 + \hat{\lambda}I = \hat{V}^H$. If $s = B$, uninformed investors set the price using Bayes’ Rule and it can be shown that $P = \frac{\theta \hat{\lambda}^2}{1 + \theta \hat{\lambda}^2} \hat{V}^H$. Therefore, for a given level of $\lambda$ chosen by the entrepreneur,

\[
E(P|\tau = H) = (1 - \theta \lambda^2) \hat{V}^H + \theta \lambda^2 \frac{\theta \hat{\lambda}^2}{1 + \theta \hat{\lambda}^2} \hat{V}^H
\]

\[
E(P|\tau = L) = \frac{\theta \hat{\lambda}^2}{1 + \theta \hat{\lambda}^2} \hat{V}^H
\]

This derivation also shows that the public market is semi-strong efficient: the share price reflects all publicly available information.

Investment Choice

We now characterize the optimal production technology choice in the public market. A high-type firm solves the following problem:

\[
\max_{\lambda} \omega(1 - \frac{\theta \lambda^2}{1 + \theta \lambda^2}) \hat{V}^H + (1 - \omega)(V_0 + \lambda I)
\]

Lemma 2.2 Let $\omega^* = \frac{(1 + \theta)I}{(3\theta + 1)I + 2\theta V_0}$, there is a unique investment level $\lambda^*$ given by

\[
\lambda^* = 1, \quad if \ \omega \leq \omega^*
\]

\[
\lambda^* = r(\omega, \theta) < 1, \quad if \ \omega > \omega^*
\]
where \( r(\omega, \theta) \) is the root of the quadratic equation:

\[
(1 - 3\omega)\theta I \lambda^2 - 2\omega \theta V_0 \lambda + (1 - \omega)I = 0
\] (2.6)

\( r(\omega, \theta) \) is strictly decreasing in \( \omega \) and \( \theta \)

All proofs are in the appendix. The intuition is as follows. Increasing \( \lambda \) (i.e., the quality of real investment) has three effects. First, it increases firm’s terminal value at \( t = 3 \). Second, it increases the price the entrepreneur receives if a liquidity shock occurs, independent of the signal realization. But the third effect is that an increase in real investment also raises the probability of generating a poor short-term signal. And when that occurs, the high-value firm is pooled with the low-value firm at \( t = 2 \). When the probability of a liquidity shock is sufficiently high, first two positive effects dominate and \( \lambda^* = 1 \). Otherwise, the high-value firm chooses a suboptimal level of investment to avoid the pooling price.

With the above lemma, we can reorganize our result around the core parameter \( \theta \).

**Proposition 2.1** For a given \( \omega \), \( V_0 \) and \( I \), there exists \( \hat{\theta} \) such that:

\[
\lambda^* = 1, \ i f \ \theta \leq \hat{\theta} \tag{2.7}
\]

\[
\lambda^* = r(\omega, \theta) < 1 \ i f \ \theta > \hat{\theta} \tag{2.8}
\]

where \( \hat{\theta} \) is defined as:

\[
\hat{\theta} = 0, \ i f \ \frac{(1 - \omega)I}{(3\omega - 1)I + 2V_0\omega} \leq 0
\]

\[
\hat{\theta} = \frac{(1 - \omega)I}{(3\omega - 1)I + 2V_0\omega}, \ i f \ \frac{(1 - \omega)I}{(3\omega - 1)I + 2V_0\omega} \in (0, 1)
\]

\[
\hat{\theta} = 1, \ i f \ \frac{(1 - \omega)I}{(3\omega - 1)I + 2V_0\omega} = 1
\]

Figure 2.2 shows the result of the proposition with \( V_0 = 4 \) and \( I = 1 \) for 3 values of \( \omega \).
Now, if firm can commit not to disclose at $t = 0$, the objective function becomes:

$$
\max_{\lambda} \omega \frac{1}{2} \hat{V}^H + (1 - \omega)(V_0 + \lambda I)
$$

(2.9)

It is easy to see that $\lambda^* = 1$. We have the following corollary:

**Corollary 2.1** Committing to non-disclosure at $t = 1$ induces an optimal level of investment.

Here, non-disclosure eliminates the third effect mentioned above: when a high-value firm cannot be distinguished from a low-value firm at $t = 1$, it will invest efficiently.\textsuperscript{11}

**Expected Payoff in the Public Market**

The expected payoff to the entrepreneur in the public market can be written as:

$$
U^{Pub} = \frac{1}{2} \left[ \omega \frac{1}{1 + \theta \lambda^2} \hat{V}^H + (1 - \omega)\hat{V}^H \right] + \frac{1}{2} \left[ \omega I \frac{\theta \lambda^2}{1 + \theta \lambda^2} \hat{V}^H + (1 - \omega) \times 0 \right] - I
$$

$$
= \frac{1}{2} \hat{V}^H - I = \frac{1}{2}(V + \hat{\lambda} I) - I
$$

(2.10)

We can see that the low-value firm is actually better off when a liquidity shock occurs: uninformed investors do not know whether a poor signal arises because the firm is truly low-value or high-value subject to “bad luck”. The low-value firm benefits from being pooled with a good type. Under the assumption that liquidity shock is observable, low-value firms are unable to strategically sell when they are not hit by a liquidity shock. We will relax this assumption and consider strategic selling later.

It is also straightforward to see that committing not to disclose increases the firm’s ex-ante payoff. Figure 2.3 shows the gross payoff (excluding the investment cost $I$) to the entrepreneur in the public market when $V_0 = 4$ and $I = 1$. Panel (a) plots the payoff as a function of $\theta$ for different values of $\omega$. As we can see, the payoff is decreasing in $\omega$ and $\theta$. Fixing $\omega$, entrepreneur will choose the most efficient production technology for low levels of $\theta$. The payoff is then flat

\textsuperscript{11}For an extensive discussion of disclosure, see Goldstein and Yang (2017a) and Goldstein and Yang (2017b).
in that region. For high level of $\theta$, the concern of being pooled with a low type increases and therefore an entrepreneur will choose a suboptimal production technology, leading to a decreasing payoff. Panel (b) plots the payoff as a function of $\omega$ for different values of $\theta$ and can be interpreted similarly.

### 2.3.2 Private Market

In the private market, when the entrepreneur is hit by a liquidity shock and finds an informed investor, a bilateral bargaining game occurs. The bilateral bargaining generates the following surplus sharing rule: The entrepreneur gets $\beta$ proportion of the firm and the remaining $1 - \beta$ goes to the informed investor. We consider several ways to microfound this reduced form result later. It is straightforward to see that optimal investment level in the private market is $\bar{\lambda} = 1$. We summarize the results in the following proposition:

**Proposition 2.2** In the private market, the price the entrepreneur obtains following a liquidity shock is $\beta(V + I)$ if $\tau = H$ and $0$ if $\tau = L$. The entrepreneur chooses the optimal level of investment, i.e. $\bar{\lambda} = 1$.

The expected payoff in the private market is then:

$$ U^{Pri} = \frac{1}{2} \left[ \beta \omega VH + (1 - \omega) VH \right] + \frac{1}{2} \times 0 - I $$

$$ = \frac{1}{2} (V + I)(\beta \omega + (1 - \omega)) - I $$ (2.11)

Figure 2.4 shows the entrepreneur’s gross payoff (excluding the investment cost $I$) in the private market with $V_0 = 4$ and $I = 1$. Panel (a) plots the payoff as a function of $\beta$ for different values of $\omega$. As we can see, the entrepreneur’s payoff is increasing in $\beta$ and decreasing in $\omega$. Panel (b) plots the pay-off as a function of $\omega$ for different values of $\beta$ and can be interpreted similarly.
2.3.3 The Choice of Public vs. Private Markets

Recall that at $t = 0$, the entrepreneur chooses whether to join the public or private market. If $\beta = 1$, when $\lambda^* = 1$, the entrepreneur is indifferent between the two markets and when $\lambda^* < 1$, she strictly prefers the private market. We focus on the more interesting case where $\beta \neq 1$. Suppose $\beta = 0$. In that case, if $U^{Pri}$ is still weakly greater than $U^{Pub}$, the entrepreneur will stay in the private market for any value of $\beta$ and $\theta$. The condition can be written as:

$$U^{Pri} \geq U^{Pub} \iff \omega \leq (1 - \lambda) \frac{I}{V + I} \quad (2.12)$$

If $\theta \leq \hat{\theta}$, then $\lambda^* = 1$ and the entrepreneur will go public. The case with $\theta > \hat{\theta}$ is more involved. If $\omega + \hat{\lambda} \frac{I}{V_0 + I} < \frac{I}{V_0 + I}$, the entrepreneur will stay in the private market. Otherwise, she will choose the public market. We can summarize the entrepreneur’s decision at $t = 0$ as follows:

**Proposition 2.3** For a given $V_0$, $I$ and $\omega$,

1. If $\theta \leq \hat{\theta}$, the entrepreneur chooses the public market.

2. $\theta > \hat{\theta}$:
   - When $\omega + \hat{\lambda} \frac{I}{V_0 + I} < \frac{I}{V_0 + I}$, the entrepreneur chooses the private market.
   - When $\omega + \hat{\lambda} \frac{I}{V_0 + I} \geq \frac{I}{V_0 + I}$, there exists a $\tilde{\beta}$ such that $\frac{1}{2}(V_0 + I)(\tilde{\beta} \omega + (1 - \omega)) = \frac{1}{2}(V_0 + \hat{\lambda} I)$. Fix this $\tilde{\beta}$, entrepreneurs with $\theta > \hat{\theta}$ chooses the private market and entrepreneurs with $\theta \leq \hat{\theta}$ chooses the public market.

When $\theta$ is low, i.e., short-term signals reveal the long-run value of the firm, the firm goes public and chooses the efficient level of investment. These firms do not worry about being pooled with lemons because the probability of that happening is low. On the other hand, when $\theta$ is high, the firm will not efficiently invest when it is in the public market.

In other words, the choice between public and private markets depends on a combination of $\omega$ and $\theta$. Loosely speaking, if $\omega$ is low, the firm anticipates that the probability of a fire sale is low
and most of the time the entrepreneur will not have to share the gain from investment. Firms will then stay private in order to benefit from the efficient level of investment. When \( \omega \) is relatively high, the concern to sell the firm in the interim period grows and the choice of a public or private market also depends on the bargaining power \( \beta \). The firm chooses the private market only if it has strong bargaining power.

For a given value of \( V_0, I \) and \( \omega \), when \( \omega > \frac{I}{I+V_0} \), we can represent the equilibrium in the following graph \(^{12}\): the blue curve denotes the combination of \( \beta \) and \( \theta \) that makes the entrepreneur indifferent between the two markets. If \((\beta, \theta)\) lies below the curve, the entrepreneur will choose the public market. In addition, if \( \theta \) is low, the firm chooses the most efficient production technology in public market (i.e. if \( \theta < \hat{\theta}, \hat{\lambda} = 1 \)); otherwise, the firm conducts suboptimal investment (i.e. \( \hat{\lambda} < 1 \)).

In the same way, Figure 2.6 shows the indifference curve for two other values of \( \omega \).

### 2.4 Extensions and Robustness Check

#### 2.4.1 Short Supply of Informed Investors

The baseline model is derived under the simplifying assumption that \( e = 1 \). If informed investors are in short supply, then \( e < 1 \). As we noted previously, in the public market, informed investors never participate. That still holds if \( e < 1 \). Thus, we restrict our attention to the private market, in which there are now fewer informed investors. For consistency, we assume that uninformed investors in the private market behave similarly to those in the public market: they bid at a price equal to the expected value of the firm, conditional on available information. Since disclosure is not mandatory in the private market, the entrepreneur can commit not to disclose anything or to voluntarily disclose when she wishes. If she commits not to disclose, the high-value firm solves:

\[
\max_{\lambda, \omega} \omega(e\beta V^H + (1 - e)(1/2)^{V^H}) + (1 - \omega)V^H
\]

\(^{12}\)Here, we set \( V_0 = 4, I = 1 \) and \( \omega = 0.3 \).
It is immediate to see that $\lambda^{Pri} = \bar{\lambda} = 1$. Since the ex-ante value of the firm is given by:

$$U^{Pri} = \frac{1}{2} e \left[ \omega \beta V^H + (1 - \omega) V^H \right] + \frac{1}{2} (1 - e) \left[ \omega V^H + (1 - \omega) V^H \right]$$

we also see that committing not to disclose yields the highest firm value at $t = 0$. The result is also intuitive: by committing not to disclose, the entrepreneur receives half of the terminal value whenever she is matched with an uninformed investor. Hence, she might as well choose the efficient production technology to boost the final value of the project.

Next, we consider the possibility that the firm cannot commit to non-disclosure. In that case, the entrepreneur voluntarily discloses the signal. Recall that in the private market, if the firm is matched with informed investor (which occurs with probability $e$), the disclosure decision is irrelevant. If the firm is matched with an uninformed investor, it ends up in a situation identical to that of the public market. For this reason, at $t = 1$, the high-value firm’s problem is given by:

$$\max_{\lambda} \ e \left[ \omega \beta V^H + (1 - \omega) V^H \right] + (1 - e) \left[ \omega (1 - \frac{\theta \lambda^2}{1 + \theta \lambda^2}) V^H + (1 - \omega) V^H \right]$$

We summarize the result of this problem and the the result we obtained in the full commitment case in the following proposition.

**Proposition 2.4** When informed investors are in short supply, i.e., $e < 1$, the firm is better off when it can commit not to disclose. In this case, the ex-ante value of the firm is also maximized. When the firm cannot commit, it always chooses a more efficient production technology in the private market.

The decision to stay private or to go public still relies on the comparison between $U^{Pri}$ and $U^{Pub}$, which is similar to proposition 2.3.
2.4.2 Strategic Selling When a Liquidity Shock is Not Observable

If liquidity shocks are not observable, a firm in the public market may choose to sell even if it is not suffering from a liquidity shock. Consider first a low-value firm. Since a high-value firm suffers from a liquidity shock with positive probability, the public market price at $t = 1$ will be positive even when the short-term signal is poor. Low-value firms therefore always choose to sell (because a positive price at $t = 1$ exceeds the long-run payoff of zero). Realizing this, a high-value firm with $s = B$ never sells unless it truly suffers from a liquidity shock, because the price it will receive is too low. When $s = G$, the choice is irrelevant because the price at $t = 1$ is fair, i.e., accurately separates the high- and low-value firms.

This implies that when $s = B$, the price in the public market is given by:

$$ P = \frac{\theta \lambda^2 \omega}{1 + \theta \lambda^2 \omega} V_H $$

(2.16)

And

$$ E(P|\theta = H) = (1 - \theta \lambda^2) V_H + \theta \lambda^2 \frac{\theta \lambda^2 \omega}{1 + \theta \lambda^2 \omega} V_H = (1 - \frac{\theta \lambda^2}{1 + \theta \lambda^2 \omega}) V_H $$

(2.17)

At the investment stage, the firm choose $\lambda$ to solve:

$$ \max_{\lambda} \omega(1 - \frac{\theta \lambda^2}{1 + \theta \lambda^2}) V_H + (1 - \omega)(V_0 + \lambda I) $$

(2.18)

We summarize the result in the following lemma.

**Lemma 2.3** Let $\omega^*$ be the solution between $[0, 1]$ that solves $-\theta I \omega(1 + \omega) - 2\theta \omega V_0 + (1 - \omega)I = 0$, there is a unique investment level $\lambda^*$ given by

$$ \lambda^* = 1, \ if \ \omega \leq \omega^* $$

(2.19)

$$ \lambda^* = R(\omega, \theta) < 1 \ if \ \omega > \omega^* $$

(2.20)
where $R(\omega, \theta)$ is the root of the quadratic equation:

$$-\theta I \omega (1 + \omega \lambda^2) - 2\theta \omega V_0 \lambda + (1 - \omega) I = 0$$

(2.21)

$R(\omega, \theta)$ is strictly decreasing in $\omega$ and $\theta$.

Using this lemma, the analysis as to the choice between a public and private market is similar to the baseline model. We reach the same qualitative conclusion.

2.4.3 Outside Financing

In the baseline model, we assume that the entrepreneur can self-finance the firm’s project. We will show here that our results are unchanged if the entrepreneur is capital-constrained and needs outside financing. Since we focus here on the firm’s IPO decision, we consider equity financing: at $t = 0$, entrepreneur sells $\alpha$ proportion of the firm to the investors in a given market, such that investors break even. We first consider the case where outside investors also suffer from the same liquidity shocks as the firm (e.g., macroeconomic risk) and then the case where they do not and can hold the firm to its maturity.

Outside Investors With Liquidity Shocks

In this extension, we assume that the entrepreneurs are capital-constrained and must raise financing from outside investors. The outside investors suffers from liquidity shocks with probability $\omega$ just as the entrepreneurs do. We will show that our results hold under these conditions as well.

Notice that the entrepreneur’s decision after $I$ is raised remains the same as in the baseline model. For outside investors to break even in the public market, the following condition must hold, letting $\alpha_1$ denote the share of equity sold in the public market:

$$\alpha_1 \left[ \frac{1}{2} \left( \omega \overline{\nu}^H (1 - \frac{\theta \lambda^2}{1 + \theta \lambda^2}) + (1 - \omega ) \overline{\nu}^H \right) + \frac{1}{2} \omega \frac{\theta \lambda^2}{1 + \theta \lambda^2} \right] = \frac{1}{2} \alpha_1 \overline{\nu}^H = I$$

(2.22)
The payoff to the entrepreneur is then:

\[
U^{Pub} = (1 - \alpha_1) \left[ \frac{1}{2} \left( \omega \bar{V}^H (1 - \frac{\theta \lambda^2}{1 + \theta \lambda^2}) + (1 - \omega) \bar{V}^H \right) + \frac{1}{2} \omega \frac{\theta \lambda^2}{1 + \theta \lambda^2} \right]
\]

\[
= \frac{1}{2} (1 - \alpha_1) \bar{V}^H = \frac{1}{2} \bar{V}^H - I
\]

(2.23)

For an investor in the private market to break even, the following condition must hold, letting \( \alpha_2 \) denote the share of equity sold in the private market:

\[
\alpha_2 \left[ \frac{1}{2} (\beta \omega + (1 - \omega)) \bar{V}^H + \frac{1}{2} \times 0 \right] = I
\]

The payoff to the entrepreneur is then:

\[
U^{Pri} = (1 - \alpha_2) \left[ \frac{1}{2} (\beta \omega + (1 - \omega)) \bar{V}^H + \frac{1}{2} \times 0 \right]
\]

\[
= \frac{1}{2} (\beta \omega + (1 - \omega)) \bar{V}^H - I
\]

(2.24)

The payoff functions are identical to the baseline model and the analysis remains the same.

**Outside Investors Without Liquidity Shocks**

If investors are patient and can hold the firm until the end of the game, we can write the break-even condition in the public market as:

\[
\frac{1}{2} \alpha_1 \bar{V}^H = I
\]

And the payoff to the entrepreneur is:

\[
U^{Pub} = (1 - \alpha_1) \left[ \frac{1}{2} \left( \omega \bar{V}^H (1 - \frac{\theta \lambda^2}{1 + \theta \lambda^2}) + (1 - \omega) \bar{V}^H \right) + \frac{1}{2} \omega \frac{\theta \lambda^2}{1 + \theta \lambda^2} \right]
\]

\[
= \frac{1}{2} (1 - \alpha_1) \bar{V}^H = \frac{1}{2} \bar{V}^H - I
\]

(2.25)
Similarly, in the private market:

$$\frac{1}{2} \alpha_2 \overline{V^H} = I$$

And the payoff to the entrepreneur is then:

$$U_{Pri} = (1 - \alpha_2) \left[ \frac{1}{2} (\beta \omega + (1 - \omega)) \overline{V^H} + \frac{1}{2} \times 0 \right]$$

$$= \frac{1}{2} (\beta \omega + (1 - \omega)) \overline{V^H} - I (\beta \omega + (1 - \omega))$$  \hspace{1cm} (2.26)

All else equal, the payoff in the private market is higher. Assuming $\omega > \frac{I}{I+V_0}$ and $\hat{\theta} \in [0, 1]$, we can solve for $\bar{\beta}^O$ which makes the entrepreneur indifferent between the two markets. This $\bar{\beta}^O$ solves the following equation:

$$\frac{1}{2} (V_0 + I) \left( \bar{\beta}^O \omega + (1 - \omega) \right) - I \left( \bar{\beta}^O \omega + (1 - \omega) \right) = \frac{1}{2} (V_0 + \lambda I) - I$$  \hspace{1cm} (2.27)

or equivalently,

$$\bar{\beta}^O = \left( \frac{V_0 + \lambda I - 2I}{V_0 + I - 2I} - (1 - \omega) \right) \frac{1}{\omega}$$  \hspace{1cm} (2.28)

It is easy to see that $\bar{\beta}^O \leq \bar{\beta}$, and equality is obtained when $\lambda = 1$.

### 2.5 Mandatory Disclosure and Real Inefficiency

To consider the efficiency implications of mandatory disclosure, we begin by defining the following measure of price informativeness:

$$-E [\Lambda(P)] = -E \left[ \frac{Var(V|P)}{Var(V)} \right]$$  \hspace{1cm} (2.29)
This is the negative of the variance of final value conditional on the price (as in Kyle (1985)), scaled by the unconditional variance.

We now compute this measure in public market under the baseline model. It is straightforward to see that $Var(V) = \frac{1}{4} V^2$. When $s = G$, the price is fully revealing and $Var(V|P) = 0$. When $s = B$, $P = \frac{\theta \lambda^2}{1 + \theta \lambda^2} V^H$. The conditional variance is then:

$$\frac{\theta \lambda^2}{1 + \theta \lambda^2} (V^H - \frac{\theta \lambda^2}{1 + \theta \lambda^2} V^H)^2 + \frac{1}{1 + \theta \lambda^2} (0 - \frac{\theta \lambda^2}{1 + \theta \lambda^2} V^H) = \frac{\theta \lambda^2}{(1 + \theta \lambda^2)^2} V^H^2$$

(2.30)

Therefore,

$$-E[\Lambda(P)] = -(1 - \theta \lambda^2) * 0 - \theta \lambda^2 \frac{\theta \lambda^2}{(1 + \theta \lambda^2)^2} = \left( -\frac{2\theta \lambda^2}{1 + \theta \lambda^2} \right)^2$$

(2.31)

Clearly, disclosure benefits shareholders of high-value firms at $t = 1$. To see why, suppose that a firm can commit not to disclose. In that case, price informativeness is simply $-1$ because no information is conveyed in prices at $t = 1$. Consider a shareholder of a high-value company who wishes to sell shares. If mandatory disclosure is required, she will receive a fair price when the short-run signal is good and an undervalued price when the signal is poor. The net loss is simply $-\frac{\theta \lambda^2}{1 + \theta \lambda^2} V^H$. If mandatory disclosure is not required and a firm commits not to disclose, her share will always be undervalued and the net loss will be $-\frac{1}{2} (V_0 + I)$, which is always bigger than $-\frac{\theta \lambda^2}{1 + \theta \lambda^2} V^H$. This result illustrates the core intuition of much of the disclosure literature: disclosure enhances price accuracy, which benefits shareholders who need to sell prior to realization of the terminal payoffs.

However, we now show that disclosure nonetheless can induce a real efficiency loss. Suppose, as in the prior section, that the investor, after financing the entrepreneur’s project, is subject to a liquidity shock with probability $\omega$ and can hold the share to the final period with probability $1 - \omega$.

