The Market for ESG Ratings

Ehsan Azarmsa and Joel Shapiro*

ABSTRACT

We present a model of competition between ESG raters who acquire information about multiple unrelated categories and sell ratings. Raters specializing in different categories maximizes the amount of information transmitted and surplus, and can be the equilibrium outcome. When investors place a high value on ESG performance across multiple categories, the unique equilibrium is for the raters to generalize – splitting their effort among the categories, resulting in

*Ehsan Azarmsa is at the University of Illinois Chicago. Joel Shapiro (corresponding author, Joel.Shapiro@sbs.ox.ac.uk) is at Saïd Business School, University of Oxford. We are indebted to Editor Lenoid Kogan, an anonymous associate editor, and two anonymous referees. We thank Jasper Knyphausen for his exceptional research assistance. We are also grateful to Sonny Biswas, Noemie Bucourt, Jesse Davis, Deeksha Gupta, Marcin Kacperczyk, Raphael Levy, Lubo Litov, Andrey Malenko, Vincent Maurin, Stefano Pegoraro, Jan Starmans, Savitar Sundaresan, Ben Tsou, Kelsey Wei, Kai Hao Yang, Liyan Yang, Aaron Yoon, seminar participants at the University of Oklahoma, University of Illinois Chicago, University of Texas at Dallas, Cleveland Fed, Aalto University, NHH, University of Toulouse, and conference participants at Midwest Finance Association, SFS Cavalcade NA, LBS Summer Finance Symposium, GRASFI, Barcelona Summer Forum, Luxembourg Conference on Sustainable Finance, Financial Intermediation Research Society Meeting, Vienna Festival of Finance Theory, Northern Finance Association, HEC-HKUST Sustainable Finance Webinar, Dutch Sustainable Finance Workshop, MSU Federal Credit Union Conference, Conference on Sustainable Financial Intermediation, and London FTG meeting for helpful comments and discussions. The authors have read *The Journal of Finance* disclosure policy and have no conflicts of interest to disclose.

less informative ratings. Greenwashing by firms can make generalization the only equilibrium. We also demonstrate that specialization maximizes ratings disagreement and, thus, empirical measures of disagreement may be poor measures of surplus.

Global sustainable investment has reached $30.3 trillion (GSIA, 2022). Underlying this boom is the market to assess investments on environmental, social, and governance (ESG) criteria. A large number[1] of ESG rating providers have sprouted up to gather, analyze, and aggregate data for investors. Nevertheless, ESG ratings have come under intense scrutiny from the media, regulators, and academics for measurement and accuracy issues.[2]

Indeed, some studies use credit ratings as a benchmark for how inaccurate ESG ratings are. However, a comparison between ESG ratings and credit ratings is like comparing apples to oranges. Aside from having well established data and methodologies, credit ratings attempt to measure one variable: the probability a default will occur.[3] ESG ratings measure a panoply of subcategories within E, S, and G that are very different from each other. For example, one of the largest raters, MSCI, states that within S, it includes

[1]The European Securities and Markets Authority states that there are 59 ESG rating providers in Europe (ESMA, 2022). Still, the market is dominated by several larger providers.

[2]Some examples are "Lack of Uniformity in ESG Ratings System Poses Risks, Opportunities" by Petrina H. McDaniel, et al., *Bloomberg Law*, May 16 , 2022, "EU overhaul of ESG ratings industry rules has further to run" by Tommy Reggiori Wilkes and Simon Jessop, *Reuters*, June 16 , 2023, and "ESG ratings: whose interests do they serve" by Kenza Bryan, *Financial Times*, October 3, 2023.

[3]Note that Moody's states that they also measure the amount of recovery upon default.

the headings (with many subheadings): Human Capital, Product Liability, Stakeholder Opposition, and Social Opportunities, and within E, it includes the headings: Climate Change, Natural Capital, Pollution & Waste, and Environmental Opportunities.

Large investors understand that ESG ratings are not credit ratings. They purchase multiple ESG ratings and use the data from the categories rather than the aggregate score (SustainAbility, 2020; ESMA, 2022).

Given this, we analyze the market structure of the ESG ratings industry. In our model, there is a project that needs funding from an investor. In addition to caring about the financial performance, the investor cares about two categories of ESG performance. These may be two categories among E, S, and G, or two of their subcategories. This preference may represent a concern about externalities, long-term performance, or potential scandals. There are two ESG raters who choose how much effort to allocate to gathering information on the two categories. The two raters then compete over the fees they charge the investor for purchasing the gathered information.[4] We assume that the ESG raters (i) provide the rating by category rather than just an aggregate, and (ii) communicate their information truthfully.

We find that, in equilibrium, the investor will purchase ratings from both raters. When the investor's value for ESG performance is low, or when positive information about one category is sufficient for the investment, the investor will pay each rater the marginal value of their ratings (the maximum amount the rater could charge the investor for acquiring the second rating

[4]We note that unlike corporate credit ratings, which are paid for by issuers, a large majority of ESG ratings are bought by investors.

if the investor had already acquired a first rating). In this case, raters will *specialize* in different categories: each rater will focus all of its effort on one category.

However, the investor may not always be willing to pay both raters their marginal values. This is because, depending on the investor's ESG preferences and the raters' ratings design, the combined value of the ratings, the maximum value the investor is willing to pay for ratings from both raters, could be less than the sum of the marginal value of each rater. In this case, the raters' share from the combined value will depend on the stand-alone value of their ratings (the maximum amount they could charge the investor if they were the only rater).

When the investor cares a lot about ESG performance and needs a positive update about the project's performance in both categories to invest, information about a single category has no value. Here, the stand-alone value plays an important role – in this case, a rater who specializes in one category has no stand-alone value, which pushes the raters away from specialization. In fact, in equilibrium, the raters will *generalize*: each rater will split its effort between the rating categories. Generalization, of course, leads to a costly duplication of effort. Nevertheless, in the current marketplace, a substantial majority of ESG raters generalize (SustainAbility (2023), Table 2).

For intermediate values of the investor's value for ESG performance, there are multiple equilibria – both specialization and generalization are equilibria. This is because, in this range, the raters' decisions to specialize (and generalize) are strategic complements.

The main theoretical result of the paper is that market competition will lead to generalization when the investor cares strongly about both categories. This result firmly relies on two key forces. The first is a complementarity between the two categories for the investor. The second is that one of the raters has some market power. The market power reduces the surplus available to the other rater, making it depend on its stand-alone value. The stand-alone value of specialization is lower than that of generalization due to the complementarity, pushing the raters into generalizing.

We find that having the raters specialize in different categories maximizes value, where we define value as the sum of the agents' payoffs. This is because specialization maximizes the amount of information transmitted. When investors place a high value on ESG performance across multiple categories, a wedge arises between the market solution and the value-maximizing solution, as the market provides less information through generalization. This occurs exactly when investors value this information the most.

Many recent papers argue that divergence among ESG ratings make it difficult for the market to assess ESG performance. For example, Berg, Kölbel, and Rigobon (2022) document that the correlations between ESG ratings range from 0.38 to 0.71; similar results are found in Chatterji et al. (2016), Christensen, Serafeim, and Sikochi (2022), and Gibson Brandon, Krueger, and Schmidt (2021). We demonstrate that the correlation between the ratings of the two raters is minimized when the raters specialize in different categories. Given the result that specialization is value maximizing, this implies that divergence is not a good measure of welfare.[5]

[5]This is in line with a statement by Jean Christophe Nicaise Chateau from the Eu-

There is much discussion of greenwashing in ESG data provision: firms' manipulation of their perfomance data so as to improve public perception (for example, see Cornaggia and Cornaggia (2025)). We show that greenwashing will impact the strategy of ESG raters – when greenwashing by firms is substantial, the unique equilibrium will be generalization. This arises due to investors' rational suspicion of ratings due to greenwashing. Particularly, investors require a consistently high ESG performance across categories for their investment, which pushes the raters toward generalization. This implies that less information provision by firms makes ratings less informative (1) through a direct effect, and (2) by influencing the industrial organization of the ESG raters.

Our results are robust to alternative specifications of the model. Some of these specifications examine asymmetric investor preferences: different valuations for the categories and different preferences over raters. Others analyze the market power of raters in more detail: allowing for unequal costs of information acquisition across categories and raters, varying the allocation of market power in pricing, and looking at the timing of investing in rating technologies and setting fees. The remaining specifications broaden the economic

ropean Commission: "Users like the diversity of ratings. It gives them the ability to go to a number of different providers depending on the type of information they're looking for. The more they know about what they're buying, the more it will help them choose to go to the relevant data provider or ratings agency that may have a more specialized approach." (See "Rating the Raters Yet Again: Increasing ESG Scrutiny Makes Current Rate the Raters Study Even More Crucial" by Aiste Brackley, Emily K. Brock, and Justin Nelson, SustainAbility, September 21, 2022.)

insights of the model: allowing for different measurement methodologies, endogenizing the ESG performance of firms, and introducing economies of scale in information production.

We also explore policy implications of the model, looking at the benefits and costs of price caps, collusion, and standardization.

Berg et al. (2022) point to three sources of divergence in ESG ratings. The first is "scope divergence," where ratings are based on different attributes. Scope divergence resembles specialization in our model, which we find to be welfare enhancing. The second is "measurement divergence," where raters measure the same attribute differently. Differences in measurement could come from different protocols in measuring a subcategory or from using different underlying data. We capture the former in our model by our notion of "precision". We examine the latter in an extension of the model, and show that divergence here will be efficient. The last is "weight divergence," which captures raters weighting different attributes differently in the aggregate rating. Our model is not able to capture weight divergence, as we do not look at aggregate ratings. Berg et al. (2022) find little evidence of weight divergence; they state that it contributes to only 6% of the divergence.