First, we consider the high-value firm to develop intuitions. Since the investor in a high-value firm

---

13Voluntary disclosure leads to the same price informativeness as mandatory disclosure. Hence we consider the case in which firm can commit not to disclose.
observes the firm’s type when she needs to sell in response to a liquidity shock, the expected net loss from the sale with disclosure is:

\[ E(P|\tau = H) - (V_0 + \hat{\lambda}I) = -\frac{\theta\hat{\lambda}^2}{1 + \theta\hat{\lambda}^2} \bar{V}^H \]  

(2.32)

Similarly, the expected net loss from the sale without disclosure is:

\[ \frac{1}{2}(V_0 + I) - (V_0 + I) = -\frac{1}{2}(V_0 + I) \]  

(2.33)

Clearly, the latter is simply the pooling price. The difference in the total payoff can be simply written as:

\[
\omega \left[ -\frac{\theta\hat{\lambda}^2}{1 + \theta\hat{\lambda}^2} \bar{V}^H + \frac{1}{2}(V_0 + I) \right] + (1 - \omega) \left[ V_0 + \hat{\lambda}I - V_0 - I \right]
\]

(2.34)

This equation highlights the tradeoff for high-value firms in the public market. Mandatory disclosure increases price informativeness at \( t = 1 \). Information disclosure help to distinguish high-value firms from low-value ones because low-value firms are more likely to generate poor signals. However, price informativeness comes at a price: disclosure puts high-value firms at the risk of being pooled with low-value ones. As a result, high-value firms may forego efficient investments. Whether the difference in the payoff is positive or negative for high-value firms depends on the parameters. For example, if \( \omega \) is large, the difference is positive.

Moreover, one can show that \( \frac{\partial \hat{\lambda}^2}{\partial \theta} = \hat{\lambda}^2 + 2\theta \hat{\lambda} \frac{\partial \hat{\lambda}}{\partial \theta} = 0 \). This implies that an increase in the \( \theta \) has two effects. A direct effect which leads to a direct increase in \( \frac{\theta\hat{\lambda}^2}{1 + \theta\hat{\lambda}^2} \) and an indirect effect working thorough \( \hat{\lambda} \) and then \( \bar{V}^H \). The indirect effect leads to a decrease in \( \frac{\theta\hat{\lambda}^2}{1 + \theta\hat{\lambda}^2} \) and \( \bar{V}^H \). On net, an increase in \( \theta \) raises the gains from disclosure but also increase the loss from inefficient investment. The result is quite intuitive. The gain from disclosure is measured as the divergence between the price at \( t = 1 \) and the long-run fundamental value. As \( \theta \) increases, although the level of real investment decreases, the divergence between the price and that lower level of real investment also shrinks.
Now, consider low-value firms. Since the investor understands the firm is low-value when she needs to sell her shares, the expected gains from a sale in response to a liquidity shock is given by:

$$E(P|\tau = L) - 0 = \frac{\theta \hat{\lambda}^2}{1 + \theta \hat{\lambda}^2} \hat{V}^H$$  \hspace{1cm} (2.35)

Similarly, the expected gain from a sale with no disclosure is:

$$\frac{1}{2} (V_0 + I) = \frac{1}{2} (V_0 + I)$$  \hspace{1cm} (2.36)

In other words, the low-value firm benefits from the pooling that results in the absence of disclosure. Since the low-value firm has no investment opportunity, the difference in the total payoff can be simply written as:

$$\omega \left[ \frac{\theta \hat{\lambda}^2}{1 + \theta \hat{\lambda}^2} \hat{V}^H - \frac{1}{2} (V_0 + I) \right]$$  \hspace{1cm} (2.37)

Intuitively, low-value firms always prefer non-disclosure: disclosure leads to some nonzero probability of separation between high- and low-value firms.

Finally, consider the expected difference in the payoff to disclosure vs. non-disclosure at $t = 0$ when firm’s type is unknown. The expected difference in the total payoff is:

$$\frac{1}{2} \omega \left[ -\frac{\theta \hat{\lambda}^2}{1 + \theta \hat{\lambda}^2} \hat{V}^H + \frac{1}{2} (V_0 + I) \right] + \frac{1}{2} (1 - \omega) \left[ V_0 + \hat{\lambda}I - V_0 - I \right]$$

$$+ \frac{1}{2} \omega \left[ \frac{\theta \hat{\lambda}^2}{1 + \theta \hat{\lambda}^2} \hat{V}^H - \frac{1}{2} (V_0 + I) \right] = \frac{1}{2} (1 - \omega)(\hat{\lambda} - 1)I < 0$$  \hspace{1cm} (2.38)

As this expression shows, mandatory disclosure is value-destroying.

### 2.6 Conclusion

The literature on mandatory disclosure has long hailed the benefits of transparency for price accuracy. Mandatory disclosure is essential for capital markets to function efficiently. But certain
disclosure metrics that now make up much of modern securities law may cause more harm than good. When the value of the best projects is poorly reflected by these metrics, and it is too costly for investors to look past these simplifications, these mandates can lead entrepreneurs to pursue lower-value projects.

More fundamentally, our project shows that price accuracy may itself come at a price: distorting real investment. When short-run signals are poor proxies for long-run fundamental value, and observable order flow limits the expected profit to information acquisition, entrepreneurs may abandon high-value projects to avoid temporary mispricing. The flight of capital to the private markets over the past decade highlights the importance of understanding whether and when transparency comes at a cost. Our framework counsels against the adoption of mandatory metrics which are noisy proxies for fundamental value.
This figure shows the firm’s optimal investment scale in public market. $\lambda$ is the investment scale. $\theta$ is the information gap between short-term disclosed signal and the firm’s fundamental value. $\omega$ is the probability of having a liquidity shock.
This figure shows the firm’s expected pay-off ($U^{\text{pub}}$) in public market under different parameter values. $\theta$ is the information gap between short-term disclosed signal and the firm’s fundamental value. $\omega$ is the probability of having a liquidity shock. Panel (a) plots the expected pay-off as a function of $\theta$. Panel (b) plots the expected pay-off as a function of $\omega$. 
This figure shows the firm’s expected pay-off ($U_{pri}$) in private market under different parameter values. $\beta$ measures the firm’s bargaining power in private market. $\omega$ is the probability of having a liquidity shock. Panel (a) plots the expected pay-off as a function of $\beta$. Panel (b) plots the expected pay-off as a function of $\omega$. 
Figure 2.5: Equilibrium When $\omega = 0.3 > \frac{1}{\tau + \nu_o} = 0.2$

This figure shows the equilibrium under certain parameter values. $\beta$ measures the firm’s bargaining power in private market. $\lambda$ is the investment scale. $\theta$ is the information gap between short-term disclosed signal and the firm’s fundamental value. The blue line is the indifference curve on which the firm is indifferent between going public or staying private.
This figure shows the equilibrium under different \( \omega \). \( \beta \) measures the firm’s bargaining power in private market. \( \theta \) is the information gap between short-term disclosed signal and the firm’s fundamental value. Lines are the indifference curves on which the firm is indifferent between going public or staying private. The firm prefers to stay private in regions above the curve.
2.8 Proofs

Lemma 2.2

Recall that a high-value firm solves the following problem:

$$\max_\lambda \omega (1 - \frac{\theta \lambda^2}{1 + \theta^2}) \widehat{V}^H + (1 - \omega)(V_0 + \lambda I)$$  \hspace{1cm} (2.39)

The F.O.C is given by:

$$-2\omega \widehat{V}^H \theta \frac{\lambda}{1 + \theta^2} + (1 - \omega)I$$  \hspace{1cm} (2.40)

And it is also clear to see that the S.O.C is negative. Hence the optimal level of $\lambda$ is completely characterized by the sign of the F.O.C. We can rewrite the F.O.C in the following way:

$$\frac{f(\lambda)}{1 + \theta \lambda^2}$$  \hspace{1cm} (2.41)

where $f(\lambda) = \lambda^2(1 - 3\omega)I \theta - 2\lambda \omega \theta V_0 + (1 - \omega)I$. The sign of F.O.C is determined by $f(\lambda)$. We now discuss in detail. Define $\Delta = 4\omega^2 \theta^2 V_0^2 - 4(1 - 3\omega)(1 - \omega)I^2 \theta$

**case 1: $\omega > \frac{1}{3}$**

In this case, $\Delta > 0$. $f(\lambda) = 0$ has root(s). Notice that $f'(0) < 0$ and that $f(\lambda)$ is inverted $U$ shape, we can conclude that it must have one positive and only one positive root. The positive root is smaller than 1 if and only if

$$f(1) < 0 \Leftrightarrow \omega > \frac{(1 + \theta)I}{(3\theta + 1)I + 2\theta V_0} = \omega^*$$  \hspace{1cm} (2.42)

Therefore, $\lambda^* = 1$ if $\omega \leq \omega^*$ and $\lambda^* = r(\omega, \theta)$ if $\omega > \frac{(1 + \theta)I}{(3\theta + 1)I + 2\theta V_0} = \omega^*$

**case 2: $\omega \leq \frac{1}{3}$**

In this case, suppose $f'(1) > 0$ which implies that $\omega < \frac{1}{3 + \frac{\theta}{V_0}}$. We will have $\frac{1 - 3\omega}{\omega} > \frac{V}{I}$. Since
\[
\frac{1 - \omega}{\omega} > \frac{1 - 3\omega}{\omega}, \text{ we can have }
\]
\[
\frac{1 - \omega}{\omega} - 3\omega \geq \frac{V_0^2}{I^2} \geq \frac{\theta V_0^2}{I^2}
\]
which implies that \( \Delta < 0 \) and \( f(\lambda) > 0 \). Hence we have \( \lambda^* = 1 \).

Suppose \( f'(1) \leq 0 \) \( (\omega \geq \frac{1}{3 + \frac{V_0}{I}}) \). Notice that \( f(0) > 0, f'(0) < 0 \) and \( f(\lambda) \) is \( U \) shaped. Then \( f(\lambda) \) has one and only one positive root between \((0, 1)\) if and only if \( f(1) < 0 \). This implies that \( \lambda^* = 1 \) if \( \omega \leq \omega^* \) and \( \lambda^* = r(\omega, \theta) \) if \( \omega > \frac{(1 + \theta)I}{(3\theta + 1)I + 2\theta V_0} = \omega^* \).

One can verify that \( \omega^* > \frac{1}{3 + \frac{V_0}{I}} \) and thus we obtain the first part of the proposition. Now we use implicit function theorem to derive the second part of the proposition. Simple computation yields:

\[
\frac{\partial f(\lambda)}{\partial \lambda} = 2(1 - 3\omega)\theta I\lambda - 2\omega V_0
\]
(2.44)
\[
\frac{\partial f(\lambda)}{\partial \omega} = -3\theta I\lambda^2 - 2V_0\theta \lambda - I
\]
(2.45)
\[
\frac{\partial f(\lambda)}{\partial \theta} = (1 - 3\omega)I\lambda^2 - 2\omega V_0 \lambda
\]
(2.46)

If \( \omega > \frac{1}{3} \), it is easy to see that \( \frac{\partial f(\lambda)}{\partial \lambda} < 0 \). If \( \omega < \frac{1}{3} \), recall that at interior solution (if there is one), it must be \( f'(0) < 0, f'(1) < 0 \), so \( \frac{\partial f(\lambda)}{\partial \lambda} < 0 \). Hence

\[
\frac{\partial \lambda}{\partial \omega} = -\frac{\partial f}{\partial \omega} \bigg|_{r < 0}
\]
(2.47)

similarly, we can have \( \frac{\partial \lambda}{\partial \theta} < 0 \).

**Proposition 2.1**

For a given \( \omega \), we define \( \bar{\theta} \) as \( \frac{(1 + \bar{\theta})I}{(3\bar{\theta} + 1)I + 2\bar{\theta} V_0} = \omega \). Notice that \( \bar{\theta} \) is decreasing in \( \omega \) and \( \bar{\theta} \) should lie between \([0, 1]\). The result follows.

**Proposition 2.4**

The first part of the proposition follows directly from the text. For the non-commitment part, we...
examine the maximization problem. Taking derivative with respect to \( \lambda \) leads to

\[
e[\omega \beta I + (1 - \omega)I] + (1 - e) \frac{f(\lambda)}{1 + \theta \lambda^2}
\]  

(2.48)

At an interior solution where \( \lambda^* < 0 \), we have \( f(\lambda^*) = 0 \). Since \( e[\omega \beta I + (1 - \omega)I] + (1 - e) \frac{f(\lambda)}{1 + \theta \lambda^2} \) is still positive, hence we have \( \lambda \geq \lambda^* \)

Notice that when \( \omega \to 1, \beta \to 0 \), \( f(1) \to -2I\theta - 2\theta V_0 \) and \( e[\omega \beta I + (1 - \omega)I] + (1 - e) \frac{f(\lambda)}{1 + \theta \lambda^2} \to -2(1 - 2) \frac{I\theta + \theta V_0}{1 + \theta \lambda^2} < 0 \). By continuity, \( \exists \) some region of \((\omega, \beta)\) such that \( \bar{\lambda} < 1 \).

**Proposition 2.3**

The F.O.C of the maximization problem can be written as:

\[
-2\theta \omega \bar{V}^H \lambda + (1 - \omega)I(1 + \omega \lambda^2) = 0
\]  

(2.49)

In equilibrium, we can define:

\[
g(\lambda) = -2\theta \omega \bar{V}^H \lambda + (1 - \omega)I(1 + \omega \lambda^2) = -\theta I\omega(1 + \omega)\lambda^2 - 2\theta \omega V_0 \lambda + (1 - \omega)I
\]  

(2.50)

Notice that \( g(0) > 0 \) and \( g'(0) < 0 \), \( g(0) \) must have one and only one positive root. The root lies between \([0, 1]\) iff \( g(1) \leq 0 \).

Notice further that \( g(1, \omega = 1) < 0 \) and \( g(1, \omega = 0) = I > 0 \) and \( \frac{\partial g(1, \omega)}{\partial \omega} < 0 \), \( \exists \) a unique \( \omega^{**} \) such that \( g(1, \omega^{**}) = 0 \). Then it follows directly that

\[
\text{if } \omega > \omega^{**}, \ g(1, \omega) < 0 \quad \text{and} \quad \lambda^{**} = R(\omega, \theta) < 1
\]  

(2.51)

\[
\text{if } \omega \leq \omega^{**}, \ g(1, \omega) \geq 0 \quad \text{and} \quad \lambda^{**} = 1
\]  

(2.52)
Furthermore,

\[
\frac{\partial g}{\partial \lambda} = -2\theta I\omega (1 + \omega)\lambda - 2\theta \omega V_0 < 0 \quad (2.53)
\]

\[
\frac{\partial g}{\partial \omega} = -\theta I\lambda^2 (1 + 2\omega) - 2\theta V_0 \lambda < 0 \quad (2.54)
\]

\[
\frac{\partial g}{\partial \theta} = -I\omega (1 + \omega)\lambda^2 - 2\omega V_0 \lambda < 0 \quad (2.55)
\]

The implicit function theorem implies that \( \lambda \) is decreasing in \( \omega \) and \( \theta \).
3.1 Introduction

The growth of passive fund management has been explosive. Passive funds typically track an index (e.g., SP 500) and charge low fees for investing. They have experienced a growth of 25% per year between 2009 and 2014. PwC predicts the asset under management in passive funds to reach 36.6 trillion dollars by 2025 (25% of global assets under management).

Despite its fast growth, theoretical analyses on the effects of passive investing are still nascent. In this chapter, we provide a benchmark model to analyze investors’ equilibrium choices and the welfare consequences of active and passive investing.

Active investing is costly, but it brings two benefits: investors can better hedge by freely trading each asset in the portfolio and can acquire information about the possible state of the world. Information acquisition decisions are strategic substitutes. Equilibrium choices between active and passive investing are determined by a set of indifference conditions: investors will become active until the net value of being active shrinks to zero. While individual investors care about costs and benefits at the level, a social planner cares about the marginal gains of being active. Hence, market equilibria may induce sub-optimal levels of active investing.

To illustrate these points, we consider a three-period, complete-market, pure-exchange economy with three assets and three states of the world. The economy is populated by competitive risk-averse agents maximizing their expected utility over final wealth. At the final date 3, investors

\footnote{This chapter is based on Bianchi et al. (2020).}
are hit by a liquidity shock in one state. At date 2, investors privately learn about the realization of this shock and have the opportunity to trade: the liquidity shocks create gains from trade. At date 1, investors choose their investment strategies and their initial portfolios.

There are three different investment profiles. First, investors can invest in passive funds. Once invested, they cannot change the weights on the risky assets in the portfolio but can only scale up or down their investments. Second, they can pay costs and invest in an active fund. Active funds can freely rebalance each asset in the portfolios. On top of that, some active fund managers acquire private information on the realization of the state. Hence, the investors may either invest in an uninformed fund (we call it uninformed but unconstrained), or an informed fund at an additional cost. Information plays a dual role in our model. On the one hand, it generates adverse selection, which has a redistributive effect and may lower the realization of gains from trade. On the other hand, its diffusion via market efficiency enables investors to better rebalance their portfolios to hedge the liquidity shocks that they randomly receive at the final date.

We start by fixing the proportion of different investors and determine the optimal demand at date 2. Due to his trading constraint, the passive investor’s demand is suboptimal: he demands too much on one risky asset and too little of the other. In the absence of financial intermediaries, active investors would be their natural counterparts and absorb the residual demand. However, prices incorporate a risk premium required by active investors to deviate from their optimal allocation. The presence of financial intermediaries lowers the magnitude of this premium. Also, prices include an adverse selection component that increases with the proportion of informed investors.

We then determine market prices using market clear conditions. We show that when aggregate hedging demand (due to liquidity shocks) and speculation demand (due to private information) go in the opposite direction, prices do not reveal investors’ private information. Intuitively, uninformed investors cannot distinguish a situation in which hedging demands are high but speculation demands are low, and a situation in which hedging demands are low but speculation demands are high: aggregate demand pressure and prices are the same. Under non-fully revealing equilibrium, the benefits of being active come from both being informed and being able to rebalance the portfo-
lio freely. We define the value of information as the pay-off difference between informed investors and uninformed but unconstrained investors. We define the value of trading flexibility as the pay-off difference between uninformed but unconstrained investors and passive investors. We show that the value of information decreases with the proportion of active investors, while the value of trading flexibility increases with the proportion of active investors.

We characterize the full set of equilibria under the non-fully revealing condition. We show that when the cost of acquiring information is sufficiently low, there will not be unconstrained but uninformed investors: investors are better off paying a small additional cost to become informed. An indifference condition determines the equilibrium proportion of informed and passive investors: investors will become informed until the value of information shrinks to zero. When the cost of information rises, we could have an equilibrium in which all three types of investors co-exits. Finally, if the cost being unconstrained is sufficiently low, we will have an equilibrium with no passive investors. In this case, being unconstrained strictly dominates being passive. The pay-off functions and equilibria that we characterize allow us to study the optimality of passive and active investing. As individual investors only care about the gains in the level, not marginal gains of being informed, equilibria may feature under-information-acquisition.

The major contribution of this chapter is to provide a theoretical framework so that analyzing equilibrium, welfare consequences and optimality of active and passive investing becomes possible. Related previous research is often empirical. Researchers find that passive investing, or ETF investing, may enhance market efficiency by facilitating informed traders to hedge industry risks (Huang et al. (2018)), increase the nonfundamental volatility of the securities in their baskets (Ben-David et al. (2018)) and increase the bid-ask spread of underlying stocks (Israeli et al. (2017)). Theoretical analyses help us to view the empirical findings in a unified framework and also allow us to draw other empirical implications.

Bhattacharya and O’Hara (2018) models the introducing of ETFs and studies how it affects the informational efficiency of underlying markets and introduces fragility via herding. Relative to their model, we do not introduce noise traders. We model behaviors of all market participants,
and this allows welfare analyses of all players. More broadly, modeling all market participants
distinguish our paper from canonical models of rational expectation, started by Grossman and
Stiglitz (1980), Hellwig et al. (1980) and Admati (1985). Exogenous noise traders in these models
make it difficult to conduct general equilibrium and welfare analyses.

The most related concurrent paper is Bond and Garcia (2019). They focus on how reducing the
cost of indexing affects price efficiency and passive investors’ welfare. In their paper, individuals’
indexing participation exhibits strategic complementarity. A reduction in the cost encourages more
indexing, making the aggregating price less informative. This, in turn, encourages risk-sharing and
improves passive investors’ welfare. Our focus is different. Instead of focusing on one particu-
lar equilibrium, we characterize the full set of equilibria. We identifies conditions under which
passive investing co-exists with active investing and conditions under which passive investing is
dominated. Our analysis allows us to contrast different equilibria with social optimum and high-
light that market equilibria may induce sub-optimal level of information acquisition.

The rest of the paper is organized as follows. Section 3.2 lays out our model. Section 3.3 solves
the optimal demand. Section 3.4 characterizes the equilibria and Section 3.5 concludes.

### 3.2 Model

Time is discrete, denoted with \(t = 1, 2, 3\). There are three states of nature indexed by \(n\),
\(\tilde{\omega} \in \{\omega_1, \omega_2, \omega_3\}\). Agents share common priors about \(\tilde{\omega}\),

\[
\Pr(\tilde{\omega} = \omega_1) = \Pr(\tilde{\omega} = \omega_2) = \frac{1 - \eta}{2}.
\]  

(3.1)

There are three assets, a bond and two stocks, denoted with \(k = \{X, Y\}\). The bond pays 1 in each
state of nature, and is taken as the numeraire (cash). One share of stock \(k\) pays a dividend \(\tilde{d}_k\). We
assume that \(\tilde{d}_X = 1\) if \(\tilde{\omega} = \omega_1\) and 0 otherwise; while \(\tilde{d}_Y = 1\) if \(\tilde{\omega} = \omega_2\) and 0 otherwise.
Investors  We consider a continuum of competitive risk-averse investors of mass 1. At $t = 1$, investors can decide to become active or passive. Becoming an active investor requires paying a cost $c_T$ and it allows to freely rebalance one’s portfolio. A passive investor instead cannot change the portfolio weights on the stocks, he can only scale up or down his risky portfolio against the bond. In addition, again at $t = 1$, investors can decide to become informed by paying a cost $c_I$. In this case, an investor receives a signal at date 2 on the realization of the state of the world. Hence, we consider three types of investors: informed investors, who are both unconstrained and informed; uninformed investors, who are unconstrained but uninformed; and passive investors, who are uninformed and subject to the above trading constraint. A generic investor is indexed by $r$. When we need to distinguish his type, we index informed investors by $i$, uninformed investors by $u$, and passive investors by $j$.

Let $e^r_C, e^r_X$ and $e^r_Y$ denote respectively the initial endowment of cash, asset $X$ and asset $Y$ by investor $r$. Assume that $e^r_C > c_I + c_T$ so each investor has enough cash to become informed if he wishes to.

Information set  At the beginning of date 2, informed investors have access to a signal $\bar{s} \in \{s_1, s_2, s_3\}$ such that

$$\begin{align*}
\Pr(\bar{s} = s_n|\bar{\omega} = \omega_n) &= \tau, \\
\Pr(\bar{s} = s_n|\bar{\omega} \neq \omega_n) &= \frac{1-\tau}{2},
\end{align*}$$

(3.2)

where $\tau \in (\frac{1}{3}, 1)$ is the signal’s precision. Let us denote $q^r_n$ the probability that investor $r$ attaches to state $\omega_n$ given his information set at date 2. We assume that all active investors receive the same signal at date 2. Hence, we can express the posterior probabilities simply as $q^i_n = q^a_n$ for all informed investors $i$ and $q^j_n = q^u_n = q^p_n$ for all passive or uninformed investors $j, u$. Denote with $\alpha$ the equilibrium proportion of informed investors, with $\upsilon$ the proportion of uninformed investors and with $\pi$ the proportion of passive investors.

Liquidity shock  At date 3, investor $r$ is hit by a liquidity shock $\bar{l}^r = (\bar{l}^r_1, \bar{l}^r_2, \bar{l}^r_3)$ where $\bar{l}^r_n$ denotes the shock received by investor $r$ in state $\omega_n$. At date 2, investor $r$, whatever his type, learns about
the realization of \( \bar{l}^r \), and this is private information. Realizations of \( \bar{l}^r \) are i.i.d. across investors. We assume that \( \bar{l}^r \in \{(L, 0, 0), (0, L, 0)\} \) and

\[
\Pr(\bar{l}^r = (L, 0, 0)) = \tilde{\sigma},
\]

(3.3)

that is, all investors get a shock of size \( L \), either in state \( \omega_1 \) or \( \omega_2 \). The fraction of investors hit by a shock in state \( \omega_1 \) is denoted \( \tilde{\sigma} \) and it is assumed to be random \((1 - \tilde{\sigma} \) is then the fraction of those who are hit by a shock in state \( \omega_2 \)). Since the realization of \( \tilde{\sigma} \) is not observed, this randomness creates aggregate uncertainty in hedging motives that can make prices only partially informative. We assume that \( \tilde{\sigma} \in \{\sigma_L, \sigma_H\} \) with \( \sigma_H > \sigma_L, \sigma_H = 1 - \sigma_L, \) and

\[
\Pr(\tilde{\sigma} = \sigma_H) = \frac{1}{2}.
\]

(3.4)

**Trading decision**  We denote with \( x^r \) and \( y^r \) the holding of asset \( X \) and asset \( Y \) by investor \( r \) at the end of period 2. Each investor \( r \) chooses \( x^r \) and \( y^r \) by trading at date 2 so as to maximize his expected utility over final wealth, according to a common mean-variance utility

\[
EU(\bar{w}^r) = E(\bar{w}^r) - \frac{\rho}{2} V(\bar{w}^r),
\]

(3.5)

where \( \bar{w} \) denotes the investor \( r \)’s final wealth and \( \rho > 0 \) is a measure of risk aversion. For any passive investor \( j \), trades must satisfy

\[
\begin{align*}
x^j - e^j_X &= \beta^j e^j_X \\
y^j - e^j_Y &= \beta^j e^j_Y,
\end{align*}
\]

where the weight \( \beta^j \) is optimally chosen by the investor.

**Financial intermediaries**  There is a mass \( m \) of competitive financial institutions in the stock markets. Intermediaries do not receive any liquidity shock, they do not receive any signal on
the realization of the state of the world, and they can freely rebalance their portfolios. They are
dowered with \( e^f_X \) units of stock \( X \) and \( e^f_Y \) units of stock \( Y \). They have a common mean-variance utility
\[
U(\tilde{w}^f) = E(\tilde{w}^f) - \frac{\rho_f}{2} V(\tilde{w}^f),
\]
where \( \tilde{w}^f \) is the final wealth of financial intermediary \( f \) at date 3 and \( \rho_f > 0 \) is a measure of risk aversion.

**Order flows and stock prices** The stock market operates at \( t = 2 \). Denote by \( \tilde{z} = (\tilde{z}_X, \tilde{z}_Y) \) the aggregate order flows in markets \( X \) and \( Y \) and by \( z = (z_X, z_Y) \) a realization of \( \tilde{z} \). Denote with \( p = (p_X, p_Y) \) the prices of assets \( X \) and \( Y \). These prices are determined by the market clearing conditions
\[
z_X = z_Y = 0.
\]
Part of active investors’ demand comes from their private information, which may be revealed into prices. We allow investors to submit limit orders, and so to condition their demand on the information revealed into the total order flows. It follows that conditional on prices, passive investors’ and financial intermediaries’ beliefs become
\[
q_1^p = \Pr(\tilde{\omega} = \omega_1|p) \text{ and } q_2^p = \Pr(\tilde{\omega} = \omega_2|p).
\]

**Timing** The timing is depicted in Figure 3.1.

\[
\begin{align*}
 & t = 1 & t = 2 & t = 3 & t \\
\text{Endowment} & (e^\ell_C, e^\ell_X \text{ and } e^\ell_Y) & \text{Observe signal} (\tilde{s}) & \tilde{l}^r, \tilde{\omega}, \tilde{d} \text{ realize} \\
\text{Choice of investment profile} & (\alpha, \nu, \pi) & \text{Learn liquidity shocks} (\tilde{l}^r) & \text{Trade} (x^r, y^r)
\end{align*}
\]

Figure 3.1: Timeline of the Model

- At date 1, investor \( r \) receives his endowment in stocks and cash and chooses whether to become informed, uninformed, or passive.
At date 2, all investors learn the realization of their liquidity shock $\tilde{l}$; informed investors receive a signal $\tilde{s}$ on the realization of the state of nature $\tilde{\omega}$. Investors submit their limit orders and execute their trades in the two markets.

At date 3, all random variables are realized.

In our model, stocks pay either in state $\omega_1$ or $\omega_2$ and shocks occur either in state $\omega_1$ or $\omega_2$. Introducing state $\omega_3$ is needed (otherwise, some asset would be redundant) but its occurrence is not very informative for our purposes. In what follows, we let the probability that $\omega_3$ occurs be arbitrarily small; that is, we let $\eta \to 0$. This allows to simplify the expressions without losing our main insights. We will remark during the analysis when the assumption is important for the result.

We also assume that all investors have identical and symmetric endowments of stocks, that is

$$e_X^r = e_Y^r = e_1 \text{ for all } r.$$  \hfill (3.7)

### 3.3 Equilibrium demand and prices at date 2

We start by deriving the optimal demand for each stock $k$ at date 2, after the realization of the signal $\tilde{s}$ and the liquidity shocks $\tilde{l}$. Given our notations, the final wealth of an investor $r$ writes as

$$\tilde{w}^r = x^r \tilde{d}_X + y^r \tilde{d}_Y - p_X (x^r - e_X^r) - p_Y (y^r - e_Y^r) + e_C^r - \tilde{l}^r. \hfill (3.8)$$

An informed investor $i$ at date 2 maximizes the expected utility of his wealth with respect to $x^i$ and $y^i$. Plugging the expected value and variance of his wealth defined in (3.8) into his expected utility defined in (3.5) and taking the first order condition yields the following demand for asset $X$:

$$x^i = \frac{E(\tilde{d}_X | I^i)}{\rho V(\tilde{d}_X | I^i)} - \frac{y^i \text{cov}(\tilde{d}_X, \tilde{d}_Y | I^i)}{V(\tilde{d}_X | I^i)} + l_1^i \frac{\text{cov}(\tilde{d}_X, 1_{\omega_1} | I^i)}{V(\tilde{d}_X | I^i)} + l_2^i \frac{\text{cov}(\tilde{d}_X, 1_{\omega_2} | I^i)}{V(\tilde{d}_X | I^i)},$$

in which $l_1^i, l_2^i$ are the realizations of the endowment shock in states $\omega_1$ and $\omega_2$ for investor $i$, $I^i$ denotes investor $i$'s information at date 2, and $1_{\omega_n}$ is an indicator equal to 1 if state $\omega_n$ occurs. The
demand for asset $Y$ is similar. As for an uninformed investor $u$, demands can be derived in the same way, replacing the information set $I^i$ with $I^u$. This yields to the following Lemma.

**Lemma 3.1** The demand of an informed investor $i$ is such that:

$$x^i = \frac{(q_1^i - p_X)(1 - q_2^i) + (q_2^i - p_Y)q_1^i}{\rho q_1^i q_3^i} + l_1^i, \quad (3.9)$$

and

$$y^i = \frac{(q_2^i - p_Y)(1 - q_1^i) + (q_1^i - p_X)q_2^i}{\rho q_2^i q_3^i} + l_2^i. \quad (3.10)$$

The demand of an uninformed investor $u$ can be obtained as in (3.9) and (3.10) replacing $q_n^i$ with $q_n^u$ for $n = 1, 2, 3$.

Asset $X$ pays off 1 in state $\omega_1$ and 0 otherwise: it enables to hedge against $l_1^i$, the liquidity shock occurring in state $\omega_1$. Similarly, asset $Y$ enables to hedge against $l_2^i$. If the financial intermediaries were risk-neutral and had the same information set as the informed investor so that $p_X = q_1^i$ and $p_Y = q_2^i$, the investor’s final position would just correspond to his liquidity shock at date 3, that is $x^i = l_1^i$ and $y^i = l_2^i$. If instead the information set of active investors differs so that $p_X \neq q_1^i$ or $p_Y \neq q_2^i$, the informed investor incorporates a speculative component to his demand, as described by the first terms in (3.9) and (3.10).

A passive investor $j$ at date 2 rebalances his portfolio of risky assets $X$ and $Y$ in proportion $\beta^j$, so his holding of asset $X$ (resp. $Y$) equals to $(1 + \beta^j)e_X^j$ (resp. $(1 + \beta^j)e_Y^j$), and his cash is equal to $(e_C^j - \beta^j (p_X e_X^j + p_Y e_Y^j))$. His final wealth (3.8) can be written as

$$\tilde{w}^j(\beta^j) = (1 + \beta^j)e_X^j \tilde{d}_X + (1 + \beta^j)e_Y^j \tilde{d}_Y + e_C^j - \beta^j (p_X e_X^j + p_Y e_Y^j) - \tilde{l}. \quad (3.11)$$

Similarly to the case of active investors, we can define the demand at date 2 that maximizes the passive investors’ expected utility, under condition (3.7).

**Lemma 3.2** A passive investor rebalances his portfolio of assets $X$ and $Y$ in proportion $\beta^j$ such
that:

\[
(1 + \beta^j)e_1 = \frac{(q_1^p - p_X) + (q_2^p - p_Y)}{\rho q_2^p (q_1^p + q_2^p)} + \frac{q_1^p l_1^j + q_2^p l_2^j}{q_1^p + q_2^p}.
\]

(3.12)

Let us compare the demand in asset \( X \) between active and passive investors. First, an active investor can freely rebalance his portfolio; second, he may have a different belief about \( \tilde{\omega} \) since he receives a signal \( \tilde{s} \) on the state of nature realized at date 3. In order to highlight the role of trading constraints, suppose there is no speculative demand and financial intermediaries are risk-neutral \( (p_X = q_1^p = q_1^a \) and \( p_Y = q_2^p = q_2^a) \) and compare the hedging demand of an active investor \( i \) and of a passive investor \( j \) who have the same initial holdings \( e_1 \) and who have received the same liquidity shock \( \hat{l} \). From Lemmas 3.1 and 3.2, we have

\[
x^j > x^i \iff \hat{l}_2 > \hat{l}_1, \quad \text{and} \quad y^j > y^i \iff \hat{l}_2 < \hat{l}_1.
\]

It follows that a passive investor who is hit by the shock in \( \omega_1 \) for instance is demanding too much of stock \( Y \) and too little of stock \( X \), relative to an active investor who is hit by the same shock. Thus in aggregate, if a large fraction of investors receives a shock in state \( \omega_1 \), the demand for \( X \) will be too low and the demand for \( Y \) will be too high relative to a case in which investors have no trading constraints. The case for those hit by the shock in \( \omega_2 \) is symmetric.

In addition, the demands of active and passive investors in Lemma 3.1 and 3.2 exhibit an additional speculative or adverse selection component that comes from the difference, if any, between the asset’s expected value and market price. If financial intermediaries were risk-neutral and posted prices that would be equal to the expected value of each asset, that is, \( p_X = q_1^p \) and \( p_Y = q_2^p \), the demand of passive investors would only include the hedging component, and their beliefs may or may not be different from the beliefs of active investors, depending on whether or not prices are fully revealing.

Financial intermediaries are uninformed, unconstrained and have no liquidity shock. Their demand can be derived from the demand of active investors, as in Lemma 3.1, substituting \( q_n^a \) with

131
\( q^p \) and setting to \( t_1^f = t_2^f = 0 \). That gives:

\[
x^f = \frac{(q_1^p - p_X)(1 - q_2^p) + (q_2^p - p_Y)q_1^p}{\rho_f q_1^p q_2^p q_3^p},
\]

(3.13)

and

\[
y^f = \frac{(q_2^p - p_Y)(1 - q_1^p) + (q_1^p - p_X)q_2^p}{\rho_f q_1^p q_2^p q_3^p}.
\]

(3.14)

Aggregating the demand for \( X \) and \( Y \) of the various investors and intermediaries one obtains the order flows \( \tilde{z} \). The order flows depend on the realization of the signal \( \tilde{s} \) and on the aggregate liquidity shock \( \tilde{\sigma} \). This implies that the realization of the signal need not be perfectly inferred by observing the order flow or conditioning on prices. Setting \( z_X = z_Y = 0 \) gives the market clearing prices for \( X \) and \( Y \), as defined in the following proposition.

**Proposition 3.1** When \( \eta \to 0 \), equilibrium prices are

\[
\tilde{p}_X = q_1^a - q_1^a q_2^a \tilde{\Lambda},
\]

(3.15)

\[
\tilde{p}_Y = q_2^a + q_1^a q_2^a \tilde{\Lambda},
\]

where

\[
\tilde{\Lambda} \equiv \frac{(m + \frac{\nu}{\rho}) (q_1^a - q_1^p) + (\alpha + \nu)(1 - 2\tilde{\sigma})q_1^p q_2^p L}{\frac{\alpha}{\rho} q_1^p q_2^p + \left( m + \frac{\nu}{\rho} \right) q_1^a q_2^a}.
\]

(3.16)

As shown in (3.15), prices can be expressed simply as the expected dividend according to informed investors and of the term \( \tilde{\Lambda} \) described in (3.16), which depends on investors’ beliefs (and so on the realization of the signal) as well as on the realization of \( \tilde{\sigma} \). It describes a premium that intermediaries require to bear the possibility of adverse selection, driven by the difference in beliefs \( q_1^a - q_1^p \), as well as to provide hedging against the aggregate liquidity shock, driven by the term in \( L \).

We now analyze the conditions for the existence of an equilibrium in which prices (or equivalently, aggregate order flows) do not allow to perfectly infer which signal active investors have
received. This is the case when hedging demand and speculative demand offset each other. For example, the demand for asset $X$ may be large due to high hedging needs in state $\omega_1$ (that is, when $\tilde{\sigma} = \sigma H$) or when investors have received a signal $\tilde{s} = s_1$. When aggregate hedging needs and speculation go in the same direction, orders are fully revealing. For example, when $\tilde{\sigma} = \sigma H$ and $\tilde{s} = s_1$ the demand for $X$ is so large and the demand for $Y$ is so low that one can infer that signal $\tilde{s} = s_1$ must have been observed. A similar argument applies when $\tilde{\sigma} = \sigma L$ and $\tilde{s} = s_2$ as well as when $\tilde{s} = s_3$. We thus examine whether it is possible to have order flows such that:

$$p_X(s_1, \sigma L) = p_X(s_2, \sigma H) \text{ and } p_Y(s_1, \sigma L) = p_Y(s_2, \sigma H),$$

which we consider in the following lemma.

**Lemma 3.3** Prices are not fully revealing if and only if

$$\frac{(3\tau - 1)(1 + \tau)}{4\rho \tau (1 - \tau)} \alpha = (\alpha + \upsilon)(\sigma H - \sigma L)L. \tag{3.17}$$

We label condition (3.17) “NFR condition”. To understand this condition, consider the demand for asset $X$: the increased speculative component by informed investors when the signal is $s_1$ as opposed to $s_2$ (on the left hand side of (3.17)) should equal the increased hedging component from informed and uninformed investors when $\tilde{\sigma} = \sigma H$ as opposed to $\sigma L$ (on the right hand side of (3.17)).

If condition (3.17) holds and the realizations of $\tilde{s}$ and $\tilde{\sigma}$ are either $(s_1, \sigma L)$ or $(s_2, \sigma H)$, then uninformed investors cannot distinguish if $\tilde{s} = s_1$ or $\tilde{s} = s_2$, which implies $q^a_1 = q^u_2$ so their demand is the same irrespective of the realization of $\tilde{s}$. By contrast for an informed investor, $q^a_1 = \Pr(\tilde{\omega} = \omega_n | \tilde{s} = s)$. Substituting these relations in equation (3.15) gives the corresponding equilibrium prices. The first term in $\tilde{\Lambda}$ decreases in $\alpha$ if and only if $q^a_1 > q^u_2$ due to increased speculative demand for asset $X$. When $\upsilon = 0$, the second term in $\tilde{\Lambda}$ decreases in $\alpha$ if and only if $\tilde{\sigma} = \sigma H$ due to higher hedging demand for asset $X$. This effect is dampened when $\upsilon > 0$ since uninformed investors, who are affected by the same liquidity shock, restrain their demand for fear of adverse
selection. We get symmetric variations for \( p_Y \) relative to \( q_2^a \).

When condition (3.17) does not hold, or when it does but the realizations of \( \tilde{s} \) and \( \tilde{\sigma} \) are neither \((s_1, \sigma_L)\) nor \((s_2, \sigma_H)\), then equilibrium prices are fully revealing (FR): conditioning on prices enables uninformed agents to condition their demand on the realization of the signal observed by active investors. It follows that \( q_n^b = q_n^a \) for all \( n \). Quite intuitively in this case, \( p_X \) increases in \( \alpha \) if and only if \( 1 - 2\tilde{\sigma} < 0 \) due to higher hedging demand for asset \( X \) while \( p_Y \) increases in \( \alpha \) if and only if \( 1 - 2\tilde{\sigma} > 0 \) for a symmetric reason.

In the following part of the paper, we focus on equilibrium when condition (3.17) holds. We call it the “non fully revealing (NFR) equilibrium”.

### 3.4 Equilibrium choice at date 1

In this section, we characterize the equilibrium at \( t = 1 \), where each investors needs to decide whether to become informed, uninformed or passive. We first derive the expected utility of each type of investors, as a function of our exogenous parameters as well as of the endogenous shares of informed, uninformed and passive investors. Comparing these expected utilities, we can define the value of being active relative to being passive as well as the value of information in our setting. Comparing these values to the costs \( c_I \) and \( c_T \) allows to derive the equilibrium shares of informed, uninformed and passive investors.

#### 3.4.1 Expected utility of informed, uninformed and passive investors

We define investors’ expected utility at \( t=1 \) in terms of expected mean variance utility, that is

\[
W^r \equiv E_1[E_2(\bar{w}^r) - \frac{\rho}{2}V_2(\bar{w}^r)],
\]

(3.18)

where \( E_1 \) denotes period-1 expectation and \( E_2, V_2 \) denote respectively period-2 expectation and variance of final wealth. Expected mean variance utility is commonly used in information choice
problems, we refer to Van Nieuwerburgh and Veldkamp (2010b) and Kacperczyk et al. (2016) for a discussion.

To compute the investors’ expected utility, we first plug the demands for $X$ and $Y$ given in Lemmas 3.1 and 3.2 and the equilibrium prices given by (3.15) in Proposition 3.1 into the investors’ wealth defined in equation (3.8) for all potential realizations of $\tilde{s}$, $\tilde{\sigma}$ and $\tilde{l}$, and the corresponding expected utility of wealth at date 2. As expressed in (3.18), the expected utility at date 1 is defined as the sum of expected utilities at date 2, weighted by the probabilities to observe the corresponding realizations of $\tilde{s}$, $\tilde{\sigma}$ and $\tilde{l}$.

Define

$$\Lambda_F \equiv \tilde{\Lambda}(s_2, \sigma_L),$$

which measures the price impact when orders are fully revealing. We have $q_n^a = q_n^p$, which substituted into (3.16) gives

$$\Lambda_F = \frac{\rho_f(\alpha + v)}{m\rho + (\alpha + v)\rho_f} (\sigma_H - \sigma_L) \rho L. \quad (3.19)$$

When financial intermediaries are risk averse, they require a premium to accommodate the excess demand or supply of assets $X$ and $Y$, captured by $\Lambda_F > 0$. Notice that $\Lambda_F$ depends on the financial intermediaries’ capacity or willingness to absorb the residual demand: the size of this premium increases with $(\sigma_H - \sigma_L)$, with $L$ and $\rho_f$, and it decreases with $m$. Notice that $\tilde{\Lambda}(s_1, \sigma_H) = -\Lambda_F$.

Similarly, define

$$\Lambda_N \equiv \tilde{\Lambda}(s_1, \sigma_L),$$

which measures the price impact when orders are not fully revealing. We have $q_1^a = \frac{2\tau}{1+\tau}$ and $q_1^p = 1/2$, which gives (after some manipulations)

$$\Lambda_N = \left( \frac{m}{\rho_f} + \frac{v}{\rho} \right) \frac{2(3\tau - 1)}{(1+\tau)} + \frac{\alpha + v(\sigma_H - \sigma_L)\rho L}{\alpha} + \left( \frac{m}{\rho_f} + \frac{v}{\rho} \right) \frac{8\tau(1-\tau)}{(1+\tau)^2}.$$

One can show that $\Lambda_N > \Lambda_F$ and $\Lambda_N \to \Lambda_F$ as $\tau \to 1/3$. In addition to a risk premium as measured by $\Lambda_F$, $\Lambda_N$ includes an adverse selection cost.
Under condition (3.17), we have

\[ \Lambda_{N}^{NFR_{cond}} = \frac{(3\tau - 1)(1 + \tau)}{4\tau(1 - \tau)}. \]

Notice that \( \tilde{\Lambda}(s_2, \sigma_H) = -\Lambda_N \). We characterize in the following proposition the expected utility of the various agents in a not-fully revealing equilibrium, that is, when the NFR condition (3.17) given in Lemma 3.3 holds and so informed investors may benefit from their private information. The proposition below characterizes the investors’ expected utility in this case.