There is some evidence of specialization in the current marketplace. The Carbon Disclosure Project (CDP) focuses on environmental ratings. Institutional Shareholder Services (ISS), which has its roots as a proxy advisor, has both ESG ratings and a stand-alone "Governance QualityScore," and has been "praised most frequently for its governance scoring" (SustainAbility, 2020, p. 14).

In our model, we are agnostic about why investors derive utility from the ratings. It could be because they care about the externalities the firms impose. Alternatively, it could be because it affects the firm's payoffs through risk (physical, reputational, litigation, or regulatory) or proxies for a long-term approach. Amel-Zadeh and Serafeim (2018) provide survey evidence that institutional investors care about all of these (see their Table 3). In our model, the only important elements are that these categories affect the investor's payoff from investing in the project and, therefore, the investor would like to learn more information about them.

In reality, the rating agencies themselves are unclear about and varied in their objectives, discussing both risk and impact (for a summary, see Larcker et al., 2022). The literature also finds very mixed evidence on the relationship between ratings and social outcomes (e.g., Raghunandan and Rajgopal, 2022) and the relationship between ratings and financial performance (e.g., Hartzmark and Sussman, 2019). Berg et al. (2024) find that combining ratings leads to a stronger relationship between ESG and financial performance.

Our results may explain why generalization is widespread in other markets where information intermediaries acquire information about multiple unrelated categories and affect investor (or consumer) decisions. Some examples include university rankings (from, e.g., U.S. News & World Report and the Financial Times; they collect information on faculty research, salaries of graduates, diversity, and other categories), real estate valuations (e.g., Zillow's Zestimate and Redfin; they collect information on property attributes, local schooling, crime, environmental quality and more), and employer ratings (from companies such as Glassdoor, Comparably, or Indeed; they aggregate employee reviews about employers in multiple categories such as culture, compensation, and management).

Our model is complementary to models of credit rating competition (e.g., Bolton, Freixas, and Shapiro (2012), Bar-Isaac and Shapiro (2013), Sangiorgi and Spatt (2017), Bouvard and Levy (2018), and Piccolo (2021)). Those papers study models with a single dimension that investors care about and for which information can be gathered (credit risk) and focus on stylized facts of the credit rating industry that are not present in the ESG ratings industry: the issuer-pays model (ESG ratings are predominantly investor-pays) and the resultant issuer shopping for ratings. Given this, it is natural for these papers to study conflicts of interest and distortions in information provision; these result in rating inflation and inefficient capital allocation. In contrast, in order to study information design with multiple dimensions, we assume that raters tell the truth and do not depend on reputation. The result that ESG raters will decrease surplus by generalizing is unique to this setting.

Of course, given that the model allows for multiple dimensions that affect investors' willingness to pay for the investment, one of those dimensions could be credit risk. In Section VI, we find that when investors value the dimensions differently, there can be a co-existence of specialist raters and generalist raters; this may explain why we see credit rating agencies providing ESG ratings alongside raters who only provide ESG ratings.

Our paper also contributes to the literature on information sales.[6] Admati and Pfleiderer (1986) study how a monopolistic information provider

[6] For a review, see Bergemann and Bonatti (2019).

should design signals about asset values when the value of its information is subject to dilution through equilibrium prices. Huang, Yang, and Xiong (2018) extend Admati and Pfleiderer (1986) to the case with multiple information sellers. A key distinguishing point of our model is that information providers decide how to gather information about multiple, possibly unrelated, variables, and then sell their assessments to investors. We show that raters may move away from specialization toward generalization, depending on the investor's preferences, while generalization by both raters results in inefficient duplication of effort, and consequently, less overall information production.

The multidimensionality of relevant information in financial decisions has been analyzed in models of financial markets (Goldstein and Yang, 2015; Goldstein et al., 2025) and bank lending (Blickle et al., 2025). We contribute to this literature by studying the incentives of information intermediaries to gather, design, and sell multidimensional information.

A growing body of literature theoretically investigates the functioning of capital markets in financing socially responsible investments (Piccolo, Schneemeier, and Bisceglia, 2022; Gupta, Kopytov, and Starmans, 2022; Piatti, Shapiro, and Wang, 2023; Oehmke and Opp, 2024), and regulations to facilitate the transition toward a green economy (Oehmke and Opp, 2022; Hong, Wang, and Yang, 2023; Huang and Kopytov, 2023; Döttling and Rola-Janicka, 2025; Inderst and Opp, 2025; Gupta and Starmans, 2024). Some studies have explored the asset pricing implications of ESG investing (e.g., see Pástor, Stambaugh, and Taylor, 2021; Sauzet and Zerbib, 2022 ; Avramov et al., 2022). This paper contributes to this literature by exploring the production of ESG information, which is a key input for the functioning of capital markets.

The production and sale of ESG information are also examined by some contemporaneous studies. Chen and Sun (2024) analyze an extension of Admati and Pfleiderer (1986) where the information available to a monopolistic information seller is subject to manipulation by firms, with an application to the design of ESG information. Lovo and Olivier (2025) examine when the "investor-pays" model (which is common for ESG ratings) is optimal for selling information, compared to the "issuer-pays" model commonly used in credit ratings. Our study examines competition among ESG raters and its impact on their trade-off between specialization and generalization across different categories and measurement methods. This analysis is crucial for understanding why ESG raters differ in their methods and measurements, factors that account for the majority of divergence in ESG ratings. Moreover, our study sheds light on how policies can enhance ESG information by recognizing its inherently multidimensional nature.

I. The Model

We consider a model of competition between two ESG raters to sell information to an investor. There are three periods in the model: $t = 0, 0.5, 1$. There is a project that requires investment $I > 0$, which generates a certain financial output of $I + \Delta$ at $t = 1$, where $\Delta > 0$. A deep-pocketed investor considers investing in the project. The investor is concerned about both the project's financial and ESG performance, where the latter is denoted by u. ESG performance may capture externalities of the project or long-term risk factors that contribute to or take away from the project's future cash flows. Therefore, ESG performance may be positive or negative. The investor assigns weight $\beta \geq 0$ to ESG performance. As such, from the investor's perspective, the total value created by the project is $\Delta + \beta u$.

The fundamental assumption in our model is that ESG performance has multiple dimensions, which can be unrelated. For instance, ESG rating agencies evaluate firms based on three major categories (environmental, social, and governance), with each category containing many sub-categories. The performance in one category (or subcategory) is not necessarily related to or informative about the performance in the others. This multidimensionality of non-financial performance differentiates our model from the equilibrium models of credit ratings, where the investors' objective is to learn about a default probability based on financial data.

For simplicity, we focus on two categories: A and B. The categories could represent the major categories (i.e., E, S, and G), or subcategories of major categories. The project's performance in A and B is characterized by (w^A, w^B), where the performance is binary in each category, i.e., $w^i \in \{L, H\}$, for $i = A, B$. A priori, the probability of the high state is $\eta \in (0, 1)$ for each category. The probabilities are independent across the categories. The project's type is publicly realized at $t = 1$.[7]

The equation below specifies the ESG performance (u) as a function of the project's type:

[7]The realization could occur with a probability of less than one without changing the insights of the model.

$$u = \begin{cases} u^{HH} & (w^A, w^B) = (H, H) \\ u^{HL} & (w^A, w^B) \in \{(H, L), (L, H)\} \\ u^{LL} & (w^A, w^B) = (L, L), \end{cases} \tag{1}$$

where:[8]

$$u^{HH} > 0, \quad u^{LL} < 0, \quad u^{HL} \in [u^{LL}, u^{HH}]. \tag{2}$$

We present our results given the assumption that the investor does not invest in the project if no information about the project's ESG performance is available. Assumption 1 formalizes this point. This assumption ensures that there is a demand for ESG information, as no investment would take place without the information.[9]

ASSUMPTION 1:

$$\Delta + \beta \mathbb{E}[u] = \Delta + \beta \big\{ \eta^2 u^{HH} + 2(1-\eta)\eta u^{HL} + (1-\eta)^2 u^{LL} \big\} < 0. \tag{3}$$

Furthermore, we assume that the investor is not indifferent between investing and not investing when the project's type is (H, L) or (L, H), as formalized by Assumption 2.

[8]Assuming symmetry in the investor's preference between (H, L) and (L, H) simplifies the demand side; the focus on different categories, for now, comes from the production side. In Section VI, we revisit the demand side and allow the investor to have asymmetric preferences as well; the results are similar.

[9]This assumption can be further relaxed; however, doing so would make the characterizations more involved without offering additional economic insight.

ASSUMPTION 2:

$$\Delta + \beta u^{HL} \neq 0. \tag{4}$$

This assumption ensures that learning the project's type in each category always has a positive marginal value for the investor.[10]

A. *ESG Rating Agencies*

Information about u is provided by two non-cooperative ESG rating agencies. The raters are indexed by $j = 1, 2$. At $t = 0$, the raters simultaneously design a rating technology that generates a rating for each of the two categories at $t = 0.5$. The ratings are denoted by $S_j = (s_j^A, s_j^B)$. The ratings are binary, i.e., $s_j^i \in \{h, l\}$, for $j = 1, 2$, and $i = A, B$. The rating technology is characterized by the pair $\boldsymbol{\lambda}_j = (\lambda_j^A, \lambda_j^B)$, where λ_j^i denotes the probability that the project receives a high rating in a category for which it has a high type, which we label as precision. Putting it differently,

$$P(s_j^i = h | w^i = H) = \lambda_j^i, \quad P(s_j^i = h | w^i = L) = 0, \quad j = 1, 2, \quad i = A, B. \tag{5}$$

[10]Note that $\Delta + \beta u^{HL} = 0$ implies that $P(\Delta + \beta u \geq 0 | w^A = H) = 1$ and $P(\Delta + \beta u \leq 0 | w^A = L) = 1$, since $u^{HH} \geq u^{HL} \geq u^{LL}$. Therefore, if the investor has perfect information about w^A, an optimal investment strategy is to invest when $w^A = H$ and not invest when $w^A = L$, regardless of her information about w^B. In this case, information about category B never impacts the investment outcome and thus has no marginal value. This possibility complicates the set of equilibrium outcomes, which we simplify by imposing Assumption 2.