**Proposition 3.2** Assume (3.17) holds and \( \eta \to 0 \). In the non-fully revealing equilibrium,

1. The active investor’s expected utility is

\[
W^i(NFR) = e^i_C + e_1 - \frac{L}{2} - \frac{\tau(1 - \tau)}{2(1 + \tau)} \left( (\Lambda_F + \Lambda_N) + \frac{1 + \tau}{4\tau} \Lambda_F \right) (\sigma_H - \sigma_L) L \\
+ \frac{\tau(1 - \tau)}{4(1 + \tau)} \left( \frac{\Lambda_N^2}{\rho} + \frac{1 + 5\tau \Lambda_F^2}{4\tau} \right) \tag{3.20}
\]

2. The unconstrained uninformed investor’s expected utility is

\[
W^u(NFR) = e^u_C + e_1 - \frac{L}{2} - \frac{\tau(1 - \tau)}{2(1 + \tau)} \left( (\Lambda_F + \Lambda_N) + \frac{1 + \tau}{4\tau} \Lambda_F \right) (\sigma_H - \sigma_L) L \\
+ \frac{\tau(1 - \tau)}{4(1 + \tau)} \left( \frac{8\tau(1 - \tau) \Lambda_N^2}{(1 + \tau)^2} + \frac{(1 + 5\tau) \Lambda_F^2}{4\tau} \right) \\
+ \frac{(3\tau - 1)^2}{8\rho(1 + \tau)} + \frac{\tau(1 - \tau)}{(1 + \tau)^2} \frac{1 - 3\tau}{\rho} \Lambda_N \tag{3.21}
\]

3. The passive investor’s expected utility is

\[
W^j(NFR) = e^j_C + e_1 - \frac{L}{2} - \frac{3\tau - 1}{8} (\sigma_H - \sigma_L) L - \rho \frac{3 + 10\tau - 9\tau^2}{32(1 + \tau)} L^2 \tag{3.22}
\]
4. Financial intermediaries’ expected utility is

\[
W^f(NFR) = e_C^f + e_1^f + \frac{(3\tau - 1)^2}{8\rho_f(1 + \tau)} + \frac{(1 - \tau)(1 + 5\tau)}{16(1 + \tau)} \frac{\Lambda_N^2}{\rho_f} \\
+ \frac{\tau(1 - \tau)}{(1 + \tau)^2} \left( (1 - 3\tau) + \frac{2\tau}{1 + \tau}(1 - \tau)\Lambda_N \right) \frac{\Lambda_N}{\rho_f} 
\]

(3.23)

Informed investors

To interpret Proposition 3.2, it is useful to note that under the NFR condition (3.17), his expected mean variance utility at date 1 writes

\[
W^{i,NFR\text{cond}}(NFR) = c_1 + \frac{1}{2}(2e_1 - L) - \frac{(1 - \tau)(5\tau + 1)}{8(1 + \tau)} \left( m \rho + v \rho_f + \alpha \rho_f \right) \left( \sigma_H - \sigma_L \right) L \Lambda_F + \frac{\Lambda_N^2}{2\rho} \\
+ \frac{\nu - \alpha}{\nu + \alpha} \frac{(3\tau - 1)^2}{64\tau^2(1 - \tau)}. 
\]

(3.24)

For the realizations of \( \tilde{s} \) and \( \tilde{\sigma} \) that are such that prices are not fully revealing, the active investor adjusts his trading volume so that his private information offsets the risk premium incorporated into prices: his expected wealth in these cases is insensitive to prices. There is however an additional component in his expected utility, namely the term in \( \frac{(3\tau - 1)^2(1 + \tau)}{64\rho^2(1 - \tau)} \), that is due to the fact that financial intermediaries incorporate an adverse selection cost in their prices.

The term \( \frac{mp}{m\rho + (\alpha + \nu)\rho_f} \left( \sigma_H - \sigma_L \right) L \Lambda_F \) that reduces the active investor’s expected utility comes for the active investor’s expected wealth. It accounts for the fact that the active investor is more likely to receive a liquidity shock inducing him to hedge using the most expensive asset. For instance, he is more likely to hedge against a shock in state \( \omega_1 \) when \( \tilde{\sigma} = \sigma_H \), situation in which buying asset \( X \) is more expensive than buying asset \( Y \). When \( \sigma_H - \sigma_L \) is large, it is indeed costly to hedge as many other investors have the same hedging demand, even though the active investor adjust his demand accordingly. More precisely, the whole negative term is the sum of two components, namely \( -\frac{(1 - \tau)(5\tau + 1)}{8(1 + \tau)} \left( \sigma_H - \sigma_L \right) L \Lambda_F \) that comes from prices, and \( + \frac{(1 - \tau)(5\tau + 1)}{8(1 + \tau)} \frac{\Lambda_N^2}{2\rho} \) that comes from the expected trading volume. The latter volume adjustment is not sufficient to
compensate the increased price impact. The active investor’s expected wealth thus increases when price impact \( \Lambda_F \) decreases, or when \((\sigma_H - \sigma_L)\) decreases.

Expected utilities also depend on the variance of investors’ wealth, that also depends on the risk premium \( \Lambda_F \) incorporated in the price they need to pay to hedge. At the limit when financial intermediaries are risk neutral, prices are inelastic and only depend on beliefs (i.e., \( \Lambda_F \to 0 \)). Since the active informed investor can fully hedge against his liquidity shock at a price that fully reveals his information, the variance of his final wealth is zero. When financial intermediaries are not risk neutral however, the expected utility of active investors decreases with \( \Lambda_F \) since there is more variation in the active investor’s traded quantity.

Let us denote by \( \Lambda_N \equiv \tilde{\Lambda}(s_1, \sigma_L) \) the price impact observed in a case where prices are not fully revealing. For the realizations of \( \tilde{s} \) and \( \tilde{\sigma} \) that are such that prices are not fully revealing, the active investor adjusts his trading volume so that his private information offsets the risk premium incorporated into prices: his expected wealth in these cases is insensitive to prices. There is however an additional component in his expected utility, namely the term in \( (3\tau - 1)^2 (1 + \tau) \) \( 64 \rho \tau (1 - \tau) \), that is due to the fact that financial intermediaries incorporate an adverse selection cost in their prices. It is indeed possible to show that the price impact observed in a case where prices are not fully revealing, i.e., \( \Lambda_N \), is equal to \( \frac{(3\tau - 1)(1 + \tau)}{4 \tau (1 - \tau)} \) when condition (3.17) holds.

**Passive investors**

Notice that given their optimal demand, the final wealth of passive investors writes

\[
\omega^j(\beta^j) = \left( \frac{(q_1^p - p_X) + (q_2^p - p_Y)}{\rho q_3^p (q_1^p + q_2^p)} + \frac{q_1^{p_1} l_1^{p_1} + q_2^{p_1} l_2^{p_1}}{q_1^p + q_2^p} \right) (1_{\omega_1} - p_X + 1_{\omega_2} - p_Y) \\
+ e_C^j + (e_X^j - l_1^j) 1_{\omega_1} + (e_Y^j - l_2^j) 1_{\omega_2} - e_X^j (1_{\omega_1} - p_X) - e_Y^j (1_{\omega_2} - p_Y).
\]  

Notice that, due to a no arbitrage condition, we have

\[
\lim_{n \to 0} (p_X + p_Y) = 1,
\]  

138
since buying one unit of each asset would give a payoff of 1 for sure when \( \eta \to 0 \).

It can also be shown that \( \lim_{\eta \to 0} (1 + \beta^j) e_1 \) is finite. Hence, at the limit when \( \eta \to 0 \) so that \( p_X + p_Y \to 1 \), the demand of the passive investor does not depend on prices. His final wealth may however still depend on prices, as it writes

\[
\tilde{w}^j (\beta^j) = (1 + \beta^j) e_X^{j} \tilde{d}_X + (1 + \beta^j) e_Y^{j} \tilde{d}_Y + e_C^j - \beta^j (p_X e_X^{j} + p_Y e_Y^{j}) - \tilde{\nu}. \tag{3.27}
\]

Whether the passive investor’s final wealth depends on prices or not thus relies on \( (e_X^{j} - e_Y^{j}) \).

Under assumption (3.7), it does not, and we have

\[
\text{If } e_X^{j} = e_Y^{j} = e_1, \text{ } \tilde{w}^j = -\tilde{\nu} \frac{1}{2} \omega_1 - \tilde{\nu} \frac{1}{2} \omega_2 + e_1 + c_1. \tag{3.28}
\]

Passive investors are constrained to buying assets \( X \) and \( Y \) in equal proportion and since \( p_X + p_Y \to 1 \) they make no losses in expectations: their losses due to an excessive demand of asset \( X \), say, are exactly compensated by what they gain by a lower demand of asset \( Y \). Hence, their expected wealth is immune to price impact \( \Lambda_F \). For the same reason, the variance of the passive investors’ wealth does not depend on \( \Lambda_F \). The passive investor’s expected utility however includes a term in \( L^2 \). The term \( L^2/4 \) in the variance of the passive investor comes from the fact that the realized liquidity shock in a given state is either 0 or \( L \) and the expected liquidity shock is \( L/2 \), hence due to his trading constraint, a passive investor has to deviate from \( L/2 \) in each of the two assets. Consequently, the trading constraint enables the passive investor to be immune to movements in prices, but he still bears a risk due to the imperfect hedge of his liquidity shock.

**The value of Information**

Let us define the value of information as follows:

\[
\Pi^I \equiv W^i (NFR) - W^u (NFR).
\]
It follows from Proposition 3.2 that

$$\Pi_I' = \frac{(3\tau - 1)^2\tau(1 - \tau)}{4\rho(1 + \tau)^3} \left( -\frac{(1 + \tau)^2}{2\tau(1 - \tau)} + \frac{4(1 + \tau)}{3\tau - 1}\Lambda_N + \Lambda_N^2 \right)$$  \text{(3.29)}

There are three components in $\Pi_I'$. The first part comes from $\frac{(3\tau-1)^2}{8\rho(1+\tau)}$ that appears (with a positive sign) in the uninformed investors’ expected utility but not in that of informed investors. It increases in $\tau$ but decreases with $\rho$. The second part comes from an adverse selection cost that is paid by uninformed investors, namely $-\frac{\tau(1-\tau)}{(1+\tau)^2}(3\tau - 1)\frac{\Lambda_N}{\rho}$. It increases then decreases in $\tau$, it increases with the price impact $\Lambda_N$. The last part comes from the coefficient $\frac{8\tau(1-\tau)}{(1+\tau)^2} < 1$ in front of $\Lambda_N^2$: without additional information, investors adjust their trade size.

**Corollary 3.1** Assume (3.17) holds and $\eta \to 0$. The value of information $\Pi_I'$ is positive, and it decreases in $\alpha$, and it increases in $\nu$.

Investments in information are strategic substitutes: the decision of one active investor to invest in information decreases the marginal value of information for the other active investors.

**The value of being active**

We now define the value of flexibility by comparing the passive and the active uninformed investor, that is, $\Pi_T^{I} \equiv W^u(NFR) - W^i(NFR)$.

It follows from Proposition 3.2 that

$$\Pi_T^{I} = \frac{(1 - \tau)(1 + 5\tau)}{8(1 + \tau)}\Lambda_F(\sigma_H - \sigma_L)L + \frac{(1 - \tau)(1 + 5\tau)}{8(1 + \tau)}\Lambda_N^2 \rho$$

$$- \frac{4\tau(1 - \tau)}{8(1 + \tau)}\Lambda_N(\sigma_H - \sigma_L)L + \frac{4\tau(1 - \tau)}{(1 + \tau)}\left( \frac{\tau(1 - \tau)}{(1 + \tau)^2}\Lambda_N - \frac{3\tau - 1}{2(1 + \tau)} \right)\Lambda_N^2 \rho$$

$$+ \frac{(3\tau - 1)^2}{8\rho(1 + \tau)} + \frac{3\tau - 1}{8}(\sigma_H - \sigma_L)L + \rho\frac{3 + 10\tau - 9\tau^2}{32(1 + \tau)}L^2$$  \text{(3.30)}

This yields the following corollary.

**Corollary 3.2** Assume (3.17) holds and $\eta \to 0$. The value of trading flexibility $\Pi_T^{I}$ is positive, it
increases in \( L, m \), and it decreases in \((\sigma_H - \sigma_L)\). \( \Pi^T \to 0 \) as \( L \to 0 \), \( \rho \to 0 \). \( \Pi^T \) is increasing in \( \alpha \).

### 3.4.2 Equilibrium shares of informed, uniformed and passive investors

We now characterize the conditions under which NFR equilibria exist.

**Proposition 3.3** Assume that the condition \((\sigma_H - \sigma_L)L = \frac{(3\tau-1)(1+\tau)}{4\rho(1-\tau)}\) holds. Let us define

\[
c_i^* = \frac{(1 + \tau)(3\tau - 1)^2}{64\rho(1 - \tau)} \quad (3.31)
\]

\[
c_T^* = \frac{3 + 10\tau - 9\tau^2}{32(1 + \tau)}\rho L^2 \quad (3.32)
\]

\[
c_{i|T}^* = \rho \left(\frac{1 - \tau}{1 + 5\tau}\right) \left(1 - \frac{(m\rho)^2}{(m\rho + \rho f)^2}\right) \left(\sigma_H - \sigma_L\right)^2 L^2 \quad (3.33)
\]

There exists an NFR equilibrium in the following cases.

1. If \( c_I < c_i^* \), and if \( c_I + c_T < c_i^* + c_T^* \), the there exists a NFR equilibrium with informed active investors, and passive investors. At equilibrium,

\[
\alpha^* = \min \left( m\rho \rho_f \left(1 - \sqrt{\frac{(c_I + c_T - c_i^* - c_T^*)}{(\sigma_H - \sigma_L)^2 L^2 \rho(1-\tau)(1+5\tau)^2} + 1}\right), 1 \right)
\]

\[
\pi^* = 1 - \alpha^* \text{, and } \nu^* = 0.
\]

The max in \( \alpha^* \) is equal to 1 when \( c_I + c_T \leq c_i^* + c_T^* - c_{i|T}^* \).

2. If \( c_I = c_i^* \) and if \( c_T < c_T^* \), then exists a NFR equilibrium in mixed strategies, in which active investors are indifferent between becoming informed or not.

   a) If \( c_T < c_T^* - c_{i|T}^* \), then there are no passive traders at equilibrium and

\[
\alpha^* = \frac{4\rho(1 - \tau)}{(3\tau - 1)(1 + \tau)}(\sigma_H - \sigma_L)L, \pi^* = 0 \text{ and } \nu^* = 1 - \alpha^*,
\]

141
\[ b) \text{If } c_T^* - c_T^{1|T} \leq c_T < c_T^*, \]

\[
\alpha^* = \frac{4\rho\tau(1 - \tau)}{(3\tau - 1)(1 + \tau)} \frac{m\rho}{\rho_f} \left( \frac{(\sigma_H - \sigma_L) L \sqrt{\rho \frac{(1 - \tau)(1 + 5\tau)}{16(1 + \tau)}}}{\sqrt{c_T - c_T^*} + \rho \frac{(1 - \tau)(1 + 5\tau)}{16(1 + \tau)}} \right) \left( \frac{(\sigma_H - \sigma_L)^2 L^2}{1} \right) - 1 \right) (\sigma_H - \sigma_L) L
\]

\[
v^* = \frac{m\rho}{\rho_f} \left( \frac{(\sigma_H - \sigma_L) L \sqrt{\rho \frac{(1 - \tau)(1 + 5\tau)}{16(1 + \tau)}}}{\sqrt{c_T - c_T^*} + \rho \frac{(1 - \tau)(1 + 5\tau)}{16(1 + \tau)}} \right) \left( 1 - \frac{4\rho\tau(1 - \tau)}{(3\tau - 1)(1 + \tau)} \right) (\sigma_H - \sigma_L) L
\]

\[
\pi^* = 1 - \alpha^* - v^*
\]

Figure 3.2 shows the equilibrium.

The equilibrium conditions are intuitive. When the cost of acquiring information is sufficiently low, there will not be unconstrained but uninformed investors. Conditional on being unconstrained, investors are better off paying an additional cost to become informed. The equilibrium proportion of informed and passive investors is determined by an indifference condition: investors will become informed until the the value of being informed shrinks to 0. When the cost of information rises, we could have a equilibrium in which all three types of investors co-exits. Finally, if the cost being unconstrained is sufficiently low, we will have an equilibrium with no passive investors. In this case, being unconstrained strictly dominates being passive.

The welfare functions characterized in Proposition 3.2 also allows us to study optimality of the equilibrium. For example, we can define social welfare \( W \) as

\[
W = \alpha (W^i - c_I - c_T) + v (W^u - c_T) + \pi W^j + mW^f. \tag{3.34}
\]

We can then investigate whether the equilibrium choices of \( \alpha, v \) and \( \pi \) are consistent with social welfare maximization.

Denote \( C = \alpha e_C^i + ve_C^u + (1 - \alpha - v)e_C^j + mc^j \) and \( E = \alpha e_1 + ve_1 + (1 - \alpha - v)e_1 + me^j_1 \). Using our notations, we can then rewrite social welfare as follows

\[
W = C_1 + \frac{1}{2} (2E - L) + \alpha (\Pi^i - c_I) + (\alpha + v)(\Pi^T - c_T) + W^j + mW^f.
\]
To determine the socially optimal proportion of informed investors in a NFR equilibrium, we are interested in

\[
\frac{\partial W}{\partial \alpha} = \Pi^I + \Pi^T - c_I - c_T + \alpha \frac{\partial \Pi^I}{\partial \alpha} + (\alpha + \upsilon) \frac{\partial \Pi^T}{\partial \alpha} + m \frac{\partial W^f}{\partial \alpha}
\] (3.35)

After we plug in the NFR condition, computation yields:

\[
\frac{\partial W(NFR)}{\partial \alpha} = \Pi^I + \Pi^T - c_I - c_T + \frac{(1 - \tau)(1 + 5\tau)}{4(1 + \tau)}(\alpha + \upsilon) \rho^2 \rho f m \left(\frac{m \rho + (\alpha + \upsilon) \rho f}{2(\sigma_H - \sigma_L)^2 L^2}\right).
\] (3.36)

Equation 3.35 highlights the difference between equilibrium and optimum. While individual investors care about the gains at the level \((\Pi^I + \Pi^T - c_I - c_T)\), a social planner also cares about the marginal contribution to the value of information and the value of flexibility. Hence equilibrium may not coincide with optimum.

Consider the case in which there exists an equilibrium in mixed strategies, with informed, uninformed and passive investors. Necessary conditions to sustain such an equilibrium is \(\Pi^I - c_I + \Pi^T - c_T = 0\). Thus

\[
\frac{\partial W(NFR)}{\partial \alpha} = \frac{(1 - \tau)(1 + 5\tau)}{4(1 + \tau)}(\alpha + \upsilon) \rho^2 \rho f m \left(\frac{m \rho + (\alpha + \upsilon) \rho f}{2(\sigma_H - \sigma_L)^2 L^2}\right).
\] (3.37)

which is strictly positive. Thus in this case, \(\alpha^{SP} > \alpha^*:\) there is under-investment in information at equilibrium. In fact, it is easy to see that in all equilibria \(\Pi^I + \Pi^T - c_I - c_T \geq 0\), apart from the corner case \(\alpha^* = 1\), our equilibria feature under-investment in information.

In general, we can derive optimal proportion of \(\alpha, \upsilon\) and \(\pi\) under different social welfare functions, and then contrast it with equilibrium. A full-fledged analysis falls out of the scope of this chapter and is further explored in Bianchi, Guo and Moinas [2020].
3.5 Conclusion

We analyze theoretically the equilibrium choice of active and passive investing. Active investors incur higher trading costs but can freely rebalance each asset in their portfolio. They can also obtain costly information on the state of the world. Passive investors incur lower costs but can only rebalance the proportion of wealth invested in risky assets. In our model, passive investors do not acquire information, but can fully hedge against the price impact of asymmetric information. In equilibrium, investors trade off the value of being active against the costs of being active. We show that when the cost of acquiring information is sufficiently low, equilibrium features the coexistence of informed and passive investors. The equilibrium proportion of informed and passive investors is determined by an indifference condition: investors will become informed until the value of being informed shrinks to zero. When the cost of acquiring information rises, we could have an equilibrium in which informed, uninformed but unconstrained and passive investors co-exist. Finally, if the cost being unconstrained is sufficiently low, we will have an equilibrium with no passive investors. In this case, being unconstrained strictly dominates being passive. While individual investors care about the value of being active at the level, a social planner also cares about the marginal contribution of being active. Hence, market equilibria may induce sub-optimal levels of active investing.
3.6 Figures

Figure 3.2: Equilibrium Proportion of Different Investors

This figure shows the equilibrium proportion of different investors. The horizontal line denotes the cost of acquiring information $c_T$. The vertical line denotes the cost of being active $c_I$. $\alpha^*$ is the equilibrium proportion of informed investors. $\nu^*$ is the equilibrium proportion of unformed but unconstrained investors. $\pi^*$ is the equilibrium proportion of passive investors.

3.7 Proof

Preliminaries: Conditional moments and probabilities Let us compute the expectation, variance and covariance under the investor’s specific information set at date 2, $I^r$. Given our assump-
tions on \( \Omega \) and \( \tilde{d}_k \) and our notation \( q^r_i = \Pr(\tilde{\omega} = \omega_i | I^r) \), we have:

\[
E(\tilde{d}_X | I^r) = q^r_1 \\
E(\tilde{d}_Y | I^r) = q^r_2 \\
V(\tilde{d}_X | I^r) = \text{Cov}(\tilde{d}_X, 1_{\omega_1} | I^r) = q^r_1(1 - q^r_1) \\
V(\tilde{d}_Y | I^r) = \text{Cov}(\tilde{d}_Y, 1_{\omega_2} | I^r) = q^r_2(1 - q^r_2) \\
\text{Cov}(\tilde{d}_X, \tilde{d}_Y | I^r) = \text{Cov}(\tilde{d}_Y, 1_{\omega_1} | I^r) = \text{Cov}(\tilde{d}_X, 1_{\omega_2} | I^r) = -q^r_1 q^r_2.
\]

Let us introduce the following notation and compute the posterior probabilities upon observing the various signals (using Bayes’ rule):

\[
\theta \equiv \Pr(\tilde{\omega} = \omega_1 | \tilde{s} = s_1) = \Pr(\tilde{\omega} = \omega_2 | \tilde{s} = s_2) = 2\tau \frac{1 - \eta}{\tau + \eta - 3\tau \eta + 1} \tag{3.39} \\
\varepsilon \equiv \Pr(\tilde{\omega} = \omega_3 | \tilde{s} = s_1) = \Pr(\tilde{\omega} = \omega_3 | \tilde{s} = s_2) = (1 - \tau) \frac{2\eta}{\tau + \eta - 3\tau \eta + 1} \\
\chi \equiv \Pr(\tilde{\omega} = \omega_1 | \tilde{s} = s_3) = \Pr(\tilde{\omega} = \omega_2 | \tilde{s} = s_3) = \frac{(1 - \eta)(1 - \tau)}{2 - 2\tau - 2\eta + 6\tau \eta} \\
\kappa \equiv \Pr(\tilde{\omega} = \omega_1 | \tilde{s} = s_1 \cup \tilde{s} = s_2) = \Pr(\tilde{\omega} = \omega_2 | \tilde{s} = s_1 \cup \tilde{s} = s_2) = \frac{(1 - \eta)(\tau + 1)}{2\tau + 2\eta - 6\tau \eta + 2}.
\]

Notice that

\[
\Pr(\tilde{\omega} = \omega_2 | \tilde{s} = s_1) = \Pr(\tilde{\omega} = \omega_1 | \tilde{s} = s_2) = 1 - \theta - \varepsilon.
\]

Besides, \( \lim_{\eta \to 0} \theta = \frac{2\tau}{1 + \tau} \), \( \lim_{\eta \to 0} \varepsilon = 0 \), \( \lim_{\eta \to 0} \chi = \frac{1}{2} \), and \( \lim_{\eta \to 0} \kappa = \frac{1}{2} \).

**Lemma 3.1 (individual demand of the unconstrained investor)**

Consider an unconstrained investor \( r \) with a mean-variance utility, and an information set \( I^r \) consisting of his liquidity shock, and beliefs \((q^r_1, q^r_2, q^r_3)\). We omit the reference to this information set in the expressions of the moments below for brevity. Since he is not constrained in his trading
quantities, the expected value and variance of his wealth respectively write:

\[
E(\tilde{w}(x^r, y^r)) = x^r E(\tilde{d}_X) + y^r E(\tilde{d}_Y) - p_X(x^r - x_1^r) - p_Y(y^r - y_1^r) + \epsilon_C - E(\tilde{l}^r), \tag{3.40}
\]

and

\[
V(\tilde{w}(x^r, y^r)) = (x^r)^2 V(\tilde{d}_X) + (y^r)^2 V(\tilde{d}_Y) + (l_1^r)^2 V(1_{\omega_1}) + (l_2^r)^2 V(1_{\omega_2}) \tag{3.41}
+ 2x^r y^r \text{Cov}(\tilde{d}_X, \tilde{d}_Y) - 2x^r l_1^r \text{Cov}(\tilde{d}_X, 1_{\omega_1}) - 2x^r l_2^r \text{Cov}(\tilde{d}_X, 1_{\omega_2})
- 2y^r l_1^r \text{Cov}(\tilde{d}_Y, 1_{\omega_1}) - 2y^r l_2^r \text{Cov}(\tilde{d}_Y, 1_{\omega_2}) + 2l_1^r l_2^r \text{Cov}(1_{\omega_1}, 1_{\omega_2}).
\]

Plugging (3.40) and (3.41) into (3.5) and taking the FOC yields the following demand for each asset:

\[
(x^r - x_1^r) = -x_1^r + \frac{E(\tilde{d}_X) - p_X}{\rho V(\tilde{d}_X)} - x^r \frac{\text{Cov}(\tilde{d}_X, \tilde{d}_Y)}{V(\tilde{d}_X)} + l_1^r \frac{\text{Cov}(\tilde{d}_X, 1_{\omega_1})}{V(\tilde{d}_X)} + l_2^r \frac{\text{Cov}(\tilde{d}_X, 1_{\omega_2})}{V(\tilde{d}_X)},
\]

and

\[
(y^r - y_1^r) = -y_1^r + \frac{E(\tilde{d}_Y) - p_Y}{\rho V(\tilde{d}_Y)} - y^r \frac{\text{Cov}(\tilde{d}_X, \tilde{d}_Y)}{V(\tilde{d}_Y)} + l_1^r \frac{\text{Cov}(\tilde{d}_Y, 1_{\omega_1})}{V(\tilde{d}_Y)} + l_2^r \frac{\text{Cov}(\tilde{d}_Y, 1_{\omega_2})}{V(\tilde{d}_Y)}.\]

Substituting the expectations, variances and covariances as computed in (3.38) yields after some simplifications:

\[
x^r = \frac{q_1^r - p_X}{\rho q_1^r(1 - q_1^r)} + y^r \frac{q_2^r}{1 - q_1^r} + l_1^r - l_2^r \frac{q_2^r}{1 - q_1^r}, \tag{3.42}
\]

and

\[
y^r = \frac{q_2^r - p_Y}{\rho q_2^r(1 - q_2^r)} + x^r \frac{q_1^r}{1 - q_2^r} - l_1^r \frac{q_1^r}{1 - q_2^r} + l_2^r. \tag{3.43}
\]
Let us plug $x^r$ defined in the first equation (3.42) into the second equation (3.43). This yields, after some manipulations:

$$y^r = \frac{(q_2^r - p_Y)(1 - q_1^r)}{\rho q_2^r(1 - q_1^r - q_2^r)} + \frac{q_1^r - p_X}{\rho(1 - q_1^r - q_2^r)} + l_2^r.$$  

This corresponds to equation (3.10) in Lemma 3.1.

Now, substituting $y^r$ into (3.42) yields after some manipulations:

$$x^r = \frac{(q_1^r - p_X)(1 - q_2^r)}{\rho q_1^r(1 - q_1^r - q_2^r)} + \frac{q_2^r - p_Y}{\rho(1 - q_1^r - q_2^r)} + l_1^r.$$  

This corresponds to equation (3.9) in Lemma 3.1.

Note that from these demands, it is straightforward to determine the demand of financial intermediaries. Both types of agents indeed have similar mean-variance utility functions. Equations (3.13) and (3.14) are obtained by replacing $\rho$ by $\rho_f$, $q_n^r$ by $q_n^p$, and $l_1^r = l_2^r = 0$.

**Lemma 3.2 (individual demand of the passive investor)**

Consider a passive investor $j$ with a mean-variance utility, and an information set $I^p$ (that we omit below for brevity) consisting of his liquidity shock, and his beliefs $(q_1^p, q_2^p, q_3^p)$. He is constrained in his trading quantities and can only adjust his portfolio in proportion. Given our notations, the expectation and variance of his final wealth defined in (3.11) would thus write:

$$E\left(\tilde{w}^j(\beta^j)\right) = (1 + \beta^j)\epsilon_X E(\tilde{d}_X) + (1 + \beta^j)\epsilon_Y E(\tilde{d}_Y) - p_X \beta \epsilon_X - p_Y \beta \epsilon_Y + \epsilon_C - E(\tilde{u}),$$  

(3.44)
and

\[
V \left( \bar{w}^j(\beta^j) \right) = (1 + \beta^j)^2(e^j_X)^2V(\bar{d}_X) + (1 + \beta^j)^2(e^j_Y)^2V(\bar{d}_Y) 
\]

(3.45)

\[
+ (l_1^j)^2V(1_{\omega_1}) + (l_2^j)^2V(1_{\omega_2})
\]

\[
+ 2(1 + \beta^j)^2e^j_X e^j_Y Cov(\bar{d}_X, \bar{d}_Y) - 2(1 + \beta^j)e^j_X l_1^j Cov(\bar{d}_X, 1_{\omega_1})
\]

\[
- 2(1 + \beta^j)e^j_Y l_2^j Cov(\bar{d}_Y, 1_{\omega_1}) - 2(1 + \beta^j)\gamma l_1^j Cov(\bar{d}_Y, 1_{\omega_1})
\]

\[
- 2(1 + \beta^j)e^j_Y l_2^j Cov(\bar{d}_Y, 1_{\omega_1}) + 2l_1^j l_2^j Cov(1_{\omega_1}, 1_{\omega_2}).
\]

Plugging (3.44) and (3.45) into (3.5) and taking the FOC with respect to \( \beta^j \) yields:

\[
\left( \rho(e^j_X)^2V(\bar{d}_X) + \rho(e^j_Y)^2V(\bar{d}_Y) + 2\rho e^j_X e^j_Y Cov(\bar{d}_X, \bar{d}_Y) \right) (1 + \beta^j)
\]

\[
= e^j_X \left( E(\bar{d}_X) - p_X \right) + e^j_Y \left( E(\bar{d}_Y) - p_Y \right) + \rho e^j_X l_1^j Cov(\bar{d}_X, 1_{\omega_1}) + \rho e^j_Y l_2^j Cov(\bar{d}_Y, 1_{\omega_1})
\]

\[
+ \rho e^j_X l_1^j Cov(\bar{d}_Y, 1_{\omega_1}) + \rho e^j_Y l_2^j Cov(\bar{d}_Y, 1_{\omega_1}).
\]

Substituting the expectations, variances and covariances as computed in (3.38) yields:

\[
1 + \beta^j = \frac{e^j_X (q^p_X - p_X) + e^j_Y (q^p_Y - p_Y) - \rho(e^j_X)^2 q^p_X(1 - q^p_X) + \rho(e^j_Y)^2 q^p_Y(1 - q^p_Y) - 2\rho e^j_X q^p_X q^p_Y}{\rho(e^j_X)^2 q^p_X(1 - q^p_X) + \rho(e^j_Y)^2 q^p_Y(1 - q^p_Y) - 2\rho e^j_X q^p_X q^p_Y + l_1^j \left( e^j_X q^p_X(1 - q^p_X) - e^j_X q^p_Y q^p_Y \right) + l_2^j \left( e^j_Y q^p_Y(1 - q^p_Y) - e^j_X q^p_X q^p_Y \right) - 2\rho e^j_X q^p_X q^p_Y}.
\]

Equation (3.12) in Lemma 3.2 follows from the simplification of this expression using the assumption (3.7).

**Proposition 3.1 (equilibrium prices)**

**Aggregate order flows**

We first derive the aggregate order flows \( \bar{z}_X \) and \( \bar{z}_Y \). From the the demand of active, unconstrained, and passive investors defined in Lemmas 3.1 and 3.2, and from that of financial interme-
However passive investors who receive a shock in uninformed investors who receive a shock in they hedge by buying asset $X$.

Under the assumption (3.7), this yields:

$$
\tilde{z}_X = \alpha \left( \frac{(q_1^0 - pX)(1 - q_2^0)}{\rho q_1^0 (1 - q_1^0 - q_2^0)} + \frac{q_2^0 - pY}{\rho (1 - q_1^0 - q_2^0)} \right) + \int_0^\alpha l_1^i \, di - \alpha e_1
$$

$$
+ \frac{q_2^0 - pY}{\rho q_1^0 (1 - q_1^0 - q_2^0)} + \int_\alpha^{\alpha+v} l_1^u \, dv - \alpha \nu e_1
$$

$$
+ (1 - \alpha - \nu) \left( \frac{(q_1^0 - pX)(1 - q_1^0)}{\rho q_1^0 (1 - q_1^0 - q_2^0)} + \frac{q_2^0 - pY}{\rho (1 - q_1^0 - q_2^0)} \right) - (1 - \alpha - \nu) e_1
$$

$$
+ \frac{q_2^0 - pY}{\rho q_1^0 (1 - q_1^0 - q_2^0)} + \int_{\alpha+v}^1 l_1^1 \, dj + \frac{q_2^0 (1 - q_1^0 - q_2^0)}{q_1^0 (1 - q_1^0) + q_2^0 (1 - q_2^0) - 2q_1^0 q_2^0} \int_\alpha^{\alpha+v} l_2^1 \, dj
$$

$$
+ \frac{m (q_1^0 - pX)(1 - q_2^0)}{\rho q_1^0 (1 - q_1^0 - q_2^0)} + \frac{q_2^0 - pY}{\rho (1 - q_1^0 - q_2^0)} = m e_1.
$$

A proportion $\sigma$ of active (resp. unconstrained uninformed) investors receive a shock in $\omega_1$, and they hedge by buying asset $X$. By contrast, the proportion $1 - \sigma$ of active (resp. unconstrained uninformed) investors who receive a shock in $\omega_1$ don’t need to hedge with $L$. Thus $\int_0^\alpha l_1^i \, di = \alpha \sigma L$.

However passive investors who receive a shock in $\omega_2$ need to trade asset $X$ to hedge. This yields $\int_\alpha^1 l_1^1 \, dj = (1 - \alpha)\sigma L$ and $\int_\alpha^1 l_2^1 \, dj (1 - \alpha) (1 - \sigma) L$. Equation (3.46) follows. Equation (3.47)
characterizing the order flow for asset $Y$ is derived similarly.

$$
\tilde{z}_X = \alpha \left( \frac{(q_1^a - p_X)(1 - q_2^a)}{\rho q_1^3 q_3^3} + \frac{q_2^a - p_Y}{\rho q_3^3} \right) + v \left( \frac{(q_1^p - p_X)(1 - q_2^p)}{\rho q_1^p q_3^3} + \frac{q_2^p - p_Y}{\rho q_3^3} \right)
$$

(3.46)

$$
\tilde{z}_Y = \alpha \left( \frac{(q_2^p - p_Y)(1 - q_1^p)}{\rho q_2^p q_3^3} + \frac{q_1^a - p_X}{\rho q_3^3} \right) + v \left( \frac{(q_2^p - p_Y)(1 - q_1^p)}{\rho q_2^p q_3^3} + \frac{q_1^p - p_X}{\rho q_3^3} \right)
$$

(3.47)

Notice that (3.46) and (3.47) do not rely on the restriction $\eta \to 0$. Now, for markets to clear, we need to have $z_X = 0$ and $z_Y = 0$.

**Market clearing condition 1**: $z_X = z_Y$

A necessary condition for market clearing is thus that $z_X = z_Y$ (both are equal to 0), which implies (under assumption (3.7)):

$$
\alpha \left( \frac{(q_1^a - p_X)(1 - q_2^a)}{\rho q_1^3 q_3^3} + \frac{q_2^a - p_Y}{\rho q_3^3} \right) + v \left( \frac{(q_1^p - p_X)(1 - q_2^p)}{\rho q_1^p q_3^3} + \frac{q_2^p - p_Y}{\rho q_3^3} \right)
$$

$$
= \alpha \left( \frac{(q_2^p - p_Y)(1 - q_1^p)}{\rho q_2^p q_3^3} + \frac{q_1^p - p_X}{\rho q_3^3} \right) + v \left( \frac{(q_2^p - p_Y)(1 - q_1^p)}{\rho q_2^p q_3^3} + \frac{q_1^p - p_X}{\rho q_3^3} \right)
$$

$$
+ m \left( \frac{(q_1^p - p_X)(1 - q_2^p)}{\rho f q_1^p q_3^3} + \frac{q_2^p - p_Y}{\rho f q_3^3} \right)
$$

$$
+ (\alpha + v)(1 - 2\tilde{\sigma})L.
$$
This rewrites:

\[
\frac{\alpha}{\rho q_1^p q_2^p} (q_1^p m - q_2^p p x) + \frac{1}{q_1^p q_2^p} \left( \frac{m}{\rho_f} + \frac{v}{\rho} \right) (q_1^p m - q_2^p p x) = (\alpha + v)(1 - 2\tilde{\sigma}) L,
\]

which yields:

\[
p_Y = \frac{\left( \frac{\alpha}{\rho q_1^p} + \frac{1}{q_1^p} \left( \frac{m}{\rho_f} + \frac{v}{\rho} \right) \right) p_X + (\alpha + v)(1 - 2\tilde{\sigma}) L}{\left( \frac{\alpha}{\rho q_2^p} + \frac{1}{q_2^p} \left( \frac{m}{\rho_f} + \frac{v}{\rho} \right) \right)}.
\]

Market clearing condition 2: \( z_X = 0 \)

We input \( p_Y \) defined in (3.48) into the condition \( z_X = 0 \) as follows:

\[
0 = \frac{\alpha}{\rho q_3^p} \left( \frac{(q_1^p - p X)(1 - q_2^p)}{q_1^p} + q_2^p - p_Y \right) + \left( \frac{v}{\rho} + \frac{m}{\rho_f} \right) \frac{1}{q_3^p} \left( \frac{(q_1^p - p X)(1 - q_2^p)}{q_1^p} + q_2^p - p_Y \right) \]

\[
+ \left( 1 - \alpha - v \right) \frac{p_Y}{\rho q_3^p} \left( \frac{(q_1^p - p X) + (q_2^p - p_Y)}{q_1^p + q_2^p} \right) \]
\[
+ (\alpha + v)\tilde{\sigma} L + \frac{q_1^p \sigma + q_2^p (1 - \tilde{\sigma})}{q_1^p + q_2^p} (1 - \alpha - v) L - e_1 - m e_1^f,
\]

After some manipulations:

\[
p_X \left( \frac{\alpha}{\rho q_3^p} \left( \frac{1 - q_2^p}{q_1^p} + \left( \frac{\alpha}{\rho q_1^p} + \frac{1}{q_1^p} \left( \frac{m}{\rho_f} + \frac{v}{\rho} \right) \right) \right) + \left( \frac{v}{\rho} + \frac{m}{\rho_f} \right) \frac{1}{q_3^p} \left( \frac{1 - q_2^p}{q_1^p} + \left( \frac{\alpha}{\rho q_1^p} + \frac{1}{q_1^p} \left( \frac{m}{\rho_f} + \frac{v}{\rho} \right) \right) \right) \right)
\]

\[
= (\alpha + v)\tilde{\sigma} L + \frac{q_1^p \sigma + q_2^p (1 - \tilde{\sigma})}{q_1^p + q_2^p} (1 - \alpha - v) L - e_1 - m e_1^f
\]

\[
+ \left( \frac{\alpha}{\rho q_3^p} + \left( \frac{v}{\rho} + \frac{m}{\rho_f} \right) \frac{1}{q_3^p} + \left( 1 - \alpha - v \right) \frac{1}{\rho q_3^p} \right) \left( \frac{1}{q_1^p + q_2^p} \right) \left( 1 - \frac{(\alpha + v)(1 - 2\tilde{\sigma}) L}{\left( \frac{\alpha}{\rho q_2^p} + \frac{1}{q_2^p} \left( \frac{m}{\rho_f} + \frac{v}{\rho} \right) \right)} \right)
\]

\[
- \frac{1 - \alpha - v}{\rho q_3^p} \left( 1 - q_1^p - q_2^p \right) \right).
\]
Let us now restrict our attention to $q_3^a = q_3^p = \eta$. Notice that this is always the case since $\varepsilon = 1 - 2\kappa$. At the limit when $\eta \to 0$, we will have:

$$p_X \left( 1 + \left( \frac{\alpha}{\rho q_1} + \frac{1}{q_1} \left( \frac{m}{\rho_f} + \frac{\nu}{\rho} \right) \right) \right) = \left( 1 - \frac{(\alpha + \nu)(1 - 2\sigma)L}{\left( \frac{\alpha}{\rho q_2} + \frac{1}{q_2} \left( \frac{m}{\rho_f} + \frac{\nu}{\rho} \right) \right)} \right),$$

which finally yields

$$p_X = \frac{\left( \frac{\alpha}{\rho q_2} + \frac{1}{q_2} \left( \frac{m}{\rho_f} + \frac{\nu}{\rho} \right) \right)}{\frac{\alpha}{\rho q_2} + \frac{1}{q_2} \left( \frac{m}{\rho_f} + \frac{\nu}{\rho} \right) + \frac{\alpha}{\rho q_1} + \frac{1}{q_1} \left( \frac{m}{\rho_f} + \frac{\nu}{\rho} \right)},$$

and

$$p_Y = \frac{\left( \frac{\alpha}{\rho q_1} + \frac{1}{q_1} \left( \frac{m}{\rho_f} + \frac{\nu}{\rho} \right) \right)}{\frac{\alpha}{\rho q_2} + \frac{1}{q_2} \left( \frac{m}{\rho_f} + \frac{\nu}{\rho} \right) + \frac{\alpha}{\rho q_1} + \frac{1}{q_1} \left( \frac{m}{\rho_f} + \frac{\nu}{\rho} \right)}. $$

We obtain the system of prices (3.15) in Proposition 3.1 by re-arranging terms in the two expressions above.

**Lemma 3.3 (condition for non revealing order flow)**

First, let us look for conditions to characterize a non fully revealing equilibrium. On the one hand, the aggregate demand for asset $X$ and $Y$ depends on the realization of the signal $\tilde{s}$ that active investors received. On the other hand, financial intermediaries and passive investors condition their demand on prices thus on the order flow. Let us conjecture that if a non fully revealing equilibrium exists, it should be such that financial intermediaries cannot distinguish between cases $\tilde{s} = s_1$ and $\tilde{\sigma} = \sigma_H$, and $\tilde{s} = s_1$ and $\tilde{\sigma} = \sigma_L$, that is:

$$p_X(s_1, \sigma_L) = p_X(s_2, \sigma_H),$$

$$p_Y(s_1, \sigma_L) = p_Y(s_2, \sigma_H).$$

Following our assumptions on the symmetry of the distribution of $\tilde{\sigma}$, financial intermediaries
would thus have the following beliefs if \( p_X(s_1, \sigma_L) = p_X(s_2, \sigma_H) \):

\[
q_1^p = q_2^p = \Pr(\tilde{\omega} = \omega_1 | \tilde{s} = s_1, \tilde{s} = s_2),
\]

which we denoted \( \kappa \) in system (3.39). Let us define

\[
\Lambda_N \equiv \frac{(\frac{m}{\rho_f} + \frac{v}{\rho})(\theta - \frac{1}{2}) + (\alpha + v)(\sigma_H - \sigma_L)\frac{1}{4}L}{\alpha + (\frac{m}{\rho_f} + \frac{v}{\rho})\theta(1 - \theta)}. \tag{3.49}
\]

When \( \eta \to 0 \), and if prices were not fully revealing, then prices defined in (3.15) would write:

\[
p_X(s_1, \sigma_L) = \theta - \theta(1 - \theta)\Lambda_N,
\]
\[
p_Y(s_1, \sigma_L) = 1 - \theta + \theta(1 - \theta)\Lambda_N,
\]

and

\[
p_X(s_2, \sigma_H) = 1 - \theta + \theta(1 - \theta)\Lambda_N,
\]
\[
p_Y(s_2, \sigma_H) = \theta - \theta(1 - \theta)\Lambda_N.
\]

The condition \( p_X(s_1, \sigma_L) = p_X(s_2, \sigma_H) \) becomes:

\[
\theta - \theta(1 - \theta)\Lambda_N = 1 - \theta + \theta(1 - \theta)\Lambda_N.
\]

We substitute \( \Lambda_N \) by its expression defined in (3.49), which yields

\[
\frac{\alpha}{\rho} (2\theta - 1) = 2\theta(1 - \theta)(\alpha + v)(\sigma_H - \sigma_L)L.
\]

Finally, since \( \theta \to \frac{2\tau}{1+\tau} \) when \( \eta \to 0 \), this is equivalent to

\[
\frac{(3\tau - 1)(1 + \tau)}{4\rho\tau(1 - \tau)} \frac{\alpha}{(\alpha + v)} = (\sigma_H - \sigma_L)L.
\]
Notice that whenever \( \nu > 0 \), the condition depends on \( \alpha \) and \( \nu \). Notice that this condition also implies that \( p_Y(s_1, \sigma_L) = p_Y(s_2, \sigma_H) \). Condition (3.17) follows.

Second, let us show that there exists no other not fully revealing equilibrium. Notice that when \( z_X(s_3, \sigma_H) = z_X(s_2, \sigma_L) \), then \( z_Y(s_3, \sigma_H) \neq z_Y(s_2, \sigma_L) \) and vice versa. Thus financial intermediaries can always distinguish whether \( \tilde{s} = s_1 \) or \( \tilde{s} = s_3 \). By similar argument, we can also rule out the existence of a non fully revealing equilibrium such that the market makers would not disentangle between \( \tilde{s} = s_1 \) and \( \tilde{s} = s_3 \). Our conjecture that if a non fully revealing equilibrium exists, it should be such that financial intermediaries cannot distinguish between cases \( \tilde{s} = s_1 \) and \( \tilde{\sigma} = \sigma_H \), and \( \tilde{s} = s_1 \) and \( \tilde{\sigma} = \sigma_L \), is thus verified.

Finally, to interpret the NFR condition, let us compute the difference in the demand in asset \( X \) when the active investor observes \( \tilde{s} = s_1 \) or \( \tilde{s} = s_2 \).

\[
x^i(\tilde{s} = s_1) - x^i(\tilde{s} = s_2) = \frac{p_X (3\tau - 1)(1 + \tau)}{2\tau(1 - \tau)},
\]

which corresponds to the left-hand-side of condition (3.17) when \( p_X = 1/2 \).