Note that under this signal structure, high-ratings perfectly reveal the underlying state.[11] In Internet Appendix IA.J, we consider a more general set of rating technologies that allow for both false-positive and false-negative errors in the ratings, and demonstrate the robustness of our results to our assumption about the signal structure.

A few remarks about the ratings are in order. The raters report their ratings truthfully. Hence, we assume away any strategic behavior in the reporting of the ratings (as in, e.g., Bolton et al. (2012); Agrawal et al. (2024)). Moreover, the raters report the rating in each category separately, rather than reporting an aggregated version. This is consistent with the fact that ESG rating agencies provide the break-down of their ratings to their subscribers, and this clearly dominates an arbitrary aggregation rule.

Since there is no false-positive error in the ratings, a high rating in a category by a rater is enough to verify that the project has a high performance in that category. Therefore, it is helpful to introduce the following notation for the combined ratings:

$$s^i = \begin{cases} h & \text{if } s_1^i = h \text{ or } s_2^i = h \\ l & \text{if } s_1^i = s_2^i = l \end{cases} \qquad i = A, B. \tag{6}$$

The conditional probabilities for $s^i = h$, namely the probability of receiving

[11]Note that (5) implies that in the extreme case of $\lambda_j^i = 0$, namely when Rater j does not assess the project in category i, Rater j assigns rating l for category i. However, this rating is uninformative.

a high rating in category i from at least one of the raters, is:

$$P(s^i = h|w^i = L) = 0,$$

$$\lambda^i \equiv P(s^i = h|w^i = H) = 1 - (1 - \lambda_1^i)(1 - \lambda_2^i) = \lambda_1^i + \lambda_2^i - \lambda_1^i \lambda_2^i, \quad i = A, B. \tag{7}$$

We assume that the raters choose their rating technology under the following technological constraint:

$$\lambda_j^A + \lambda_j^B \leq 1. \tag{8}$$

This constraint implies that each rater can perfectly disclose the project's performance in one of the categories, or provide a noisy rating for both categories. As a result, it is possible for the raters to perfectly reveal the project's type collectively.[12] The ratings are independent across the raters conditional on the project's performance, which can capture different measurement protocols employed by ESG raters. The raters choose their technologies simultaneously; this is meant to capture the fact that even established raters continue adapting their design of ratings in response to the evolving market

[12]In Section VI, we demonstrate the robustness of our results to alternative cost structures. First, we allow information acquisition costs to differ across categories and between raters; when the costs differ across categories, the raters cannot perfectly reveal the project's type, while when they differ across raters, one may have a competitive advantage. Second, we study economies of scale in information acquisition – i.e., allowing the cost of acquiring information on a second category to depend on whether information was collected on a first category. The assumption above of symmetric costs helps us simplify the characterizations in the main model.

landscape.[13]

Two types of rating technologies are particularly important in our analysis:

DEFINITION 1: *We say rater $j \in \{1,2\}$* ***specializes*** *if $\boldsymbol{\lambda}_j \in \{(1,0),(0,1)\}$. We denote the specialized rating technologies in categories A and B by $\boldsymbol{\lambda}^{SP_A}$ and $\boldsymbol{\lambda}^{SP_B}$, respectively. Moreover, we say rater $j \in \{1,2\}$* ***generalizes*** *if $\boldsymbol{\lambda}_j = \boldsymbol{\lambda}^{GN} = (\frac{1}{2}, \frac{1}{2})$.*

B. The Ratings Market

At $t = 0$, after designing the rating technologies simultaneously, the raters observe the other rater's choice of rating technology and sequentially set fees ϕ_1 and ϕ_2. Rater 1 sets its fee first. The sequential fee-setting enables us to refine the equilibrium outcomes since the equilibrium would not always be unique had we assumed a simultaneous fee-setting.[14] In Section VI, we show the robustness of our results in the case where both raters set their fees first with a positive probability. And, while ESG is often considered to be a new

[13]In Section VI, we also study the case where rating technologies are chosen sequentially and find similar results.

[14]To see this point, consider the following conceptually similar game: Suppose two players simultaneously request a fraction of a cake, analogous to the overall surplus created by the ratings. The requests are accepted as long as their sum does not exceed one. Then, any pair of fractions that sum up to one constitutes an equilibrium. In Internet Appendix IA.G, we confirm that the equilibria in the main model (with sequential fee setting) are also subgame-perfect when fee-setting is simultaneous, although they may not be unique.

factor in financial markets, the ESG data provision market has firms who are long established and have significant market share (and, thus, may have a first mover advantage); e.g., MSCI, which has 25% market share,[15] uses the methodology of Innovest, who they acquired in 2010 and was founded in 1992 (Eccles, Lee, and Stroehle, 2020).

After observing the fees and rating technologies, the investor decides whose ratings to purchase.[16]

The investor can purchase ratings from one rater, from both raters, or not purchase at all. The investor cannot see the ratings before purchasing them; the information for each set of ratings is revealed after purchasing them. This is in line with the subscription model offered by ESG raters where payments occur prior to the realization of ratings (ESMA, 2022).[17]

We define $V(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)$ as the expected value of the ratings to the investor

[15] "MSCI shares will show the future of ESG," by Ross Kerber, *Reuters*, September 12, 2024.

[16] In Section VI, we show that the main insights of the model hold in extensions with multiple investors who have heterogeneous preferences either over categories or across raters.

[17] The investor may benefit from acquiring information sequentially from the raters, as the ratings from one rater may be sufficient for the investor, making it unnecessary to purchase from a second rater. Nevertheless, it is unclear whether the investor actually gains more surplus if sequential rating acquisition is anticipated by the raters. The analysis for this case is complex, because the rater who is approached second, say Rater 2, faces a screening problem: Its optimal fee depends on the distribution of the investor's willingness to pay after observing (s_1^A, s_1^B). This, in turn, complicates Rater 1's fee-setting, and Rater 1 may even prefer being approached second.

when bought together. Moreover, let $\boldsymbol{O}$ denote the uninformative rating technology, i.e., $\boldsymbol{O} = (0, 0)$. Therefore, $V(\boldsymbol{\lambda}_1, \boldsymbol{O})$ and $V(\boldsymbol{O}, \boldsymbol{\lambda}_2)$ are the "standalone" values of the ratings provided by Raters 1 and 2, respectively. Equation 9 formally defines the value function:

$$V(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) = \sum_{s^A, s^B \in \{h,l\}} \max\{0, \Delta + \beta \mathbb{E}[u|s^A, s^B]\} P(s^A, s^B). \tag{9}$$

In equation 9, s^A and s^B denote the possible realizations of the combined ratings. The expression $P(s^A, s^B)$ denotes the unconditional probability that $(s^A, s^B) \in \{h, l\}^2$ realizes. The maximum operator indicates that the investor may decide to invest or not given the realization of s^A and s^B.

Figure 1 displays the moves and the investor's payoff. We assume that the investor breaks ties in her purchasing decision according to the following order, with the first being the most favored: (1) purchasing from both raters, (2) purchasing only from Rater 1, (3) purchasing only from Rater 2, (4) no purchase. We analyze and discuss the equilibrium fees in detail in Section III.A.

Figure 2 presents the timeline of the model. Key proofs are in the Appendix. The Internet Appendix provides the details of the analyses and results for the extensions discussed in Section VI.

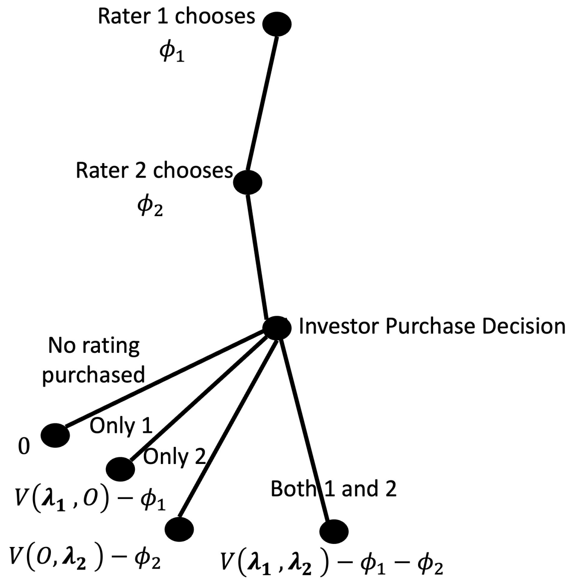


Figure 1. Sequence of actions and the investor's payoff in the ratings market stage

II. An Example

To illustrate how the investor's ESG preferences, as captured by β, impact the equilibrium design of the ratings, we first analyze a simple example.[18] Specifically, suppose the project's performance in each category is either H or L with equal probabilities (i.e., $\eta = 0.5$). Let $u^{HH} = 100$, and $u^{HL} = u^{LL} = -200$, meaning that the ESG performance is positive only when the project has a high performance in both categories. The financial payoff is set at $\Delta = 120$.

We simplify the raters' set of feasible rating technologies as follows: Rater 1 chooses between specialization in category A and generalization (i.e., $\boldsymbol{\lambda}_1 \in \{\boldsymbol{\lambda}^{GN}, \boldsymbol{\lambda}^{SP_A}\}$), and Rater 2 chooses between specialization in category B and

[18]We thank Deeksha Gupta for suggesting this.