The following property is worth noticing. Under the NFR condition (3.17), \( \Lambda_N \) defined in (3.49) rewrites (after some manipulations):

\[
\Lambda_{NFR}^N = \frac{(3\tau - 1)(1 + \tau)}{4\tau(1 - \tau)}.
\]  

(3.50)

**Proposition 3.2 (welfare in the non fully revealing equilibrium)**

To prove Proposition 3.2, we will proceed in three steps. First, we determine the investors’ final wealth, given their optimal demand for the assets at date 2. Second, we compute the expectation and variance of this final wealth at date 2, conditional on the realization of signals and liquidity shocks. Third, we compute their expected mean variance utility at date 1, defined as

\[
U_1 = E_1 \left[ E_2 (\bar{w}^r) - \frac{\rho}{2} V_2 (\bar{w}^r) \right],
\]
where $E_1$ the expectation at date 1, that is, prior to the realization of signals and shocks.

**Step 1. Final wealth at equilibrium**

The following result reports the final wealth of investors based on their optimal demands.

The final wealth of an unconstrained investor at equilibrium is

$$\tilde{w}^r = \frac{p_X}{\rho q_1^2} 1_{\omega_1} - \frac{p_Y}{\rho q_2^2} 1_{\omega_2} + \frac{(p_X)^2}{\rho q_1^2} + \frac{(p_Y)^2}{\rho q_2^2} + p_X (x_1^r - l_1^r) + p_Y (y_1^r - l_2^r) + e_C^r, \quad (3.51)$$

where $\Lambda$ is defined in (3.16). The final wealth of a passive investor at equilibrium is

$$\tilde{w}^j = (e_j^X - l_1^j) 1_{\omega_1} + (e_j^Y - l_2^j) 1_{\omega_2} + e_C^j. \quad (3.52)$$

To get the result of the unconstrained investor, we first re-write the demand of the unconstrained and constrained investors as:

$$x^u = \frac{1}{\rho} \frac{1 - p_X - p_Y}{q_3^u} - \frac{p_X}{\rho q_1^u} + l_1^u \quad (3.53)$$

$$y^u = \frac{1}{\rho} \frac{1 - p_X - p_Y}{q_3^u} - \frac{p_Y}{\rho q_2^u} + l_2^u \quad (3.54)$$

$$(1 + \beta^j) e_1 = \frac{1}{\rho} \frac{1 - p_X - p_Y}{q_3^j} - \frac{1}{\rho} \frac{q_1^{pj} l_1^j + q_2^{pj} l_2^j}{q_1^p + q_2^p} \quad (3.55)$$

We then plug the demand in assets $X$ and $Y$ given by the above equations into the wealth of the unconstrained investor, defined in equation (3.8). We use two properties of our model, namely i. $d_X = 1_{\omega_1}$ and $d_Y = 1_{\omega_2}$, and ii. $\bar{r} = l_1^r 1_{\omega_1} + l_2^r 1_{\omega_2}$.

$$\tilde{w}^r = \left( \frac{(q_1^r - p_X)(1 - q_3^1)}{\rho q_1^r q_3^1} + \frac{q_2^r - p_Y}{\rho q_3^2} \right) (1_{\omega_1} - p_X) + \left( \frac{(q_2^r - p_Y)(1 - q_3^1)}{\rho q_2^r q_3^1} + \frac{q_1^r - p_X}{\rho q_3^2} \right) (1_{\omega_2} - p_Y) + p_X (e_x^r - l_1^r) + p_Y (e_Y^r - l_2^r) + e_C^r. \quad (3.56)$$
When $\eta \to 0$, the no arbitrage condition in (3.26) yields

\[
\tilde{w}^r = -\frac{p_X}{\rho q_1^p} 1_{\omega_1} - \frac{p_Y}{\rho q_2^p} 1_{\omega_2} + \frac{(p_X)^2}{\rho q_1^p} + \frac{(p_Y)^2}{\rho q_2^p} + p_X (e_X^r - l_X^r) + p_Y (e_Y^r - l_Y^r) + e_C^r \\
+ \frac{y}{\rho} (1_{\omega_1} + 1_{\omega_2} - p_X - p_Y). \tag{3.57}
\]

where $y = \frac{1 - p_X - p_Y}{q_3^p}$. One can show that $\lim_{\eta \to 0} y$ is finite. Hence, equation (3.51) in Result 3.7 is obtained under $1_{\omega_1} + 1_{\omega_2} - p_X - p_Y = 0$ when $\eta \to 0$.

To get the result of the passive investor, we use two properties of our model, namely i. $\tilde{d}_X = 1_{\omega_1}$ and $\tilde{d}_Y = 1_{\omega_1}$, and ii. $\tilde{\omega}^i = l_1^i 1_{\tilde{\omega}_1} + l_2^i 1_{\tilde{\omega}_2}$. In addition, we use the assumption that $e_{X}^i = e_{Y}^i = e_1$. Rearranging terms in (3.11), the passive investor’s wealth writes

\[
\tilde{w}^j(\beta^j) = \beta^j e_X^j (1_{\omega_1} - p_X) + \beta^j e_Y^j (1_{\omega_2} - p_Y) + e_C^j + (e_X^j - l_X^j) 1_{\omega_1} + (e_Y^j - l_Y^j) 1_{\omega_2}.
\]

We plug the passive investor’s demand defined in (3.12) into this equation. Equation (3.58) follows, and we use this equation to characterize the passive investor’s final wealth in the various cases.

\[
\tilde{w}^j(\beta^j) = \left( \frac{q_{1}^{p} - p_X}{\rho q_{1}^{p}} + \frac{q_{2}^{p} - p_Y}{\rho q_{2}^{p}} \right) (1_{\omega_1} - p_X + 1_{\omega_2} - p_Y) + e_C^j + (e_X^j - l_X^j) 1_{\omega_1} + (e_Y^j - l_Y^j) 1_{\omega_2} - e_X^j (1_{\omega_1} - p_X) - e_Y^j (1_{\omega_2} - p_Y). \tag{3.58}
\]

Ultimately, when $\eta \to 0$, $p_X + p_Y \to 1$ and we get (3.52).

**Step 2. Expected utility at date 2**

Prices contain information on the realization of $\tilde{\sigma}$ and the beliefs of uninformed agents. By conditioning on prices at date 2, investors condition on the realization of $\tilde{\Lambda}$ that depends on $\tilde{\sigma}$ and
s, namely Λ. Expected utilities at date 2 write as follows:

\[ U_2(\tilde{w}|p_2, l^u) = e^u_C + \left(1 - \frac{q_a^u}{2}\right) \frac{(p_X)^2}{\rho q_1^u} + \left(1 - \frac{q_a^u}{2}\right) \frac{(p_Y)^2}{\rho q_2^u} \]
\[ + \frac{pxp_Y - 1}{\rho} + px(e_X - l_1^u) + py(e_Y - l_2^u). \] (3.59)

\[ U_2(\tilde{w}|p_2, l^t) = e^t_C + \frac{\Lambda^2}{2\rho} + q_a^t (1 - q_a^u) e^{t}_X - l_1^t \]
\[ + q_a^t (1 + q_a^u) (e_Y - l_2^t). \] (3.60)

\[ U_2(\tilde{w}|p_2, l^f) = e^f_C + (e^X - l_1^f)q_a^f + (e^Y - l_2^f)q_a^f \]
\[ - \frac{p}{2} \left((e^X - l_1^f) - (e^Y - l_2^f)\right)^2 q_a^f. \] (3.61)

\[ U_2(\tilde{w}|p_2, l^p) = \frac{1}{\rho_f} \left(1 - \frac{q_a^u}{2}\right) \frac{(1 - q_a^u)\Lambda}{\rho_f q_1^p} + \frac{(1 + q_a^u)\Lambda}{\rho_f q_2^p} \]
\[ + q_a^u (1 - q_a^u) e^f_X + q_a^t (1 + q_a^u) e^f_Y \]
\[ - \frac{1}{2\rho_f} \left(\frac{(1 - q_a^u)\Lambda}{q_1^u} - \frac{(1 + q_a^u)\Lambda}{q_2^u}\right)^2 q_1^p q_2^p. \] (3.62)

Let’s derive these results.

**Unconstrained investors**

It follows that at date 2, the expectation and variance of an unconstrained investor’s wealth given in (3.51) in Result 3.7 write

\[ E_2(\tilde{w}^u|p_2, l^u) = -\frac{1}{\rho} + \frac{(p_X)^2}{\rho q_1^u} + \frac{(p_Y)^2}{\rho q_2^u} + px(e_X - l_1^u) + py(e_Y - l_2^u) + e^u_C, \]

and

\[ V_2(\tilde{w}^u|p_2, l^u) = \left(\frac{p_X}{\rho q_1^u} - \frac{p_Y}{\rho q_2^u}\right)^2 q_1^u q_2^u, \]

which finally yields (3.59).

**Active investors**

For the active investor, given his beliefs \((q_a^u)\) and prices defined in (3.15), (3.59) simplifies as
follows:

\[ E_2 (\tilde{w}^i | p_2, l^i, s) = q_1^a q_2^a \left( \frac{\Lambda^2}{\rho} \right) + q_1^a (1 - q_2^a \Lambda) (e_X^i - l_1^i) + q_2^a (1 + q_1^a \Lambda) (e_Y^i - l_2^i) + e_C^i, \]

and

\[ V_2 (\tilde{w}^i | p_2, l^i, s) = \frac{q_1^a q_2^a \Lambda^2}{\rho^2}, \]

which finally yields (3.60).

**Passive investors**

Similarly the expectation and variance at date 2 of a passive investor’s final wealth when \( \eta \to 0 \) given in (3.52) of Result 3.7 write

\[ E_2 (\tilde{w}^j (\beta^j) | p_2, l^j) = e_C^j + (e_X^j - l_1^j)q_1^p + (e_Y^j - l_2^j)q_2^p, \]

and

\[ V_2 (\tilde{w}^j (\beta^j) | p_2, l^j) = (e_X^j - l_1^j)^2 q_1^p (1 - q_1^p) + (e_Y^j - l_2^j)^2 (1 - q_2^p) q_2^p - 2(e_X^j - l_1^j)(e_Y^j - l_2^j) q_1^p q_2^p. \]

When \( \eta \to 0 \),

\[ V_2 (\tilde{w}^j (\beta^j) | p_2, l^j) \to ( (e_X^j - l_1^j) - (e_Y^j - l_2^j) )^2 q_1^p q_2^p. \]

which finally yields (3.61).

**Financial intermediaries**

Finally, the expectation and variance at date 2 of a financial intermediary’s wealth when \( \eta \to 0 \) given in Result 3.7 write

\[ E_2 (\tilde{w}^f | p_2) = e_C^f - \frac{1}{\rho_f} + \frac{(q_1^a)^2 (1 - q_2^a \Lambda)^2}{\rho_f q_1^p} + \frac{(q_2^a)^2 (1 + q_1^a \Lambda)^2}{\rho_f q_2^p} + q_1^a (1 - q_2^a \Lambda) e_X^f + q_2^a (1 + q_1^a \Lambda) e_Y^f, \]
and

\[ V_2 \left( \tilde{w}^f | p_2 \right) = \left( \frac{q_1^a (1 - q_2^a \Lambda)}{\rho \tilde{w}^F} - \frac{q_2^a (1 + q_1^a \Lambda)}{\rho \tilde{w}^F} \right)^2 q_1^p q_2^p, \]

which finally yields (3.62).

**Step 3. Expected mean variance utilities at date 1**

Proposition 3.2 characterizes the investors’ expected mean variance utility at date 1 when the NFR condition (3.17) holds. In this third 3, we compute the expected value at date 1 of the expected utilities at date 2. The expectation at date 1 accounts for all possible realizations of signals and liquidity shocks at date 2. We cluster realizations into cases.

In Cases 1, 2 and 3, prices are fully revealing but prices have different informational content. Let us consider \( \Lambda_F \) defined in (3.19) as

\[
\Lambda_F \equiv \frac{\rho \rho_f (\alpha + \nu)}{m \rho + (\nu + \alpha) \rho_f} (\sigma_H - \sigma_L) L.
\]

Whenever prices are fully revealing and when \( \eta \to 0 \), then prices defined in (3.15) write:

\[
p_X = q_1^a - q_1^a q_2^a \Lambda_F \left( 1_{\tilde{s} = \sigma_L} - 1_{\tilde{s} = \sigma_H} \right), \tag{3.63}
\]

\[
p_Y = q_2^a + q_1^a q_2^a \Lambda_F \left( 1_{\tilde{s} = \sigma_L} - 1_{\tilde{s} = \sigma_H} \right).
\]

In Case 4, prices are not fully revealing: the realizations of \( \tilde{s} \) and \( \tilde{\sigma} \) are either \((s_1, \sigma_L)\) or \((s_2, \sigma_H)\). In the latter case, when \( \eta \to 0 \), then prices defined in (3.15) write:

\[
p_X(s_1, \sigma_L) = \theta - \theta (1 - \theta) \Lambda_N, \tag{3.64}
\]

\[
p_Y(s_1, \sigma_L) = 1 - \theta + \theta (1 - \theta) \Lambda_N.
\]
and

\[ p_X(s_2, \sigma_H) = 1 - \theta + \theta(1 - \theta)\Lambda_N, \]
\[ p_Y(s_2, \sigma_H) = \theta - \theta(1 - \theta)\Lambda_N. \]

(3.65)

where \( \Lambda_N \) is defined in (3.49).

**Active investors**

_**Case 1:** \( \tilde{s} = s_1, \tilde{\sigma} = \sigma_H, q_1^p = \theta, q_2^p = 1 - \theta \) (with probability \( \frac{1+\tau}{8} \))

If \( \tilde{s} = s_1 \) and \( \tilde{\sigma} = \sigma_H \) then the equilibrium is perfectly revealing that \( \tilde{s} = s_1 \), which yields \( q_1^a = q_1^p = \theta \) and \( q_2^a = q_2^p = 1 - \theta - \varepsilon \). When \( \eta \to 0 \), inputting \( q_1^a = 1 - \theta \) and prices defined in (3.63) conditional on \( \tilde{\sigma} = \sigma_H \) into (3.60), we get:

\[
U_2 (\tilde{w}^i|p_2, l_i, s_1, \text{ case 1}) = e_C^i + \theta(1 - \theta) \left( \frac{\Lambda_F^2}{2\rho} \right)
+ \theta(e_X^i - l_1^i) + (1 - \theta)(e_Y^i - l_2^i) + \theta(1 - \theta)(e_X^i - e_Y^i + e_Y^i + l_2^i)\Lambda_F
\]

_**Case 2:** \( \tilde{s} = s_2, \tilde{\sigma} = \sigma_L, q_1^p = 1 - \theta, q_2^p = \theta \) (with probability \( \frac{1+\tau}{8} \))

If \( \tilde{s} = s_2 \) and \( \tilde{\sigma} = \sigma_L \) then the equilibrium is perfectly revealing that \( \tilde{s} = s_2 \), which yields \( q_2^a = \theta \) and \( q_1^a = 1 - \theta - \varepsilon \). When \( \eta \to 0 \), inputting \( q_1^a = 1 - \theta \) and prices defined in (3.63) conditional on \( \tilde{\sigma} = \sigma_L \) into (3.60), we get:

\[
E_2 (\tilde{w}^i|p_2, l_i, s_2, \text{ case 2}) = e_C^i + \theta(1 - \theta) \left( \frac{\Lambda_F^2}{2\rho} \right)
+ (1 - \theta)(e_X^i - l_1^i) + \theta(1 - \theta)(e_X^i - l_1^i - e_Y^i + l_2^i)\Lambda_F
\]

_**Case 3:** \( \tilde{s} = s_3, q_1^a = q_2^a = q_1^p = q_2^p = \chi \) (with probability \( \frac{1+\tau}{2} \))

If \( \tilde{s} = s_3 \) then the equilibrium is perfectly revealing that \( \tilde{s} = s_3 \), which yields \( q_1^a = q_2^a = \chi \). With probability \( \frac{1}{2} \), \( \tilde{\sigma} = \sigma_H \). When \( \eta \to 0 \), inputting \( q_1^a = \frac{1}{2} \) and prices defined in (3.63)
conditional on $\tilde{\sigma} = \sigma_H$ into (3.60), we get:

$$U_2 (\tilde{w}^i|p_2, l_i, s_3, \text{case 3a}) = e^i_C + \frac{1}{4} \left( \frac{\Lambda^2_N}{2 \rho} \right) + \frac{1}{2} (e^i_X - l^i_1) + \frac{1}{2} (e^i_Y - l^i_2) + \frac{1}{4} (e^i_X - l^i_1 - e^i_Y + l^i_2) \Lambda_F$$

With probability $\frac{1}{2}$, $\tilde{\sigma} = \sigma_L$. When $\eta \to 0$, inputting $q^a_1 = \frac{1}{2}$ and prices defined in (3.63) conditional on $\tilde{\sigma} = \sigma_L$ into (3.60), we get:

$$U_2 (\tilde{w}^i|p_2, l_i, s_3, \text{case 3b}) = e^i_C + \frac{1}{4} \left( \frac{\Lambda^2_N}{2 \rho} \right) + \frac{1}{2} (e^i_X - l^i_1) + \frac{1}{2} (e^i_Y - l^i_2) - \frac{1}{4} (e^i_X - l^i_1 - e^i_Y + l^i_2) \Lambda_F$$

**Case 4:** $\tilde{s} = s_1, \tilde{\sigma} = \sigma_L$ or $\tilde{s} = s_2, \tilde{\sigma} = \sigma_H$, and $q^a_1 = q^a_2 = \kappa$ (with probability $\frac{1+\tau}{4}$)

In Case 4, prices are not fully revealing. With probability $\frac{1}{2}$, active investors observe $\tilde{s} = s_1$ and $q^a_1 = \theta, q^a_2 = 1 - \theta$. They infer from prices that $\tilde{\sigma} = \sigma_L$. Prices are defined in (3.64) for $\tilde{\sigma} = \sigma_L$. Inputting $q^a_1 = \theta$ into equation (3.60) yields

$$U_2 (\tilde{w}^i|p_2, l_i, s_1, \text{NFR, case 4a}) = e^i_C + \theta (1 - \theta) \left( \frac{(\Lambda_N)^2}{2 \rho} \right) + \theta (1 - (1 - \theta) \Lambda_N) (e^i_X - l^i_1) + (1 - \theta) (1 + \theta \Lambda_N) (e^i_Y - l^i_2)$$

With probability $\frac{1}{2}$, active investors observe $\tilde{s} = s_2$ and $q^a_1 = 1 - \theta, q^a_2 = \theta$. They infer from prices that $\tilde{\sigma} = \sigma_H$. Prices are defined in (3.65) for $\tilde{\sigma} = \sigma_H$. Inputting $q^a_1 = 1 - \theta$ into equation (3.60) yields

$$U_2 (\tilde{w}^i|p_2, l_i, s_2, \text{NFR, case 4b}) = e^i_C + \theta (1 - \theta) \left( \frac{\Lambda^2_N}{2 \rho} \right) + (1 - \theta) (1 + \theta \Lambda_N) (e^i_X - l^i_1) + \theta (1 - (1 - \theta) \Lambda_N) (e^i_Y - l^i_2)$$

We aggregate over all cases to compute the agents’ expected mean variance utility at date 1.
that is, before receiving information or observing prices. $W_i^i(NFR)$ is the expectation of the active
investor’s expected utilities at date 2 in the non fully revealing equilibrium in the different cases.
Date 1 probabilities used to compute this expectation take into account the probability with which
the investor receives a liquidity shock in state $\omega_1$ or $\omega_2$ depending on the realization of $\tilde{\sigma}$.

Notice that the contribution of the expected utilities in Cases 1, 2 or 3 to the investor’s expected
mean variance utility at date 1 does not depend on whether the parameters are such that the equi-
librium is fully revealing or not. Only the contribution of the expected utilities in Case 4 will thus
change.

The realization of $\tilde{\sigma}$ influences the probability with which the investor will receive a shock in
state $\omega_1$ or $\omega_2$. In cases 1, 3a, and 4b, we have $\tilde{\sigma} = \sigma_H$ while in cases 2, 3b and 4a we have $\tilde{\sigma} = \sigma_L$.
This yields:

$$W_i^i(NFR) = \frac{1 + \tau}{8} (\sigma_H U_2 (\tilde{w}^i|p_2, l_1 = L, s_1, \text{case 1}) + (1 - \sigma_H) U_2 (\tilde{w}^i|p_2, l_2 = L, s_1, \text{case 1}))$$
$$+ \frac{1 - \tau}{8} (\sigma_L U_2 (\tilde{w}^i|p_2, l_1 = L, s_2, \text{case 2}) + (1 - \sigma_L) U_2 (\tilde{w}^i|p_2, l_2 = L, s_2, \text{case 2}))$$
$$+ \frac{1 - \tau}{4} (\sigma_H U_2 (\tilde{w}^i|p_2, l_1 = L, s_3, FR, \text{case 3a}) + (1 - \sigma_H) U_2 (\tilde{w}^i|p_2, l_2 = L, s_3, \text{case 3a}))$$
$$+ \frac{1 - \tau}{4} (\sigma_L U_2 (\tilde{w}^i|p_2, l_1 = L, s_3, \text{case 3b}) + (1 - \sigma_L) U_2 (\tilde{w}^i|p_2, l_2 = L, s_3, \text{case 3b}))$$
$$+ \frac{1 + \tau}{8} (\sigma_L U_2 (\tilde{w}^i|p_2, l_1 = L, s_1, NFR, \text{case 4a}) + (1 - \sigma_L) U_2 (\tilde{w}^i|p_2, l_2 = L, s_1, NFR, \text{case 4a}))$$
$$+ \frac{1 + \tau}{8} (\sigma_H U_2 (\tilde{w}^i|p_2, l_1 = L, s_2, NFR, \text{case 4b}) + (1 - \sigma_H) U_2 (\tilde{w}^i|p_2, l_2 = L, s_2, NFR, \text{case 4b}))$$

We replace expected utilities in each case using computations in Step 2 and take $\eta \to 0$. Compu-
tations yield

$$W_i^i(NFR) = e^i_C + \frac{1}{2} (e^i_X + e^i_Y - L) - \frac{1 + \tau}{4} \theta (1 - \theta) (\sigma_H - \sigma_L) L (\Lambda_F + \Lambda_N)$$
$$- \frac{1 - \tau}{8} (\sigma_H - \sigma_L) L \Lambda_F + \frac{1 + \tau}{8} \theta (1 - \theta) \left( \frac{\Lambda_F^2}{\rho} + \frac{\Lambda_N^2}{\rho} \right) + \frac{1 - \tau}{16} \left( \frac{\Lambda_F^2}{\rho} \right)$$

163
Since $\theta \to \frac{2\tau}{1+\tau}$ as $\eta \to 0$, we have $\theta(1-\theta) = \frac{2\tau(1-\tau)}{(1+\tau)^2}$, which yields:

\[
W^i(NFR) = e^i_C + \frac{1}{2}(e^i_X + e^i_Y - L) - \frac{\tau(1-\tau)}{2(1+\tau)} \left( (\Lambda_F + \Lambda_N) + \frac{1+\tau}{4\tau} \Lambda_F \right) (\sigma_H - \sigma_L)L + \frac{\tau(1-\tau)}{4(1+\tau)} \left( \frac{\Lambda_N^2}{\rho} + \frac{1+5\tau}{4\tau} \frac{\Lambda_F^2}{\rho} \right)
\]

(3.66)

To interpret the result, it is interesting to notice that under the NFR condition (3.17), $\Lambda_N = \frac{(3\tau-1)(1+\tau)}{4\tau(1-\tau)}$ and $(\sigma_H - \sigma_L)L = \frac{(3\tau-1)(1+\tau)}{4\tau(1-\tau)} \frac{\alpha}{(\alpha+\nu)}$, which yields

\[
W^{i,NFR\text{Cond}}(NFR) = e^i_C + \frac{1}{2}(2e_1 - L) - \frac{(1-\tau)(5\tau+1)}{8(1+\tau)} \left( (\sigma_H - \sigma_L) \Lambda_F - \frac{\Lambda_F^2}{2\rho} \right) + \frac{\nu - \alpha (1+\tau)(3\tau - 1)^2}{\nu + \alpha} \frac{64\rho \tau (1-\tau)}{}
\]

Since

\[
(\sigma_H - \sigma_L) \Lambda_F - \frac{\Lambda_F^2}{2\rho} = \left( \frac{m\rho}{m\rho + \nu \rho_f + \alpha \rho_f} (\sigma_H - \sigma_L) \Lambda_F + \frac{\Lambda_F^2}{2\rho} \right)
\]

It follows that

\[
W^{i,NFR\text{Cond}}(NFR) = c_1 + \frac{1}{2}(2e_1 - L) - \frac{(1-\tau)(5\tau+1)}{8(1+\tau)} \left( \frac{m\rho}{m\rho + \nu \rho_f + \alpha \rho_f} (\sigma_H - \sigma_L) \Lambda_F + \frac{\Lambda_F^2}{2\rho} \right) + \frac{\nu - \alpha (1+\tau)(3\tau - 1)^2}{\nu + \alpha} \frac{64\tau \rho (1-\tau)}{}
\]

**Uninformed, unconstrained investors**

**Cases 1, 2 and 3:**

Whenever prices are fully revealing, we have $q_s^u = q_s^a$. Since active and unconstrained, uninformed investors have the same final wealth and the same information set, their expected mean variance utilities at date 2 are similar. We input these beliefs into prices and plug in $p_X$ and $p_Y$ into
We finally get:

\[
U_2(\tilde{w}^u|p_2, l_u, s, \text{cases } 1 - 2 - 3) = e_C + q_1^a q_2^a \left( \frac{\Lambda_F^2}{2 \rho} \right) + q_1^a (e_X^u - l_1^u) + q_2^a (e_Y^u - l_2^u) - q_1^a q_2^a (e_X^u - l_1^u - e_Y^u + l_2^u) \Lambda_N (1 - \tilde{\sigma} = \sigma_L - 1 - \tilde{\sigma} = \sigma_H) \tag{3.67}
\]

**Case 4**: \( q_1^u = q_2^u = \kappa \) (with probability \( \frac{1 + \tau}{4} \))

When prices are not fully revealing in the two subcases below, unconstrained uninformed traders cannot perfectly infer from the order flow whether the active investor is characterized by \( \tilde{s} = s_1 \) and \( \tilde{\sigma} = \sigma_L \), or \( \tilde{s} = s_2 \) and \( \tilde{\sigma} = \sigma_H \). Indeed, for a given realization of \( \tilde{\omega} \), the same order flow is observed in both cases. Accordingly, the uninformed traders’ beliefs are such that: \( q_1^u = q_2^u = \kappa = \frac{1}{2} \), and (3.59) yields

\[
U_2(\tilde{w}^u|p_2, NFR, l^u) = e_C + e_C^u - q_1^a (1 - q_2^a \Lambda_{\tilde{\sigma}, Q}) l_1^u - q_2^a (1 + q_1^a \Lambda_{\tilde{\sigma}, Q}) l_2^u + \frac{1}{2 \rho} - \frac{2q_1^a q_2^a (1 - q_2^a \Lambda_{\tilde{\sigma}, Q}) (1 + q_1^a \Lambda_{\tilde{\sigma}, Q})}{\rho} \tag{3.68}
\]

With probability \( \frac{1}{2} \), active investors observe \( \tilde{s} = s_1 \) (so \( q_1^a = \theta, q_2^a = 1 - \theta \)) and they may infer from prices that \( \tilde{\sigma} = \sigma_L \). In this case, the price impact is \( \Lambda_N \) and inputting \( q_1^a = \theta \) into equation (3.68) yields

\[
U_2(\tilde{w}^u|p_2, l^u, \text{case 4a}) = e_1 + e_C^u - \theta (1 - (1 - \theta) \Lambda_N) l_1^u - (1 - \theta) (1 + \theta \Lambda_N) l_2^u + \frac{1}{2 \rho} - \frac{2\theta(1 - \theta)(1 - (1 - \theta) \Lambda_N)(1 + \theta \Lambda_N)}{\rho} \tag{3.69}
\]

With probability \( \frac{1}{2} \), active investors observe \( \tilde{s} = s_2 \) (so \( q_1^a = 1 - \theta, q_2^a = \theta \)) and they may infer
from prices that $\bar{\sigma} = \sigma_H$. Inputting $q_i^c = 1 - \theta$ into equation (3.68) yields

$$
U_2 \left( \tilde{w}^u | p_2, l^u, \text{case 4b} \right) = e_1 + e_C^u - (1 - \theta) (1 + \theta \Lambda_N) l_1^u - \theta (1 - (1 - \theta) \Lambda_N) l_2^u \\
+ \frac{1}{2\rho} \left[ 2(1 - \theta) \theta (1 + \theta \Lambda_N) (1 - (1 - \theta) \Lambda_N) \right] (3.70)
$$

In the end, when prices are not fully revealing, $U_2(\tilde{w}^u | p_2, l^u, \text{case 4})$ for uninformed, unconstrained investors writes:

$$
U_2(\tilde{w}^u | p_2, l^u, \text{case 4}) = \frac{1 + \tau}{4} (e_1 + e_C^u) + \frac{1 + \tau}{4} \frac{1}{2\rho} \\
- \frac{1 + \tau}{4} \frac{2\theta(1 - \theta)}{\rho} \left( 1 - (1 - 2\theta) \Lambda_N - \theta(1 - \theta)(\Lambda_N)^2 \right) \\
- \frac{1 + \tau}{4} \left( \sigma_L \theta + \sigma_H (1 - \theta) + (\sigma_H - \sigma_L) \theta(1 - \theta) \Lambda_N \right) L \quad (3.71)
$$

We now compute the expectation at date 1 of the unconstrained, uninformed investor’s utilities at date 2. This yields:

$$
W^u(NFR) = (e_1 + e_C^u) - \frac{L}{2} + \frac{1 + \tau}{4} \left( \frac{1}{2\rho} \right) \\
+ \left( \frac{1 + \tau}{4} \theta(1 - \theta) + \frac{1 - \tau}{8} \right) \frac{\Lambda_F^2}{2\rho} \\
- \frac{1 + \tau}{2} \frac{\theta(1 - \theta)}{\rho} \left( 1 - (1 - 2\theta) \Lambda_N - \theta(1 - \theta)(\Lambda_N)^2 \right) \\
+ \left( \sigma_H - \sigma_L \right) \left( -\frac{1 - \tau}{8} \Lambda_F - \frac{1 + \tau}{4} \theta(1 - \theta) (\Lambda_N + \Lambda_F) \right) L
$$

Since $\theta = \frac{2\tau}{1+\tau}$, $1 - \theta = \frac{1-\tau}{1+\tau}$, and $\theta(1 - \theta) = \frac{2\tau(1-\tau)}{(1+\tau)^2}$, we have

$$
W^u(NFR) = e_1 + e_C^u - \frac{L}{2} + \left( \sigma_H - \sigma_L \right) \left( -\frac{1 - \tau}{8} \Lambda_F - \frac{\tau(1 - \tau)}{2(1 + \tau)} (\Lambda_F + \Lambda_N) \right) L \\
+ \left( \frac{(3\tau - 1)^2}{8\rho(1 + \tau)} + \frac{(1 - \tau)(1 + 5\tau)}{16(1 + \tau)} \right) \frac{\Lambda_F^2}{\rho} \\
+ \frac{\tau(1 - \tau)}{(1 + \tau)^2} \left( 1 - 3\tau + \frac{2\tau}{1+\tau} (1 - \tau) \Lambda_N \right) \frac{\Lambda_N}{\rho}
$$

166
To interpret the result, it is interesting to notice that under the NFR condition (3.17),

\[ W^{u,NFR_{cond}}(NFR) = e^u_C + \frac{1}{2} (2e_1 - L) - \frac{(1 - \tau)(5\tau + 1)}{8(1 + \tau)} \left( (\sigma_H - \sigma_L) L \Lambda_F - \frac{\Lambda^2_F}{2\rho} \right) \]

which can be rewritten as

\[ W^{u,NFR_{cond}}(NFR) = e^u_C + \frac{1}{2} (2e_1 - L) - \frac{(1 - \tau)(5\tau + 1)}{8(1 + \tau)} \left( m\rho + v\rho_f + \alpha\rho_f \right) (\sigma_H - \sigma_L) L \Lambda_F + \frac{\Lambda^2_F}{2\rho} \]

\[ - \frac{\alpha}{v + \alpha} \left( 1 + \tau \right) (3\tau - 1)^2 32\rho\tau (1 - \tau) \]

The first row corresponds to the hedging demand and it is the same as for the active investor.

The second row corresponds to the impact of adverse selection; we have the same element in the welfare of passive investors.

**Passive investors**

**Case 1:** \( \tilde{s} = s_1, \tilde{\sigma} = \sigma_H, q^p_1 = \theta, q^p_2 = 1 - \theta \) (with probability \( \frac{1 + \tau}{8} \))

In this case, uninformed agents infer from the order flow that the active investors have observed a signal \( \tilde{s} = s_1 \). Thus \( q^p_1 = q^a_1 = \theta \) and \( q^p_2 = q^a_2 = 1 - \theta \). Imputing \( q^p_1 = \theta \) into (3.61), we get:

\[ U_2 (\tilde{w}^j(\beta^j)|p_2, l^j, \text{case 1}) = c^j_C + \theta (e^j_X - l^j_1) + (1 - \theta) (e^j_Y - l^j_2) - \frac{\rho}{2} \theta (1 - \theta) \left( (e^j_X - l^j_1) - (e^j_Y - l^j_2) \right)^2. \]

**Case 2:** \( \tilde{s} = s_2, \tilde{\sigma} = \sigma_L, q^p_1 = 1 - \theta, q^p_2 = \theta \) (with probability \( \frac{1 + \tau}{8} \))

This is the symmetric case of Case 1. In this case, uninformed agents infer from the order flow that the active investors have observed a signal \( \tilde{s} = s_2 \). Thus \( q^p_1 = q^a_1 = 1 - \theta \) and \( q^p_2 = q^a_2 = \theta \). Imputing \( q^p_1 = 1 - \theta \) into (3.61), we get:

\[ U_2 (\tilde{w}^j(\beta^j)|p_2, l^j, \text{case 2}) = c^j_C + (1 - \theta) (e^j_X - l^j_1) + \theta (e^j_Y - l^j_2) - \frac{\rho}{2} \theta (1 - \theta) \left( (e^j_X - l^j_1) - (e^j_Y - l^j_2) \right)^2. \]
Case 3: \( q_1^p = q_2^p = \chi \) (with probability \( \frac{1-\tau}{2} \))

In this case, uninformed agents infer from the order flow that the active investors have observed a signal \( \tilde{s} = s_3 \), thus \( q_1^p = q_2^p = \chi \). When \( \eta \rightarrow 0 \), inputing \( q_1^p = \frac{1}{2} \) into (3.61), we get:

\[
U_2 \left( \tilde{\omega}(\beta) | p_2, l^j, \text{case 3} \right) = e_C^j + \frac{1}{2} (e_X^j - l_1^j) + \frac{1}{2} (e_Y^j - l_2^j) - \frac{\rho}{\delta} \left( (e_X^j - l_1^j) - (e_Y^j - l_2^j) \right)^2.
\]

Case 4: \( \tilde{s} = s_1, \tilde{\sigma} = \sigma_L, q_1^p = \theta, q_2^p = 1 - \theta \) or \( \tilde{s} = s_2, \tilde{\sigma} = \sigma_H, q_1^p = 1 - \theta, q_2^p = \theta \), and \( q_1^p = q_2^p = \kappa \) (with probability \( \frac{1+\tau}{2} \))

In this case, when the condition (3.17) holds, uninformed traders cannot perfectly infer from the order flow whether \( \tilde{s} = s_1 \) and \( \tilde{\sigma} = \sigma_L \), or \( \tilde{s} = s_2 \) and \( \tilde{\sigma} = \sigma_H \). Indeed, for a given realization of \( \tilde{\omega} \), the same order flow is observed in both cases. Accordingly, the uninformed traders’ beliefs are such that: \( q_1^p = q_2^p = \kappa \). Inputing \( q_1^p = \frac{1}{2} \) into (3.61), we get:

\[
U_2 \left( \tilde{\omega}(\beta) | p_2, l^j, \text{case 4} \right) = e_C^j + \frac{1}{2} (e_X^j - l_1^j) + \frac{1}{2} (e_Y^j - l_2^j) - \frac{\rho}{\delta} \left( (e_X^j - l_1^j) - (e_Y^j - l_2^j) \right)^2.
\]

We now compute the expectation at date 1 of the passive investor’s expected utilities at date 2. We take into account the probability with which the investor receives a liquidity shock in state \( \omega_1 \) or \( \omega_2 \) depending on the realization of \( \tilde{\sigma} \).

Notice that the contribution of the expected utilities in Cases 1, 2 or 3 to the investor’s expected mean variance utility does not depend on whether the parameters are such that the equilibrium is fully revealing or not. Only the contribution of the expected utilities in Case 4 will thus change.

The realization of \( \tilde{\sigma} \) influences the probability with which the investor will receive a shock in state \( \omega_1 \) or \( \omega_2 \). In cases 1, 3a, and 4b, we have \( \tilde{\sigma} = \sigma_H \) while in cases 2, 3b and 4a we have \( \tilde{\sigma} = \sigma_L \).
This yields:

$$W^j(NFR) = \frac{1 + \tau}{8} \left( \sigma_H U_2 \left( \tilde{w}^j|p_2, l_1^j = L, s_1, \text{case 1} \right) + (1 - \sigma_H) U_2 \left( \tilde{w}^j|p_2, l_2^j = L, s_1, \text{case 1} \right) \right)$$

$$+ \frac{1 + \tau}{8} \left( \sigma_L U_2 \left( \tilde{w}^j|p_2, l_1^j = L, s_2, \text{case 2} \right) + (1 - \sigma_L) U_2 \left( \tilde{w}^j|p_2, l_2^j = L, s_2, \text{case 2} \right) \right)$$

$$+ \frac{1 - \tau}{4} \left( \sigma_H U_2 \left( \tilde{w}^j|p_2, l_1^j = L, s_3, \text{case 3a} \right) + (1 - \sigma_H) U_2 \left( \tilde{w}^j|p_2, l_2^j = L, s_3, \text{case 3a} \right) \right)$$

$$+ \frac{1 - \tau}{4} \left( \sigma_L U_2 \left( \tilde{w}^j|p_2, l_1^j = L, s_3, \text{case 3b} \right) + (1 - \sigma_L) U_2 \left( \tilde{w}^j|p_2, l_2^j = L, s_3, \text{case 3b} \right) \right)$$

$$+ \frac{1 + \tau}{8} \left( \sigma_L U_2 \left( \tilde{w}^j|p_2, l_1^j = L, s_1, \text{NFR, case 4a} \right) + (1 - \sigma_L) U_2 \left( \tilde{w}^j|p_2, l_2^j = L, s_1, \text{NFR, case 4a} \right) \right)$$

$$+ \frac{1 + \tau}{8} \left( \sigma_H U_2 \left( \tilde{w}^j|p_2, l_1^j = L, s_2, \text{NFR, case 4b} \right) + (1 - \sigma_H) U_2 \left( \tilde{w}^j|p_2, l_2^j = L, s_1, \text{NFR, case 4b} \right) \right)$$

When $\eta \to 0, \theta \to \frac{2\tau}{1+\tau}$, $\theta(1 - \theta) = \frac{2\tau(1 - \tau)}{(1+\tau)^2}$, which yields:

$$W^j(NFR) = e_C \frac{1}{2} (e_X + e_Y - L) - \frac{1 + \tau}{4} (\theta - \frac{1}{2})(\sigma_H - \sigma_L)L$$

$$- \rho \left( (e_X - e_Y)^2 + L^2 \right) \left( \frac{1 + \tau}{8} \theta(1 - \theta) + \frac{3 - \tau}{32} \right)$$

Since $\theta \to \frac{2\tau}{1+\tau}$ as $\eta \to 0$, we have $\theta(1 - \theta) = \frac{2\tau(1 - \tau)}{(1+\tau)^2}$, which yields:

$$W^j(NFR) = e_C \frac{1}{2} (e_X + e_Y - L) - \frac{3\tau - 1}{8} (\sigma_H - \sigma_L)L - \rho \frac{3 + 10\tau - 9\tau^2}{32(1 + \tau)} \left( (e_X - e_Y)^2 + L^2 \right)$$

To interpret the result, it is interesting to notice that under the NFR condition (3.17), $(\sigma_H - \sigma_L)L = \frac{(3\tau - 1)(1 + \tau)}{4\rho(1 - \tau)} \frac{\alpha}{(\alpha + v)}$ thus

$$W^{j,NFRecond}(NFR) = e_C \frac{1}{2} (e_X + e_Y - L) - \frac{(3\tau - 1)^2(1 + \tau)}{32\rho(1 - \tau)} \frac{\alpha}{(\alpha + v)}$$

$$- \rho \frac{3 + 10\tau - 9\tau^2}{32(1 + \tau)} \left( (e_X - e_Y)^2 + L^2 \right)$$

**Financial intermediaries**

The expected utilities of a financial intermediary at date 2 can be computed from that of an uninformed unconstrained investor, by replacing $\rho$ and $l^*_1 = l^*_2 = 0$ in each case. When
\( \eta \to 0 \), we get:

\[
U_2(\tilde{w}^f|p_2, \text{case } 1) = c_1^f + \theta(1-\theta) \left( \frac{\Lambda_N^2}{2\rho_f} \right) + \theta(e_X^f) + (1-\theta)(e_Y^f) + \theta(1-\theta)(e_X^f - e_Y^f)\Lambda_F
\]

\[
U_2(\tilde{w}^f|p_2, \text{case } 2) = c_1^f + \theta(1-\theta) \left( \frac{\Lambda_N^2}{2\rho_f} \right) + (1-\theta)(e_X^f) + \theta(e_Y^f) - \theta(1-\theta)(e_X^f - e_Y^f)\Lambda_F
\]

\[
U_2(\tilde{w}^f|p_2, \text{case } 3a) = c_1^f + \frac{1}{4} \left( \frac{\Lambda_N^2}{\rho_f} \right) + \frac{1}{2}(e_X^f) + \frac{1}{2}(e_Y^f) + \frac{1}{4}(e_X^f - e_Y^f)\Lambda_F
\]

\[
U_2(\tilde{w}^f|p_2, \text{case } 3b) = c_1^f + \frac{1}{4} \left( \frac{\Lambda_N^2}{\rho_f} \right) + \frac{1}{2}(e_X^f) - \frac{1}{2}(e_Y^f) + \frac{1}{4}(e_X^f - e_Y^f)\Lambda_F
\]

\[
U_2(\tilde{w}^f|p_2, \text{case } 4a) = c_1^f + \theta(1-(1-\theta)\Lambda_N) e_X^f + (1-\theta)(1+\theta\Lambda_N) e_Y^f
\]

\[
+ \frac{1}{4\rho_f} - \frac{2\theta(1-\theta)(1-(1-\theta)\Lambda_N)(1+\theta\Lambda_N)}{\rho_f}
\]

\[
U_2(\tilde{w}^f|p_2, \text{case } 4b) = c_1^f + (1-\theta)(1+\theta\Lambda_N) e_X^f + \theta(1-(1-\theta)\Lambda_N) e_Y^f
\]

\[
+ \frac{1}{4\rho_f} - \frac{2(1-\theta)(1+\theta\Lambda_N)(1-(1-\theta)\Lambda_N)}{\rho_f}
\]

We now compute the expectation at date 1 of the financial intermediary's utilities at date 2.

Computation yields:

\[
W^f(NFR) = c_1^f + \frac{1}{2} \left( e_X^f + e_Y^f \right) + \frac{(3\tau - 1)^2}{8\rho_f(1+\tau)} + \frac{(1-\tau)(1+5\tau)}{16(1+\tau)} \frac{\Lambda_N^2}{\rho_f}
\]

\[
+ \frac{\tau(1-\tau)}{(1+\tau)^2} \left( 1 - 3\tau \right) + \frac{2\tau}{1+\tau} (1-\tau)\Lambda_N \ \frac{\Lambda_N}{\rho_f}
\]

To interpret the result, it is interesting to notice that under the NFR condition (3.17), \( \Lambda_N = \frac{(3\tau-1)(1+\tau)}{4\tau(1-\tau)} \) which yields:

\[
W^{f,NFRcond}(NFR) = c_1^f + \frac{1}{2} \left( e_X^f + e_Y^f \right) + \frac{(1-\tau)(1+5\tau)}{16\rho_f(1+\tau)} \frac{\Lambda_F^2}{\rho_f}
\]
Corollary 3.1

It follows from Proposition 3.2 that

\[ \Pi^I \equiv W^i(NFR) - W^u(NFR) = -\frac{\tau(1 - \tau)}{2(1 + \tau)} \left( \Lambda_F + \Lambda_N \right) + \frac{1 + \tau}{4\tau} \Lambda_F \left( \sigma_H - \sigma_L \right) L \]

\[ + \frac{\tau(1 - \tau)}{4(1 + \tau)} \left( \frac{\Lambda_N^2}{\rho} + \frac{1 + 5\tau \Lambda_N^2}{4\tau} \right) \left( \frac{\tau(1 - \tau)}{2(1 + \tau)} \left( \Lambda_F + \Lambda_N \right) + \frac{1 + \tau}{4\tau} \Lambda_F \right) \left( \sigma_H - \sigma_L \right) L \]

\[ - \frac{\tau(1 - \tau)}{4(1 + \tau)} \left( \frac{8\tau(1 - \tau) \Lambda_N^2}{\rho} + \frac{(1 + 5\tau) \Lambda_F^2}{4\tau} \right) \]

\[ - \frac{(3\tau - 1)^2}{8\rho(1 + \tau)} - \frac{\tau(1 - \tau)}{(1 + \tau)^2} \frac{\Lambda_N}{\rho} \]

Computations yield

\[ \Pi^I = \left( \frac{3\tau - 1}{4\rho(1 + \tau)} \right)^2 \left( -\frac{1}{2} + \frac{4\tau(1 - \tau)}{(3\tau - 1)(1 + \tau)} \Lambda_N + \frac{\tau(1 - \tau)}{(1 + \tau)^2} \Lambda_N^2 \right) \]

(3.72)

Using the NFR condition (3.17), simple computations yield

\[ \Pi^{I,NFRcond} = \frac{(1 + \tau)(3\tau - 1)^2}{64\rho\tau(1 - \tau)}, \]

(3.73)

which is strictly positive. This proves the first part of the corollary.