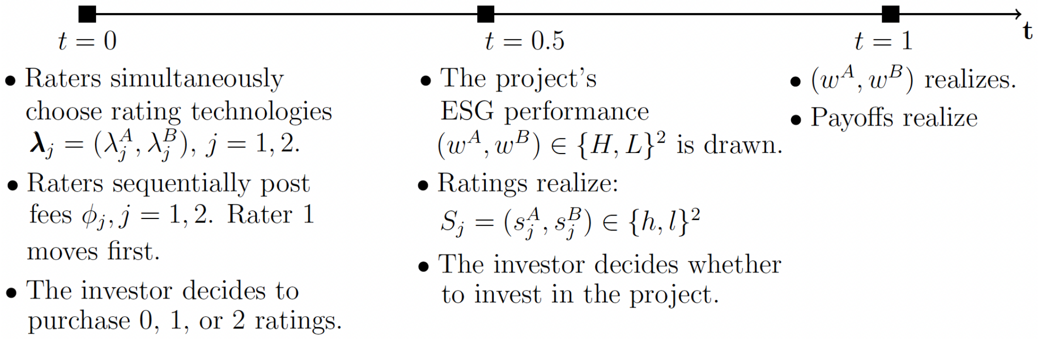


Figure 2. Timeline.

generalization (i.e., $\boldsymbol{\lambda}_2 \in \{\boldsymbol{\lambda}^{GN}, \boldsymbol{\lambda}^{SP_B}\}$). Specialization in a category perfectly reveals the project's performance in that category. In contrast, generalization generates a noisy rating for both categories. The timeline of events follows the main setup.

Note that the project's performance is perfectly revealed when both raters specialize, thereby maximizing the informativeness of the ratings. We analyze whether the outcome where both raters specialize is an equilibrium for $\beta = 1$ and $\beta = 3$.

A. *Equilibrium Outcome When* $\beta = 1$

When both raters specialize, the investor has four choices regarding which raters to purchase ratings from:

- The investor buys no ratings: The parameter values satisfy Assumption 1. Therefore, the investor does not invest in the absence of ESG information and earns an expected payoff of zero.

- The investor buys only from Rater 1 or 2: In this case, the investor perfectly learns the project's performance in one category. For instance, she learns w^B if she buys from Rater 2. If the ratings reveal that $w^B = H$, which occurs with probability $\frac{1}{2}$, the project's type is either (H, H) or (L, H) with equal probabilities, and her expected payoff (before fees) is:

$$V(\boldsymbol{O}, \boldsymbol{\lambda}^{SP_B}) = \frac{1}{2}\max\{0, \frac{1}{2}(120 + 1 \times 100) + \frac{1}{2}(120 + 1 \times -200)\} = 35. \tag{10}$$

The investor earns the same expected payoff if she only buys ratings from Rater 1.

- The investor buys ratings from both raters: The project's performance is perfectly revealed in this case. The investor only invests when the project has a high performance in both categories:

$$V(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{\lambda}^{SP_B}) = \frac{1}{4}(120 + 1 \times 100) = 55. \tag{11}$$

Therefore, when the raters specialize in different categories, the combined value of the ratings is 55, while the stand-alone value of each rater is 35. Rater 1 sets its fee at $\phi_1 = 55 - 35 = 20$, the marginal value of its ratings, as it cannot charge more. To see this, suppose that Rater 1 sets $\phi_1 = 21$. In response, Rater 2 sets ϕ_2 slightly less than 21, say $\phi_2 = 20.9$, and the investor purchases ratings only from Rater 2, as this maximizes her surplus:

$$\underbrace{V(\boldsymbol{O}, \boldsymbol{\lambda}^{SP_B}) - 20.9}_{35-20.9=14.1} > \max\{0, \underbrace{V(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{O}) - 21}_{35-21=14}, \underbrace{V(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{\lambda}^{SP_B}) - 21 - 20.9}_{55-41.9=13.1}\}. \tag{12}$$

Following $\phi_1 = 20$, Rater 2 also sets $\phi_2 = 20$, which is Rater 2's marginal value, and the investor purchases ratings from both raters. Note that $20 + 20 < 55$; the sum of the marginal values of the specialized rating technologies does not exceed their combined value.

Given a conjectured equilibrium where both raters specialize, if a rater, say Rater 1, switches to generalization, the combined value of the ratings declines because the generalization leads to noisier information about category A, without improving information produced about category B. In fact, the combined value of the ratings will decline to:[19]

$$\begin{aligned} V(\boldsymbol{\lambda}^{GN}, \boldsymbol{\lambda}^{SP_B}) &= \frac{1}{2} \times \frac{1}{4}(120 + 1 \times 100) \\ +\frac{1}{2} \times \frac{3}{4} \max\{0, \frac{2}{3}(120 + 1 \times -200) &+ \frac{1}{3}(120 + 1 \times 100)\} = 35. \end{aligned} \tag{13}$$

Since $V(\boldsymbol{O}, \boldsymbol{\lambda}^{SP_B}) = 35$, Rater 1 adds no value to Rater 2's information when their ratings bought together. Rater 1 cannot charge a positive value since its marginal value is zero. As a result, Rater 1, and similarly Rater 2, have no incentive to switch to generalization.

Therefore, specialization by both raters is an equilibrium in this case, and it can be shown that it is the unique equilibrium as well.[20]

[19] In Equation 13, the first term represents the case in which $s_1^A = s_2^B = h$, revealing that the project's performance is (H, H). The second term corresponds to the signal realization $s_1^A = l$ and $s_2^B = h$, which implies a conditional probability of $\frac{2}{3}$ for $(w^A, w^B) = (L, H)$ and $\frac{1}{3}$ for $(w^A, w^B) = (H, H)$. Note that if $s_2^B = l$, then $w^B = L$ with probability one, implying a negative investment payoff regardless of the project's performance in category A.

[20] The combined value when both raters generalize is $V(\boldsymbol{\lambda}^{GN}, \boldsymbol{\lambda}^{GN}) = \frac{9}{16} \times 55 = 30.94$, which is less than the stand-alone value of specialization. Generalization by both raters is

We now show that specialization by both raters is not an equilibrium for $\beta = 3$, i.e., when the investor places a larger value on ESG performance.

B. Equilibrium Outcome When $\beta = 3$

When $\beta = 3$, the stand-alone value of specialization is zero (i.e., $V(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{O}) = V(\boldsymbol{O}, \boldsymbol{\lambda}^{SP_B}) = 0$). To see this, suppose Rater 2 specializes in category B, and the investor buys ratings only from Rater 2. The investor's expected payoff (before fees) is:

$$V(\boldsymbol{O}, \boldsymbol{\lambda}^{SP_B}) = \frac{1}{2}\max\{0, \frac{1}{2}(120+3\times100)+\frac{1}{2}(120+3\times-200)\} = \max\{0, -15\} = 0. \tag{14}$$

In other words, the investor is so averse to projects with negative ESG performance that even an h-realization of Rater 2's rating in category B does not lead to investment. Therefore, the investor is not willing to pay a positive fee for these ratings when no other ESG information is available.

Similar to the previous case, the project's type is perfectly revealed when both raters specialize. The combined value of the ratings is:

$$V(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{\lambda}^{SP_B}) = \frac{1}{4}(120 + 3 \times 100) = 105. \tag{15}$$

When both raters specialize, Rater 1 sets $\phi_1 = 105$, and Rater 2 charges zero. This is because, with a zero stand-alone value, there is no $\phi_2 > 0$ that Rater 2 can set to attract the investor. Rater 2 can increase the stand-alone value of its ratings by generalizing, as it identifies projects with type (H, H)

not an equilibrium since both raters can charge more if they switch to specialization.

with probability $\frac{1}{4}$:

$$V(\boldsymbol{O}, \boldsymbol{\lambda}^{GN}) = \frac{1}{4} \times \frac{1}{4}(120 + 3 \times 100) = 26.25. \quad (16)$$

However, generalization reduces the combined value of the ratings:

$$V(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{\lambda}^{GN}) = \frac{1}{2} \times \frac{1}{4}(120 + 3 \times 100) = 52.5. \quad (17)$$

In this case, both raters set $\phi_1 = \phi_2 = 26.25$, and the investor buys ratings from both raters. Therefore, Rater 2 has benefited from switching to generalization, but this harms the overall information produced by the ratings. Specialization by both raters is not an equilibrium. In fact, the unique equilibrium is generalization by both raters.[21]

By comparing these two cases, we see that the quality of information produced about the project's ESG performance declines as its importance to the investor increases. When it becomes very important for the investor that the project has a high performance in both categories, she requires positive information about both categories to invest. This reduces the stand-alone value of specialization and increases that of generalization, and stand-alone values impact the raters' equilibrium fees. However, generalization by both raters leads to an inefficient duplication of effort, which harms the ratings' informativeness about the project's ESG performance. The following sections demonstrate this insight within the more general framework presented in

[21] When both raters generalize, the equilibrium fees are $\phi_1 = 32.81$, Rater 1's marginal value, and $\phi_2 = 26.25$, Rater 2's stand-alone value.

Section I.

III. Equilibrium Choice of Rating Technology

This section solves for the market equilibrium in the main framework, and contrasts the results with the rating technologies that maximize surplus.

A. *Equilibrium in the Ratings Market Stage*

In this section, we analyze the raters' equilibrium fees given a pair of rating technologies $(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)$. Lemma 1 provides the equilibrium fees.

LEMMA 1: *If the rating technologies chosen by the raters are* $(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)$*, then the following fees are set in equilibrium:*

$$\begin{aligned} \phi_1 &= V(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) - V(\boldsymbol{O}, \boldsymbol{\lambda}_2) \\ \phi_2 &= \min\{V(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) - V(\boldsymbol{\lambda}_1, \boldsymbol{O}), V(\boldsymbol{O}, \boldsymbol{\lambda}_2) - V(\boldsymbol{O}, \boldsymbol{O})\}. \end{aligned} \tag{18}$$

Lemma 1 states that Rater 1 is always paid the marginal value of its ratings, while Rater 2 may be paid less. The reason is that the sum of the marginal values can exceed the total value created by the ratings, depending on the choice of rating technologies. In other words,

$$\underbrace{V(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) - V(\boldsymbol{O}, \boldsymbol{\lambda}_2)}_{\text{Marginal value of Rater 1}} + \underbrace{V(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) - V(\boldsymbol{\lambda}_1, \boldsymbol{O})}_{\text{Marginal value of Rater 2}} \geq \underbrace{V(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) - V(\boldsymbol{O}, \boldsymbol{O}).}_{\text{Total value created by the ratings}} \tag{19}$$

Consequently, the investor is not willing to pay both raters their marginal values.[22] As Rater 1 moves first, it can always charge its marginal value. In this case, the maximum that Rater 2 can charge is:

$$\underbrace{V(\boldsymbol{\lambda}_1,\boldsymbol{\lambda}_2)-V(\boldsymbol{O},\boldsymbol{O})}_{\text{Total value created by the ratings}}-\underbrace{(V(\boldsymbol{\lambda}_1,\boldsymbol{\lambda}_2)-V(\boldsymbol{O},\boldsymbol{\lambda}_2))}_{\text{Marginal value of Rater 1}}=\underbrace{V(\boldsymbol{O},\boldsymbol{\lambda}_2)-V(\boldsymbol{O},\boldsymbol{O})}_{\text{Stand-alone value of Rater 2}} \tag{20}$$

We label this the stand-alone value of Rater 2's rating technology, as it is equal to the maximum amount the rater can charge if the investor were to buy only Rater 2's ratings. Therefore, Rater 2's fee is bounded by the stand-alone value of its ratings, affecting the rater's incentives in the design of its rating technology.[23] In equilibrium, the investor purchases both sets of ratings.

[22]By rearranging equation 19, we obtain:

$$V(\boldsymbol{\lambda}_1,\boldsymbol{\lambda}_2)-V(\boldsymbol{\lambda}_1,\boldsymbol{O})\geq V(\boldsymbol{O},\boldsymbol{\lambda}_2)-V(\boldsymbol{O},\boldsymbol{O}). \tag{19'}$$

This directly connects equation 19 to Rater 2's equilibrium payoff. Note that we can see from equation 19′ that the rating technologies are complements (the marginal value of Rater 2 is larger than its stand-alone value) when the inequality holds, and substitutes otherwise. We use this formulation in the appendix.

[23]In Section VI, we demonstrate the robustness of our results when both raters move first with a positive probability, and consequently, both raters' payoffs will depend on their stand-alone value with a positive probability.

B. Market Outcomes

With the characterization of the fees, we can analyze the equilibrium design of the ratings. Note that $(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)$ is an equilibrium outcome when $\boldsymbol{\lambda}_j$, $j = 1, 2$, maximizes fee ϕ_j, given the choice of the other rater. Proposition 1 characterizes the equilibrium outcomes in pure strategies. In section VI, we analyze the set of mixed strategy equilibria.

PROPOSITION 1: *Given Assumptions 1 and 2, the only possible pure strategy equilibrium outcomes are generalization by both raters and specialization by each rater in different categories. The following provides the characterization in detail (up to symmetries in the actions):*

a) *If $u^{HL} \geq 0$, the only equilibrium outcome is that the raters specialize in different categories.*

b) *If $u^{HL} < 0$, define:*

$$\beta^*(\lambda) = \sup\{\beta | \frac{(1-\eta)(\Delta + \beta u^{HL})}{\eta(\Delta + \beta u^{HH})} \geq \lambda - 1\}, \tag{21}$$

which is a decreasing function of λ.

b.1) *If $\beta^*(\frac{1}{4})$ is finite and $\beta > \beta^*(\frac{1}{4})$, then the unique equilibrium is generalization by both raters.*

b.2) *If $\beta \leq \beta^*(\frac{1}{4})$, specialization in different categories is an equilibrium. This is the only equilibrium outcome, except when $\beta \in [\beta^*(\frac{9}{16}), \beta^*(\frac{17}{32})] \cup \{\beta^*(\frac{1}{4})\}$. In this case, generalization by both raters is also an equilibrium.*

Proposition 1 states that if information about the project's ESG performance is essential for the investment (Assumption 1), and information about both categories always has a positive marginal value (Assumption 2), the only pure strategy equilibria are that the raters either both generalize or specialize in different categories. In Section VI, where we analyze the mixed strategy equilibria, we show that there is one additional equilibrium for some parameter values: one where the raters fail to coordinate on the category in which they specialize, and they randomize between specializing in categories A and B with equal probabilities.

Part (a) of Proposition 1 states that specialization in different categories is the only form of equilibrium when $u^{HL} \geq 0$. The intuition is that high performance in a single category is sufficient for the investor to invest, which implies a positive stand-alone value for specialization. In fact, specialization has the largest stand-alone value in this case. Moreover, specialization in different categories identifies the project type most efficiently, which means specialization also has the largest marginal value given the other rater specializes in the other category. Therefore, in response to the other rater specializing, specialization in the other category is the best response for both raters. More generally, the best response to any strategy is to specialize in one of the two categories, as illustrated in Figure 3A. In addition, both ratings are purchased in equilibrium since each rater identifies high-performing projects in a different category; thus, both ratings have a positive marginal value.

When $u^{HL} < 0$, the equilibrium outcome depends on the investor's decision when the project has a mix of positive and negative ratings, i.e., $s^A = h$ and $s^B = l$, or vice versa. To this end, function $\beta^*(\lambda)$, where $\lambda \in [0, 1]$, represents the maximum value of β at which the investor is willing to invest when (s^A, s^B) is (h, l) or (l, h), and the combined precision of the category with the low rating is λ. A lower value of λ is associated with a higher probability of a false-negative error in the low rating, and consequently, a greater likelihood of the project type being (H, H); this implies there is a larger range of β for which an investor would invest and, consequently, the threshold $\beta^*(\lambda)$ is higher. For instance, when $\beta^*(0)$ is finite and $\beta > \beta^*(0)$, β is so large that a single high performance never convinces the investor to invest, even if the available ratings in the other category are perfectly uninformative. Specifically, we have:

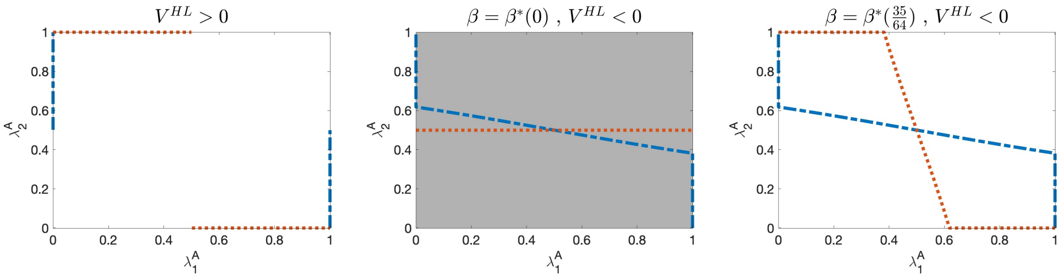
$$\begin{aligned} \beta > \beta^*(0) &\Rightarrow \eta(\Delta + \beta u^{HH}) + (1-\eta)(\Delta + \beta u^{HL}) < 0 \\ &\Rightarrow \mathbb{E}[\Delta + \beta u | w^A = H] = \mathbb{E}[\Delta + \beta u | w^B = H] < 0. \end{aligned} \tag{22}$$

In this case, learning only w^A or w^B holds no value for the investor since she invests only when she receives a positive update about both categories. Therefore, Rater 2 cannot charge a positive fee when it specializes, pushing the rater toward generalization. In fact, as depicted in Figure 3B, generalization is Rater 2's best response to any choice of Rater 1 since generalization yields the highest stand-alone value. In response, Rater 1 moves away from specialization and chooses to generalize.

More broadly, stand-alone values play a role in the market equilibrium when β is sufficiently large. In our model, generalization has the highest stand-alone value when $\beta^*(\frac{1}{4})$ is finite and $\beta > \beta^*(\frac{1}{4})$. In other words, when this condition holds, the investor prefers obtaining a noisy signal about each category to a perfectly-revealing signal about a single category. Given this, Rater 2 generalizes, and consequently, the unique equilibrium is generalization by both raters. Here, the investor's strong preference for ESG investments leads, surprisingly, to less information being provided.

Another key force that shapes the market equilibrium is the strategic complementarity in the design of ratings. To see the strategic complementarity, note that the marginal value of specialization in a category is the highest when the other rater specializes in the other category. As the other rater moves from specialization towards generalization, the overlap between the two rating technologies increases, which reduces the marginal value of specialization. Conversely, the marginal value of generalization increases as the other rater moves from specialization to generalization. This strategic complementarity can result in both an equilibrium where both raters specialize in different categories or an equilibrium where both generalize, as illustrated by Figure 3C. This occurs for some intermediate values of β, specifically (1) when $\beta \geq \beta^*(\frac{9}{16})$, corresponding to when the marginal value of generalization is higher than the marginal value of specialization, given the other rater generalizes $(V(\boldsymbol{\lambda}^{GN}, \boldsymbol{\lambda}^{GN}) - V(\boldsymbol{O}, \boldsymbol{\lambda}^{GN}) \geq V(\boldsymbol{\lambda}^{SP_i}, \boldsymbol{\lambda}^{GN}) - V(\boldsymbol{O}, \boldsymbol{\lambda}^{GN}))$, $i = A, B$; and (2) $\beta \leq \beta^*(\frac{17}{32})$, which ensures that Rater 2's payoff is equal to the marginal value of generalization, not its stand-alone value.[24] Figure 4

[24]The intuition for this condition is that β impacts the comparative value (ratio) of the expected payoffs from investing in projects with a positive rating in only one category over the expected payoffs from investing in projects with positive ratings in both categories. Given $u^{HL} < 0$, this ratio decreases with β, which means that for low values of β there is a lower premium in getting a high rating from both categories. This implies that both raters will be able to charge their marginal values when β is sufficiently low. Put differently, equation 19 will not hold and there is enough surplus for both raters to charge



Panel A Panel B Panel C

Figure 3. This figure displays the best response of each rater (blue (dashed) line for Rater 1 and red (dotted) line for Rater 2) for different values of β and u^{HL}. The intersection points designate the equilibrium outcomes. The white (gray) area represents the outcomes in which Rater 2 charges the marginal value (stand-alone value) of its ratings.

illustrates how the set of market equilibria varies with β when $u^{HL} < 0$.