Without using the NFR condition before taking the derivative, we have

\[ \frac{\partial \Pi^I}{\partial \alpha} = \frac{(3\tau - 1)\tau(1 - \tau)}{\rho(1 + \tau)^2} \frac{\partial \Lambda_N}{\partial \alpha} + \frac{(3\tau - 1)^2\tau(1 - \tau)}{2\rho(1 + \tau)^3} \Lambda_N \frac{\partial \Lambda_N}{\partial \alpha} \]

\[ = \frac{(3\tau - 1)\tau(1 - \tau)}{\rho(1 + \tau)^2} \frac{\partial \Lambda_N}{\partial \alpha} \left( 1 + \frac{(3\tau - 1)}{2(1 + \tau)} \Lambda_N \right) \]

Recall that under NFR

\[ \Lambda^{NFRcond}_N = \frac{(3\tau - 1)(1 + \tau)}{4\tau(1 - \tau)}. \]
We simply replace terms associated with $\Lambda_N$ with the NFR.

$$\frac{\partial \Pi^I}{\partial \alpha} = \frac{(3\tau - 1)\tau(1-\tau)}{\rho(1+\tau)^2} \frac{\partial \Lambda_N}{\partial \alpha} \left( 1 + \frac{(3\tau - 1)(3\tau - 1)(1+\tau)}{2(1+\tau)4\tau(1-\tau)} \right)$$

$$= \frac{(3\tau - 1)\tau(1-\tau)}{\rho(1+\tau)^2} \frac{\partial \Lambda_N}{\partial \alpha} \left( 1 + \frac{(3\tau - 1)^2}{8\tau(1-\tau)} \right)$$

Thus $\frac{\partial \Pi^I}{\partial \alpha}$ has the sign of $\frac{\partial \Lambda_N}{\partial \alpha}$. When we compute $\frac{\partial \Lambda_N}{\partial \alpha}$ and use the NFR condition, $\frac{\partial \Lambda_N}{\partial \alpha}^{NFRcond} \leq 0$, thus $\Pi^I$ is decreasing in $\alpha$: the value of information decreases with the proportion of informed investors. This shows the second part of the corollary.

Similarly,

$$\frac{\partial \Pi^I}{\partial \upsilon} = \frac{(3\tau - 1)\tau(1-\tau)}{\rho(1+\tau)^2} \frac{\partial \Lambda_N}{\partial \alpha} + \frac{(3\tau - 1)^2\tau(1-\tau)}{2\rho(1+\tau)^3} \Lambda_N \frac{\partial \Lambda_N}{\partial \upsilon}$$

$$= \frac{(3\tau - 1)(1-\tau)}{\rho(1+\tau)^2} \frac{\partial \Lambda_N}{\partial \upsilon} \left( 1 + \frac{(3\tau - 1)^2}{2(1+\tau)\Lambda_N} \right)$$

Recall that under NFR

$$\Lambda_N^{NFRcond} = \frac{(3\tau - 1)(1+\tau)}{4\tau(1-\tau)}.$$  

We simply replace terms associated with $\Lambda_N$ with the NFR.

$$\frac{\partial \Pi^I}{\partial \upsilon} = \frac{(3\tau - 1)\tau(1-\tau)}{\rho(1+\tau)^2} \frac{\partial \Lambda_N}{\partial \alpha} \left( 1 + \frac{(3\tau - 1)(3\tau - 1)(1+\tau)}{2(1+\tau)4\tau(1-\tau)} \right)$$

$$= \frac{(3\tau - 1)\tau(1-\tau)}{\rho(1+\tau)^2} \frac{\partial \Lambda_N}{\partial \alpha} \left( 1 + \frac{(3\tau - 1)^2}{8\tau(1-\tau)} \right)$$

Thus $\frac{\partial \Pi^I}{\partial \upsilon}$ has the sign of $\frac{\partial \Lambda_N}{\partial \upsilon}$. When we compute $\frac{\partial \Lambda_N}{\partial \upsilon}$ and use the NFR condition, $\frac{\partial \Lambda_N}{\partial \upsilon}^{NFRcond} \geq 0$, thus $\Pi^I$ is increasing in $\upsilon$: the value of information increases with the proportion of uninformed investors.
Corollary 3.2

It follows from Proposition 3.2 that

\[ W^u(NFR) = e_1 + e^u_C - \frac{L}{2} + \left( (\sigma_H - \sigma_L) \left( -\frac{1}{8} \Lambda_F - \frac{\tau(1-\tau)}{2(1+\tau)} (\Lambda_F + \Lambda_N) \right) \right) L \]
\[ + \frac{(3\tau - 1)^2}{8\rho(1+\tau)} + \frac{(1-\tau)(1+5\tau)\Lambda_F^2}{16(1+\tau)} \]
\[ + \frac{\tau(1-\tau)}{(1+\tau)^2} \left( (1 - 3\tau) + \frac{2\tau}{1+\tau} (1-\tau)\Lambda_N \right) \frac{\Lambda_N}{\rho} \]

\[ W^j(NFR) = e^j_C + \frac{1}{2} (e^j_X + e^j_Y - L) - \frac{3\tau - 1}{8} (\sigma_H - \sigma_L) L \]
\[ - \rho \frac{3 + 10\tau - 9\tau^2}{32(1+\tau)} ((e^j_X - e^j_Y)^2 + L^2) \]

Thus

\[ \Pi^T = \frac{(1-\tau)(1+5\tau)}{8(1+\tau)} \Lambda_F (\sigma_H - \sigma_L) L + \frac{(1-\tau)(1+5\tau)\Lambda_F^2}{8(1+\tau)} \frac{1}{2\rho} \]
\[ - \frac{4\tau(1-\tau)}{8(1+\tau)} \Lambda_N (\sigma_H - \sigma_L) L + \frac{4\tau(1-\tau)}{(1+\tau)^2} \left( \frac{\tau(1-\tau)}{(1+\tau)\Lambda_N} - \frac{3\tau - 1}{2(1+\tau)} \right) \frac{\Lambda_N}{2\rho} \]
\[ + \frac{(3\tau - 1)^2}{8\rho(1+\tau)} + \frac{3\tau - 1}{8} (\sigma_H - \sigma_L) L + \rho \frac{3 + 10\tau - 9\tau^2}{32(1+\tau)} L^2 \]

Using the NFR condition (3.17), simple computations yield

\[ \Pi^{T,NFR}_{cond} = \left( \frac{3 + 10\tau - 9\tau^2}{32(1+\tau)} - \frac{(1-\tau)(5\tau + 1)}{16(1+\tau)} \right) \left( 1 - \frac{(m\rho)^2}{(m\rho + (\alpha + \nu)\rho f)^2} \right) (\sigma_H - \sigma_L)^2 \rho L^2 \]

(3.74)

which is positive since \( \frac{3 + 10\tau - 9\tau^2}{32(1+\tau)} - \frac{(1-\tau)(5\tau + 1)}{16(1+\tau)} > 0 \). This proves the first part of the corollary.

Second, let us compute the derivative \( \frac{\partial \Pi^T}{\partial \alpha} \) before plugging the NFR condition.

173
\[
\frac{\partial \Pi^T}{\partial \alpha} = \frac{(1 - \tau)(1 + 5\tau)}{8(1 + \tau)} \left[ (\sigma_H - \sigma_L)L \frac{\partial \Lambda_F}{\partial \alpha} + \frac{\Lambda_F}{\rho} \frac{\partial \Lambda_F}{\partial \alpha} \right] \\
- \frac{4\tau(1 - \tau)}{8(1 + \tau)} (\sigma_H - \sigma_L) \frac{\partial \Lambda_N}{\partial \alpha} + \frac{4\tau^2(1 - \tau)^2 \Lambda_N}{(1 + \tau)^3} \frac{\partial \Lambda_N}{\partial \alpha} \\
- \frac{(3\tau - 1)\tau(1 - \tau)}{\rho(1 + \tau)^2} \frac{\partial \Lambda_N}{\partial \alpha} \\
= \frac{(1 - \tau)(1 + 5\tau)}{8(1 + \tau)} \left[ (\sigma_H - \sigma_L)L + \frac{\Lambda_F}{\rho} \right] \frac{\partial \Lambda_F}{\partial \alpha} \\
+ \frac{\tau(1 - \tau)}{1 + \tau} \left( \frac{1}{2} (\sigma_H - \sigma_L)L + \frac{4\tau(1 - \tau)\Lambda_N}{(1 + \tau)^2} \rho - \frac{(3\tau - 1)}{\rho(1 + \tau)} \right) \frac{\partial \Lambda_N}{\partial \alpha}
\]

Recall that under NFR
\[
\Lambda_N^{NFRcond} = \frac{(3\tau - 1)(1 + \tau)}{4\tau(1 - \tau)}.
\]

We simply replace terms associated with \( \Lambda_N \) with the NFR.

\[
\frac{\partial \Pi^T}{\partial \alpha} = \frac{(1 - \tau)(1 + 5\tau)}{8(1 + \tau)} \left[ (\sigma_H - \sigma_L)L + \frac{\Lambda_F}{\rho} \right] \frac{\partial \Lambda_F}{\partial \alpha} - \frac{\tau(1 - \tau)}{1 + \tau} \frac{1}{2} (\sigma_H - \sigma_L)L \frac{\partial \Lambda_N}{\partial \alpha}
\]

Since \( \frac{\partial \Lambda_F}{\partial \alpha} > 0 \) and \( \frac{\partial \Lambda_N}{\partial \alpha} \leq 0 \), we have \( \frac{\partial \Pi^T}{\partial \alpha} \geq 0 \).

**Proposition 3.3**

Proposition 3.3 follows from the results detailed below.

Under the NFR condition (3.17), the value of being informed, conditional on the investor being unconstrained and on the presence of active investors, does not depend on \( \alpha \) nor \( \nu \), and it increases in \( \tau \). Let us define \( c^*_I \) as

\[
c^*_I \equiv \Pi^{I,NFRcond} = \frac{(1 + \tau)(3\tau - 1)^2}{64\tau(1 - \tau)}.
\]  

(3.75)

Besides, we have

\[
\Pi^{T,NFRcond} = \rho \frac{3 + 10\tau - 9\tau^2}{32(1 + \tau)} L^2 - \rho \frac{(1 - \tau)(5\tau + 1)}{16(1 + \tau)} \left( 1 - \frac{(m\rho)^2}{(m\rho + (\alpha + \nu)\rho_f)^2} \right) (\sigma_H - \sigma_L)^2 L^2.
\]
Under the NFR condition (3.17), the value of flexibility reaches a maximum when \( \alpha + \nu \to 0 \), and this maximum is \( \frac{3 + 10\tau - 9\tau^2}{32(1 + \tau)} \rho L^2 \). Let us define

\[
c^*_T \equiv \lim_{\alpha + \nu \to 0} \Pi^{T,NFRcond} = \frac{3 + 10\tau - 9\tau^2}{32(1 + \tau)} \rho L^2.
\] (3.76)

Part 1. Non existence of NFR equilibrium

First, it follows from the definition of \( c^*_I \) that under the NFR condition (3.17)

\[
W^{i,NFRcond}(\text{NFR}) - W^{u,NFRcond}(\text{NFR}) < 0 \iff c_I > c^*_I.
\]

Thus if \( c_I > c^*_I \), and if a NFR equilibrium exists, it is characterized by \( \alpha^* = 0 \) since active investors strictly prefer being uninformed than informed. But if \( \alpha^* = 0 \) there cannot be any NFR equilibrium, which yields a contradiction.

Second, we also have

\[
W^{i,NFRcond}(\text{NFR}) - W^{j,NFRcond}(\text{NFR}) = \Pi^{I,NFRcond} + \Pi^{T,NFRcond} - c^*_I - c^*_T
\]

\[
= -\rho \left( \frac{(1 - \tau)(1 + 5\tau)}{16(1 + \tau)} \right) \left( 1 - \frac{(m\rho^2)^2}{(mp + (\alpha + v)\rho f)^2} \right) (\sigma_H - \sigma_L)^2 L^2
\]

\[
\leq 0, \forall \alpha, \nu,
\]

\[
+ \frac{(3\tau - 1)^2(1 + \tau)}{64\rho \tau (1 - \tau)} + \frac{3 + 10\tau - 9\tau^2}{32(1 + \tau)} \rho L^2
\]

Hence if \( c_I + c_T > c^*_I + c^*_T \), and if a NFR equilibrium exists, it is characterized by \( \alpha^* = 0 \) since investors strictly prefer being passive than active and informed. But if \( \alpha^* = 0 \) there cannot be any NFR equilibrium, which yields a contradiction.

Consequently, if \( c_I > c^*_I \) or if \( c_I + c_T > c^*_I + c^*_T \), there exists no NFR equilibrium. A necessary condition for a NFR equilibrium to exist is \( c_I \leq c^*_I \) and \( c_I + c_T \leq c^*_I + c^*_T \).
Part 2. Case $c_I < c_I^*$ and $c_I + c_T \leq c_T^* + c_T^*$: $v^* = 0$ at equilibrium

Since $c_I < c_I^*$, active investors have incentives to acquire information. Staying uninformed is strictly dominated by becoming informed for active investors, i.e., $v^* = 0$. If a NFR equilibrium exists, then it would be such that investors would either be active and informed, or passive. We therefore focus on $W_i^{NFRcond}(NFR) - W_j^{NFRcond}(NFR)$, which we have defined as $\Pi^{I,NFRcond} + \Pi^{T,NFRcond}$, evaluated at $v^* = 0$, that is

\[
(W_i^{NFRcond}(NFR) - W_j^{NFRcond}(NFR)) \big|_{v^*=0} - (c_I - c_T)
\]

\[
= -\rho \frac{(1 - \tau)(1 + 5\tau)}{16(1 + \tau)} \left(1 - \frac{(m\rho)^2}{(m\rho + \alpha\rho_I)^2}\right) (\sigma_H - \sigma_L)^2 L^2
\]

\[
+ \frac{(3\tau - 1)^2(1 + \tau)}{64\rho\tau(1 - \tau)} + \frac{3 + 10\tau - 9\tau^2}{32(1 + \tau)} \rho L^2 - c_I - c_T
\]

\[
\leq 0, \forall \alpha
\]

\[
\geq 0, \forall \alpha
\]

First, notice that if $c_I + c_T = c_I^* + c_T^*$ then the second part disappears and

\[
(W_i^{NFRcond}(NFR) - W_j^{NFRcond}(NFR)) \big|_{v^*=0} - (c_I - c_T) \leq 0
\]

with an equality when $\alpha = 0$. If $\alpha > 0$ then passive investors are strictly better off which implies that $\alpha > 0$ cannot be an equilibrium. If $\alpha = 0$ then investors are indifferent between being passive or informed provided that the equilibrium is NFR, which contradicts $\alpha = 0$. Consequently, there exists no NFR equilibrium if $c_I + c_T = c_I^* + c_T^*$.

Second, when $c_I + c_T < c_I^* + c_T^*$, $W_i^{NFRcond}(NFR) - W_j^{NFRcond}(NFR) 0|_{v^*=0} - (c_I - c_T)$ is decreasing in $\alpha$, and is strictly positive in $\alpha = 0$. Consequently, if

\[
(W_i^{NFRcond}(NFR) - W_j^{NFRcond}(NFR)) \big|_{v^*=0} - (c_I - c_T) \big|_{\alpha=1} > 0,
\]
then \( (W^{i,NFR_{cond}}(NFR) - W^{j,NFR_{cond}}(NFR)) |_{\nu^*=0} - c_I - c_T > 0 \) for all \( \alpha \). Let us define

\[
c_{I,T}^{**} = -\rho \frac{(1 - \tau)(1 + 5\tau)}{16(1 + \tau)} \left( 1 - \frac{(m\rho)^2}{(m\rho + \rho_f)^2} \right) (\sigma_H - \sigma_L)^2 L^2 + \frac{(3\tau - 1)^2(1 + \tau)}{64\rho(1 - \tau)} \frac{1}{(1 + \tau)^2} \rho L^2.
\]

Given our definition, \( c_{I,T}^{**} = c_I^* + c_T^* - c_{I,T}^* \).

If \( c_I + c_T < c_{I,T}^{**} \) and provided that the NFR condition holds when \( \nu^* = 0 \), there exists a NFR equilibrium in pure strategies in which \( \alpha^* = 1 \) since \( (W^{i,NFR_{cond}}(NFR) - W^{j,NFR_{cond}}(NFR)) |_{\nu^*=0} - c_I - c_T > 0 \) for all \( \alpha \).

If \( c_I + c_T \geq c_{I,T}^{**} \) and provided that the NFR condition holds when \( \nu^* = 0 \), there exists a NFR equilibrium in mixed strategies in which \( \alpha^* \leq 1 \) (with \( \alpha^* = 1 \) when \( c_I + c_T = c_{I,T}^{**} \)). Formally, \( \alpha^* \) is such that \( (W^{i,NFR_{cond}}(NFR) - W^{j,NFR_{cond}}(NFR)) |_{\nu^*=0} - c_I - c_T = 0 \), that is:

\[
(m\rho + \alpha^* \rho_f)^2 = \frac{(m\rho)^2}{(\sigma_H - \sigma_L)^2 L^2 \rho(1 - \tau)(1 + 5\tau) + 1}
\]

which yields

\[
\alpha^* = \frac{m\rho}{\rho_f} \frac{1 - \sqrt{\frac{c_I + c_T - c_I^* - c_T^*}{(\sigma_H - \sigma_L)^2 L^2 \rho(1 - \tau)(1 + 5\tau) + 1}}}{\sqrt{\frac{c_I + c_T - c_I^* - c_T^*}{(\sigma_H - \sigma_L)^2 L^2 \rho(1 - \tau)(1 + 5\tau) + 1}}}
\]

Notice that when \( c_I + c_T = c_{I,T}^{**} \), we have

\[
\alpha^* = 1.
\]

**Part 3.** \( c_I = c_I^* \) (and \( c_T \leq c_T^* \))

First, notice that since \( c_T \leq c_T^* \), becoming active and uninformed is not strictly dominated by becoming passive: there may exist NFR equilibrium characterized by \( \nu^* > 0 \).

Now, since \( c_I = c_I^* \), active investors are indifferent between informed or not. Under this
restriction, we thus have

\[ W^{i,NFR}_{i,NFR}(NFR) - c_I - c_T - W^{j,NFR}_{j,NFR}(NFR) = W^{u,NFR}_{u,NFR}(NFR) - c_T - W^{j,NFR}_{j,NFR}(NFR) = \]

\[ -\beta \frac{(1-\tau)(1+5\tau)}{16(1+\tau)} \left( 1 - \frac{(m\rho)^2}{(m\rho + (\alpha + \nu)\rho f)^2} \right) (\sigma_H - \sigma_L)^2 L^2 \]

\[ \leq 0, \forall \alpha, \nu \]

\[ + c^*_T - c_T \]

\[ \geq 0 \]

\[ W^{i,NFR}_{i,NFR}(NFR) - c_I - c_T - W^{j,NFR}_{j,NFR}(NFR) \] is decreasing in \( \alpha \) and in \( \nu \), and is positive at \( \alpha + \nu = 0 \) (strictly if \( c_T < c^*_T \)).

Thus when \( (W^{i,NFR}_{i,NFR}(NFR) - W^{j,NFR}_{j,NFR}(NFR) - c_I - c_T) |_{\alpha+\nu=1} > 0 \), we have that \( W^{i,NFR}_{i,NFR}(NFR) - c_I - c_T - W^{j,NFR}_{j,NFR}(NFR) > 0 \) for all \( \alpha \) and \( \nu \). Let us define

\[ c^{***}_T = c^*_T - \beta \frac{(1-\tau)(1+5\tau)}{16(1+\tau)} \left( 1 - \frac{(m\rho)^2}{(m\rho + (\alpha + \nu)\rho f)^2} \right) (\sigma_H - \sigma_L)^2 L^2. \]

Given our definitions, we have \( c^{***}_T = c^*_T - c^*_I |_{T}. \)

If \( c_T < c^{***}_T \) and provided that the NFR condition holds when \( \alpha^* + \nu^* = 1 \), there exists a NFR equilibrium in which \( \pi^* = 0 \) and \( \alpha^* + \nu^* = 1 \). To identify the values of \( \alpha^* \) and \( \nu^* \), we use the equilibrium NFR condition, which is

\[ \text{NFR condition: } \frac{(3\tau - 1)(1+\tau)}{4\rho \tau (1-\tau)} \alpha = (\alpha + \nu)(\sigma_H - \sigma_L)L, \]

which yields

\[ \alpha^* = \frac{4\rho \tau (1-\tau)}{(3\tau - 1)(1+\tau)}(\sigma_H - \sigma_L)L \]

If \( c_T \geq c^{***}_T \) and provided that the NFR condition holds, there exists a NFR equilibrium in mixed strategies in which becoming a passive investor is not strictly dominated. To identify the
values of $\alpha^*$ and $\nu^*$, we use the two equilibrium conditions, which are

Indifference between passive and active: $\Pi^{T,NFR} = c_T$

NFR condition:

$$\frac{(3\tau - 1)(1 + \tau)}{4\rho\tau(1 - \tau)} \alpha = (\alpha + \nu)(\sigma_H - \sigma_L)L$$

This yields

$$0 = -\rho \frac{(1 - \tau)(1 + 5\tau)}{16(1 + \tau)} \left( 1 - \frac{(m\rho)^2}{(m\rho + (\alpha + \nu)\rho f)^2} \right) (\sigma_H - \sigma_L)^2 L^2 + c_T^* - c_T$$

$$\frac{(3\tau - 1)(1 + \tau)}{4\rho\tau(1 - \tau)} \alpha = (\alpha + \nu)(\sigma_H - \sigma_L)L$$

From the first equation of the system,

$$\alpha + \nu = \frac{m\rho}{\rho f} \left( \frac{(\sigma_H - \sigma_L) L \sqrt{\rho \frac{(1-\tau)(1+5\tau)}{16(1+\tau)}}}{\sqrt{c_T - c_T^* + \rho \frac{(1-\tau)(1+5\tau)}{16(1+\tau)}(\sigma_H - \sigma_L)^2 L^2}} - 1 \right) (\sigma_H - \sigma_L) L$$

Finally, given the NFR condition,

$$\alpha^* = \frac{4\rho\tau(1 - \tau)}{(3\tau - 1)(1 + \tau)} \frac{m\rho}{\rho f} \left( \frac{(\sigma_H - \sigma_L) L \sqrt{\rho \frac{(1-\tau)(1+5\tau)}{16(1+\tau)}}}{\sqrt{c_T - c_T^* + \rho \frac{(1-\tau)(1+5\tau)}{16(1+\tau)}(\sigma_H - \sigma_L)^2 L^2}} - 1 \right) (\sigma_H - \sigma_L) L$$

$$\nu^* = \frac{m\rho}{\rho f} \left( \frac{(\sigma_H - \sigma_L) L \sqrt{\rho \frac{(1-\tau)(1+5\tau)}{16(1+\tau)}}}{\sqrt{c_T - c_T^* + \rho \frac{(1-\tau)(1+5\tau)}{16(1+\tau)}(\sigma_H - \sigma_L)^2 L^2}} - 1 \right) \left( 1 - \frac{4\rho\tau(1 - \tau)}{(3\tau - 1)(1 + \tau)} (\sigma_H - \sigma_L)L \right)$$

Notice that when $c_T = c_T^*$, $\lim_{c_T \to c_T^*}(\pi^*) = 1$ so there can be no NFR equilibrium. Notice also that if $c_T = c_T^{***}$ and provided that the NFR condition holds when $\pi^* = 0$, we have $\pi^* = 0$. 

179
Bibliography