In summary, Proposition 1 reveals that the raters' incentives to specialize or generalize depend on whether the project's performance in the two categories are substitutes or complements for the investor. When $u^{HL} \geq 0$, high performance in a single category is sufficient for investment. This ensures that specialization has substantial stand-alone value, pushing the raters toward specialization. In contrast, when $u^{HL} < 0$ and β is sufficiently large (specifically, when the condition in equation 22 holds), high performance in a single category is never sufficient for investment. As a result, specialization has little stand-alone value, pushing the raters toward generalization.

their marginal values.

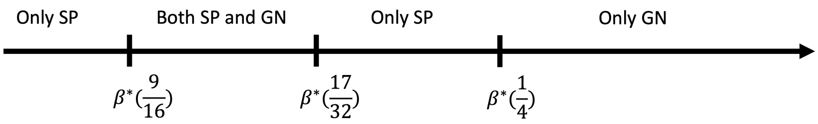


Figure 4. This diagram illustrates the set of pure strategy market equilibria for different values of β when $u^{HL} < 0$ and $\beta^*(\frac{1}{4})$ is finite. "GN" and "SP" represent generalization by both raters and specialization in different categories, respectively.

C. Value-maximizing Ratings

Next, we analyze the choice of rating technologies that maximize the expected investment value. In particular, we call pair $(\boldsymbol{\lambda}_1^*, \boldsymbol{\lambda}_2^*)$ "value-maximizing" if it is a global maximizer of $V(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)$, as defined in equation 9, given the technological constraint in equation 8.

A value-maximizing pair can be interpreted in a number of ways. First, it can describe the solution to the social planner's problem when the planner maximizes the total surplus of the investor and the raters. Second, it can describe the situation where the social planner's objective is to maximize the expected social value of the investment and the social planner has the same preferences (i.e., β) as the investor. Third, if the raters could collude and jointly decide about their rating technologies, they would choose a value-maximizing pair. As such, value-maximizing pairs provide us with a natural benchmark to study the impact of raters' competition on the production of ESG information. Proposition 2 demonstrates that the only value-maximizing outcome is specialization in different categories.

PROPOSITION 2: *Given Assumptions 1 and 2, the only value-maximizing outcome is specialization in different categories, i.e.,* $(\boldsymbol{\lambda}_1^*, \boldsymbol{\lambda}_2^*) \in$

$\{(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{\lambda}^{SP_B}), (\boldsymbol{\lambda}^{SP_B}, \boldsymbol{\lambda}^{SP_A})\}$.

Note that when the raters specialize in different categories, the project's type is perfectly revealed to the investor, which results in perfect investment efficiency. In other words, the investment takes place iff it generates a positive value from the perspective of the investor (i.e., $\Delta + \beta u > 0$). No other pair of rating technologies implements this investment outcome since they result in some inefficient overlap in the information produced for each category.

By juxtaposing Propositions 1 and 2, we learn that competition in the production of ESG information results in generalization when β is large and $u^{HL} < 0$, while the value-maximizing solution is specialization. Note that the deviation from the value-maximizing outcome happens when the investor is the most concerned about the project's ESG performance. Interestingly, this implies a non-monotonic relationship between the amount of information provided and the investor's preference β; when β is low, there are no ratings or information production; when β is intermediate, specialization in different categories is an equilibrium and perfect information transmission can occur; and when β is large, raters generalize, resulting in noisy signals.

The underlying insight here is that, while the specialization outcome maximizes the investment value, Rater 2 has incentives to generalize because its payoff is tied to the stand-alone value of its ratings. In other words, the strong demand for information about both categories A and B creates an incentive for raters to generalize in equilibrium, as specialized ratings hold no stand-alone value for the investor. Of course, as demonstrated earlier, strategic complementarities in the design of ratings can lead to generalization as well.

IV. Greenwashing and the Design of Ratings

In our model, we assume that the ESG raters can assess the true type of the project. However, in practice, ESG raters mostly utilize self-disclosed information provided by firms for their assessments. This means that there is room for manipulation, which is typically referred to as "greenwashing" in this context (for example, see Baker et al. (2024a) and Cornaggia and Cornaggia (2025)). For instance, a firm might announce a plan to reduce carbon emissions, which would help the firm receive a better rating in the environmental category. However, it would likely be difficult to assess the firm's commitment to the plan. In this section, we analyze how greenwashing impacts the raters' design of their ratings.

Specifically, suppose the raters assess a potentially manipulated type of the project, which we denote by (w_M^A, w_M^B), where $w_M^A, w_M^B \in \{H, L\}$. In particular, with probability α, the project designer can successfully manipulate the type for any category with a low performance. The equations below describe the relationship between manipulated and actual types:

$$Prob(w_M^i = H | w^i = H) = 1, \quad Prob(w_M^i = H | w^i = L) = \alpha \in [0, 1), \quad i = A, B. \tag{23}$$

The probability α is lower when ESG disclosure requirements are tightened, or when greenwashing is costlier. The main model in the text corresponds to the case that greenwashing is not possible – i.e., $\alpha = 0$.

We maintain the mapping between types and ratings in equation 5 with the difference being that the types are the manipulated types:

$$P(s^i_j = h | w^i_M = H) = \lambda^i_j, \quad P(s^i_j = h | w^i_M = L) = 0, \quad j = 1, 2, \quad i = A, B. \tag{24}$$

Note that the equations in (24) indicate that greenwashing results in the manipulation of the ratings. In our model, the investor forms beliefs rationally. Hence, she correctly accounts for the possibility of greenwashing in her evaluation of the project's ESG performance given a set of ESG ratings.

We assume that α is such that the ratings still induce investment when the project receives a high rating in both categories. In particular,

$$\begin{aligned} &\Delta + \beta \mathbb{E}[u | s^A = s^B = h] = \Delta + \beta \mathbb{E}[u | w^A_M = w^B_M = H] \\ &= \Delta + \frac{\beta}{(\eta + (1-\eta)\alpha)^2} \Big\{ \eta^2 u^{HH} + 2\eta(1-\eta)\alpha u^{HL} + (1-\eta)^2 \alpha^2 u^{LL} \Big\} > 0. \end{aligned} \tag{25}$$

Note that greenwashing introduces additional noise to the ESG ratings. For instance, a high rating in both categories does not necessarily indicate high performance in both. Consequently, greenwashing causes investors to discount the expected ESG performance based on these ratings. For example, if the investor requires a high performance only in one category when there is no greenwashing, she might require a high rating in two categories to account for the possibility of greenwashing. This can be thought of as a hedging mechanism against the greenwashing risk. In Proposition 3, we describe how greenwashing impacts the equilibrium design of ratings, through its effect on the investor's demand for ESG information.

PROPOSITION 3: *For any values of β and u^{HL}, there exists an α^* such that the unique equilibrium is generalization by both raters when $\alpha > \alpha^*$ and α satisfies the condition in equation 25.*[25]

Proposition 3 states that when the amount of greenwashing is sufficiently large, the unique equilibrium is generalization by both raters.[26] The intuition is that when α is large enough, the investor requires a high rating in both categories to invest in the project. As a result, the stand-alone value of specialization is zero, which causes the raters to move away from the specialization outcome. Note that specialization by both raters in different categories remains the unique value-maximizing outcome since any other pair of rating technologies, including generalization by both raters, produces a garbled version of the information provided by the specialization outcome.

This result reveals a propagation mechanism through which greenwashing contaminates ESG ratings. Greenwashing does not only result in noisier ESG information, but also might push raters away from specialization, as investors demand high ratings across a larger number of categories to hedge against greenwashing concerns. Therefore, greenwashing indirectly harms the quality of ESG ratings through its influence on the organization of the rating industry.

---

[25]The proof in the appendix confirms that the set of α's satisfying these conditions is non-empty.

[26]The amount of greenwashing (i.e., α) must not be so large as to make the information in ESG ratings valueless (this is embodied in equation 25).

V. Implications of the Model

In this section, we consider the empirical and policy implications of the model. Some of these are more tentative and present opportunities for future research.

A. Divergence in Ratings

Investors rely on ESG ratings and information to incorporate ESG considerations in their investments. Nevertheless, the available ESG ratings vastly disagree in their assessment of firms' ESG performance, potentially creating confusion for investors (Chatterji et al., 2016; Berg et al., 2022). In this section, we analyze the disagreement in the ratings implied by different rating technology choices.

To analyze disagreement, we examine the correlation between the expected investor payoffs implied by each rater's ratings; that is,

$$\begin{aligned} Agg(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) &= Corr(\mathbb{E}[\Delta + \beta u | s_1^A, s_1^B], \mathbb{E}[\Delta + \beta u | s_2^A, s_2^B]) \\ &= Corr(\mathbb{E}[u | s_1^A, s_1^B], \mathbb{E}[u | s_2^A, s_2^B]). \end{aligned} \tag{26}$$

We use this method of aggregation rather than being limited to an arbitrary aggregation rule. Figure 5 illustrates how the disagreement varies with the choice of rating technology. Lighter (darker) points indicate less (more) agreement. We see that specialization in different categories, which is value-maximizing, results in the highest level of disagreement. The agreement level is intermediate when both raters generalize. The highest level of agreement would be obtained if both raters specialize in the same category.

We formalize this result in Proposition 4.

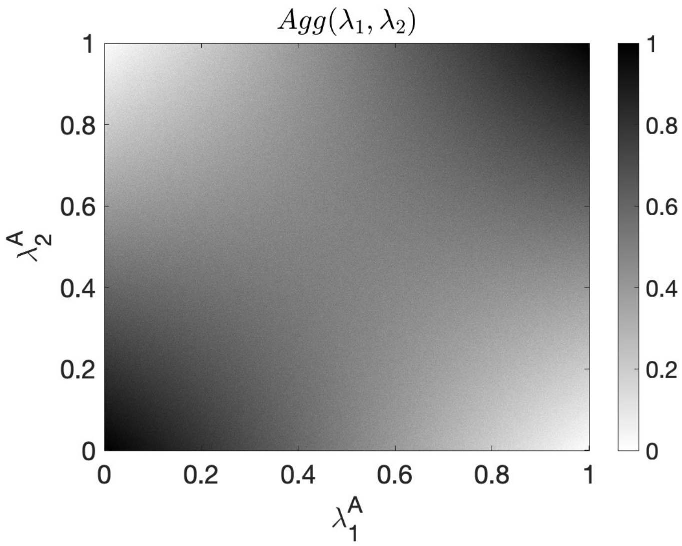


Figure 5. This figure illustrates agreement in the expected investor payoffs implied by the ratings ($Agg(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)$), as specified in equation 26. Darker values correspond to higher levels of agreement. Top-left and bottom-right points correspond to the outcomes in which the raters specialize in different categories. The center point corresponds to the generalization outcome.

PROPOSITION 4: *$Agg(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)$ is minimized when the raters specialize in different categories, and maximized when the raters specialize in the same category.*

To understand this result, note that disagreement arises from two sources: 1) providing inaccurate, but independent, ratings for a category, or 2) allocating resources differently among the categories. When both raters generalize, the former is the only source of discrepancy, and when the raters specialize in different categories, the latter is the only source. As a result, the disagreement observed in data can be attributed to a combination of these two sources, which stem from specialization and technological limitations. Therefore, disagreement in ESG ratings is efficient for investors when specialization is the driver.[27]

Berg et al. (2022) find that 38% of the discrepancy in the category ratings provided by major ESG rating agencies can be attributed to the differences in the subcategories examined (scope). This indicates that indeed a significant portion of the disagreement reflects specialization by the raters, and can thus be beneficial from a social welfare perspective.[28]

B. *Further Empirical Implications*

Changes in investor weights on ESG performance. We mentioned earlier in the text that when the weight investors put on ESG performance β is low, there will be no market for ratings. When β is intermediate, specialization is always an equilibrium, and when β is large, generalization is the unique equilibrium if $u^{HL} < 0$. This indicates that rating design could change discretely with investor valuations of ESG factors. Pástor, Stambaugh, and Taylor (2022) and Ardia et al. (2023) measure concerns about climate change using a media index and focus on shocks to these concerns, while Baker, Egan, and Sarkar (2024b) use a revealed preference approach in

[27]This intuition is robust to the way the ratings are aggregated. For instance, had the raters reported a precision-weighted average of their ratings, specialization in different (the same) categories would still induce the lowest (highest) correlation in the aggregated ratings.

[28]Berg et al. (2022) find that divergence in measurement methodologies also substantially contributes to the discrepancy in the ratings. In Section VI, we analyze the raters' trade-offs in choosing among different measurement methods.

estimating demand for ESG-oriented index funds. Similarly, this comparison could be done in reverse with financial returns (measured in our model by Δ) – when financial returns dominate ESG returns, we would expect little ESG information, while when financial returns go down, information will first become more accurate (specialization) and then less so (generalization).

Greenwashing. The model predicts that changes in greenwashing can (discretely) affect the information provision of ESG raters. When there is little greenwashing, our main model holds. With an intermediate level of greenwashing, the unique equilibrium is generalization. With a large level of greenwashing, there will be no ratings. Regulatory shocks that reduce greenwashing opportunities, such as EU disclosure mandates, should lead to weakly more informative ratings.

C. Policy Implications

Price Caps. The key friction in the model is that Rater 2's fee is bounded by the stand-alone value of its ratings, due to the pricing power of Rater 1. Therefore, when the investor is very concerned about the project's low ESG performance in both categories, Rater 2 can charge a larger fee by generalizing, instead of specializing. In this case, specialization in different categories, which is the efficient outcome, does not constitute an equilibrium, as shown in Proposition 1.

Setting a price cap may be a solution to improve Rater 2's pricing power, and thereby restore the specialization in different categories outcome as an equilibrium. Specifically, suppose that fees are not allowed to exceed $\bar{\phi}$. With this price cap, the equilibrium fees for a given pair of rating technologies $(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)$ is:

$$\begin{aligned} \phi_1 &= \min\{\bar{\phi}, V(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) - V(\boldsymbol{O}, \boldsymbol{\lambda}_2)\} \\ \phi_2 &= \min\{\bar{\phi}, V(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) - V(\boldsymbol{\lambda}_1, \boldsymbol{O}), V(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) - \phi_1\}. \end{aligned} \tag{27}$$

One can show that by setting $\bar{\phi} = \frac{1}{2}V(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{\lambda}^{SP_B})$, Rater 2's best response to specialization by Rater 1 is always specialization in the other category. This recovers specialization in different categories as an equilibrium for all values of β and u^{HL} in our framework.[29]

There is a potential drawback. We do not model explicit costs of information acquisition in the model,[30] but we note that Rater 1 can charge the same price-capped fee $\bar{\phi}$ even if it slightly reduces the accuracy of its rating (e.g., from $\boldsymbol{\lambda}_1 = (1, 0)$ to $\boldsymbol{\lambda}_1 = (1-\varepsilon, 0)$). If this would result in a lower information acquisition cost, the price cap would thus distort Rater 1's incentives for information production. Therefore, by imposing a price cap, policymakers may be able to incentivize specialization by ESG raters, though this may come at a cost of lower rating quality.

[29]For large values of β, where $V(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{\lambda}^{SP_B}) > V(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{O}) + V(\boldsymbol{O}, \boldsymbol{\lambda}^{SP_B})$, Rater 1's price cap binds when the raters specialize in different categories, i.e., $\phi_1 = \bar{\phi}$. In response to Rater 1's specialization, specialization in the other category is a best response for Rater 2 since ϕ_2 in equation 27 is weakly increasing in $V(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)$ when $\phi_1 = \bar{\phi}$.

[30]In our model, we assume that all rating technologies that satisfy the technological constraint in equation 8 are equally costly to obtain tractability and focus on the trade-off between specialization and generalization.

Collusion. The model's results suggests that collusion would facilitate the outcome of specialization in different categories, as it would eliminate any coordination failure and maximize the joint surplus of the raters. This is because competition can cause raters to provide an inefficient amount of information. However, while this maximizes total surplus in the model, this could easily come at a cost to investor surplus. Furthermore, this could make entry into and innovation in (both unmodeled here) the ESG ratings market more costly.

Standardization: We prove that both a lack of correlation between ESG ratings and using different measures/data may be good for surplus. Therefore, attempts to standardize ESG ratings (beyond creating transparency) may be misguided and reduce information provision.

VI. Robustness of the Model

This section discusses the results under alternative assumptions and provides additional insights. The timing of these extensions are the same as the baseline model unless otherwise specified in the text. To keep the section concise, details are provided in the Internet Appendix.[31]

[31] In addition to the analyses described in this section, in Internet Appendices IA.I, IA.J, and IA.K, we extend the analysis to incorporate mixed strategies, false-positive errors in the ratings, and unbundling of the ratings in the design of rating technologies and demonstrate the robustness of our main results.

A. Asymmetric Investor Preferences

Investor Heterogeneity in Valuation. The baseline model considers the demand of a single investor for ESG ratings, abstracting away from potential heterogeneity in investor preferences. In Internet Appendix IA.A, we analyze an extension that features multiple types of investors that vary in their ESG preferences; specifically, there are type-A and type-B investors, where type-i investors ($i = A, B$) are primarily concerned about the project's low performance in category i (which we refer to as their primary category) and are less concerned about low performance in the other category (their secondary category).[32] By incorporating investor heterogeneity, the analysis enables us to capture additional incentives for raters to specialize or generalize that are absent in the baseline model. For instance, by specializing, the raters can cater to the information needs of a particular investor group, while by generalizing, they can broaden their client base.[33]

We find that generalization by both raters is the unique equilibrium when investors are sufficiently concerned about the project's performance in both categories, in line with our baseline results. If investors are only moderately concerned about their secondary category, the generalization outcome

[32] We assume that each type-i investor is sufficiently concerned about the project's performance in category i that they require positive information about it to invest. The parametric restrictions used for the analysis are formally specified in equation IA.1 and Assumption IA.1.

[33] We assume that the project is flexibly scalable, such that both investment costs and payoffs are defined on a per-unit basis.

is not an equilibrium (except for a knife-edge case), whereas specialization in different categories is an equilibrium. In this equilibrium, investors purchase ratings from both raters, maximizing the total surplus. However, in the case where investors are only minimally concerned about their secondary category, the raters specialize in different categories and investors purchase ratings *only* from the rater that specializes in their primary category. As a result, our model predicts that extreme heterogeneity in ESG preferences results in a segmented market for ESG ratings.

Asymmetric Preferences Over Categories. The baseline model considers symmetric preferences over the project's performance in the two categories. Specifically, we assume that the investor values projects with ESG performance $(w^A, w^B) = (H, L)$ and (L, H) equivalently. Although asymmetric preferences are considered for each investor in the part above on Investor Heterogeneity in Valuation (analyzed in Internet Appendix IA.A), the preferences – and consequently the market outcomes – are symmetric in the aggregate. We relax this symmetry in Internet Appendix IA.B, by considering a single investor who values project types (H, L) and (L, H) differently.

The results provide insights on how the raters' equilibrium behavior changes as the investor becomes more concerned about only one category. For instance, suppose we start from the scenario in which the investor has symmetric preferences and $\beta > \beta^*(\frac{1}{4})$, where the unique equilibrium is generalization by both raters. If the investor becomes less concerned about category B, our results in Internet Appendix IA.B imply that this shift can lead to either specialization by both raters, or to an outcome where Rater 1 specializes in category A while Rater 2 chooses an interior rating technology.

Now, suppose we start from the symmetric case when $u^{HL} < 0$ and $\beta < \beta^*(\frac{9}{16})$, where the only equilibrium is that the raters specialize in different categories. Suppose Rater 1 specializes in category A and Rater 2 specializes in category B. Our results reveal that if the investor becomes sufficiently more concerned about category A, Rater 2 moves away from specialization as the stand-alone value of specialization in category B declines, and starts producing some information about category A. This implies a decline in the precision of Rater 2's rating in category B as a result of the increased importance of category A.

These results predict a co-existence of generalists and specialists when investors value performance across the categories asymmetrically. In fact, while most ESG rating providers generalize across multiple categories, some agencies specialize in a narrower set of variables. A notable example is the Carbon Disclosure Project (CDP), which focuses specifically on rating the environmental impact of various entities.

Furthermore, if we interpret the two categories broadly as credit risk and ESG information, this result is then in line with the recent moves by major credit rating agencies to expand into providing ESG information. In particular, in 2019, S&P Global acquired the ESG rating unit of RobecoSAM and Moody's acquired Vigeo Eiris,[34] which was previously a European ESG rating agency. Fitch also launched its ESG ratings product in 2021.[35] How-

[34]Moody's stopped offering stand-alone ESG ratings in 2024, as a result of a partnership with MSCI.

[35]In response to the recent shift toward ESG investing, several other major providers

ever, MSCI focuses only on producing ESG information and not credit risk information.[36]

Heterogeneous Preferences over Raters. In our baseline model, the investor has no preference between the two ESG rating providers and faces no restrictions on purchasing ratings from either one. In practice, however, some investors may prefer to obtain ratings from specific ESG providers – either due to proximity or experience with other products the rater provides. This could lead to segmentation in the ESG ratings market, which in turn would affect the raters' pricing behavior, market power, and their incentives in designing ratings.

We analyze the effect of this segmentation on the equilibrium outcome in Internet Appendix IA.C. Specifically, we consider 3 groups of infinitesimal investors: those who buy ESG ratings only from Rater 1, those who buy only from Rater 2, and those who buy from both raters. Let $Z_1 = Z_2$ be the mass of captive investors for each rater, and let Z_{12} denote the mass of investors the two raters compete for. In the extreme case where $Z_{12} = 0$ and investors only buy ratings from a single rater, the raters are monopolists and

of financial information have also entered the ESG ratings market. Examples include Morningstar's acquisition of Sustainalytics and LSEG's acquisition of Refinitiv.

[36]In this analysis, it is important to note that the business model for credit ratings is different from ESG ratings (Lovo and Olivier, 2025). Credit ratings are typically paid for by issuers, whereas most ESG rating providers, including those that are owned by credit rating agencies, follow an investor-pays model. Additionally, in the model, ratings for multiple categories are bundled and sold together; this does not typically happen with credit and ESG ratings.

maximize the stand-alone value of their ratings, given the investors' preferences. Specifically, if $u^{HL} < 0$, generalization maximizes the stand-alone value when $\beta > \beta^*(\frac{1}{4})$, whereas specialization in either category maximizes the stand-alone value when $\beta < \beta^*(\frac{1}{4})$.[37] In the latter case, the two raters may specialize in the same category or in different categories.

We demonstrate the robustness of our main results to segmentation in the ESG ratings market given general values of Z_{12}. Specifically, we show that generalization by both raters is an equilibrium only when $u^{HL} < 0$ and $\beta \geq \beta^*(\frac{1}{4})$, while specialization in different categories is an equilibrium only when $u^{HL} > 0$, or $u^{HL} < 0$ and $\beta^* \leq \beta^*(\frac{1}{4})$.[38] A difference with the equilibrium outcomes in the baseline model, characterized in Proposition 1, is that generalization by both raters is no longer an equilibrium in the intermediate region of $\beta \in [\beta^*(\frac{9}{16}), \beta^*(\frac{17}{32})]$ when $u^{HL} < 0$.[39] Proposition

[37]By comparing this result with the baseline result in Proposition 1, we see that the shift to generalization occurs at the same threshold value of $\beta^*(\frac{1}{4})$. This is because in the baseline model, Rater 2's fee is bounded by its stand-alone value, which is maximized by generalization when $\beta > \beta^*(\frac{1}{4})$, and the rater attains this maximum stand-alone value.

[38]We do not demonstrate the uniqueness of these equilibrium outcomes, though we know from the baseline model that when $Z_1 = Z_2 = 0$, generalization by both raters will be unique when $u^{HL} < 0$ and $\beta > \beta^*(\frac{1}{4})$.

[39]The intuition is that if a rater, say Rater 2, switches from generalization to specialization, the marginal value of generalization by Rater 1 falls to zero (i.e., $V(\boldsymbol{\lambda}^{GN}, \boldsymbol{\lambda}^{SP_B}) = V(\boldsymbol{O}, \boldsymbol{\lambda}^{SP_B})$). As such, instead of setting its fee equal to this marginal value of zero, Rater 1 charges its captive investors its stand-alone value (i.e., $\phi_1 = V(\boldsymbol{\lambda}^{GN}, \boldsymbol{O})$). Note that this mechanism is absent in the baseline model, as there are no captive investors. In response, Rater 2 charges its stand-alone value to both the shared group and its captive group, yielding a payoff of $V(\boldsymbol{O}, \boldsymbol{\lambda}^{SP_B})(Z_2 + Z_{12})$. We show that this payoff exceeds Rater 2's

IA.3 in Internet Appendix IA.C details the equilibrium fees and the resulting market organization under different equilibrium outcomes.

B. Market Power

Unequal Information Acquisition Costs Across Raters and Categories. Our baseline model assumes that the raters face the same information acquisition costs, and the costs are equal for the two categories. In practice, raters may differ in their information acquisition costs or in the resources available to them. Furthermore, performance may be easier to measure in some categories than others (for example, emissions data is standardized, while biodiversity risk is more complex to assess (Giglio et al., 2025)).

In Internet Appendix IA.D, we analyze the market equilibrium and value-maximizing outcomes when the cost of information acquisition differs between (i) raters and (ii) categories. Specifically, we consider the following extension of the technological constraint in the main model for the raters:

$$\begin{aligned} &\lambda_1^A + \frac{\lambda_1^B}{b_1} \leq 1, && b_1 < 1, \\ &\lambda_2^A + \frac{\lambda_2^B}{b_2} \leq \bar{\lambda}, && b_2 \leq b_1, \bar{\lambda} \in (0,1]. \end{aligned} \tag{28}$$

In (28), b_1^{-1} and b_2^{-1} capture the relative difficulty of acquiring information

payoff under generalization, in response to generalization by Rater 1. Therefore, generalization by both raters cannot be an equilibrium in this intermediate region. Instead, specialization in different categories constitutes an equilibrium in this region.

about category B compared to category A for Rater 1 and Rater 2, respectively. The parameter $\bar{\lambda}$ reflects the amount of resources available to Rater 2 to allocate between the two categories. Conditions $\bar{\lambda} \leq 1$ and $b_2 \leq b_1$ imply that Rater 1 is superior in rating design; Rater 1 finds it easier to collect information about category B and has more resources. Any rating technology that is feasible for Rater 2 is also feasible for Rater 1, but not vice versa.

We denote the specialized rating technologies for the raters by $\boldsymbol{\lambda}_j^{SP_i}$, $j = 1, 2$ and $i = A, B$, where:

$$\boldsymbol{\lambda}_1^{SP_A} \equiv (1, 0), \quad \boldsymbol{\lambda}_1^{SP_B} \equiv (0, b_1), \quad \boldsymbol{\lambda}_2^{SP_A} \equiv (\bar{\lambda}, 0), \text{ and } \quad \boldsymbol{\lambda}_2^{SP_B} \equiv (0, b_2\bar{\lambda}). \tag{29}$$

Note that the generalization rating technology $\boldsymbol{\lambda}^{GN} = (\frac{1}{2}, \frac{1}{2})$ is not feasible with the technological constraints in (28).

In line with the baseline results, we find that if $u^{HL} < 0$ and β is sufficiently large ($\beta > \beta^*(\frac{b_2\bar{\lambda}}{4})$), the unique equilibrium outcome is that both raters choose interior rating technologies – the analogue of generalization in this model. Note that this threshold is above the generalization threshold in the baseline model (i.e., $\beta^*(\frac{1}{4})$). This implies that specialization by raters is more likely when information acquisition is costlier. The reason is that specializing in category A (the cheaper category) maximizes Rater 2's standalone value for a wider range of β than in the baseline model. Therefore, the shift to the generalization-like outcome occurs at a larger threshold.

For other parameter values, specialization in different categories is an equilibrium. In particular, $(\boldsymbol{\lambda}_1^{SP_B}, \boldsymbol{\lambda}_2^{SP_A})$ is always an equilibrium in these cases, while $(\boldsymbol{\lambda}_1^{SP_A}, \boldsymbol{\lambda}_2^{SP_B})$ is an equilibrium for a subset of these parameter values.[40] Overall, we observe that allowing for heterogeneities in information acquisition costs leads to similar equilibrium outcomes.

Both of these specialization outcomes are value-maximizing if $b_1 = b_2$. However, when $b_2 < b_1$, the unique value-maximizing outcome is $(\boldsymbol{\lambda}_1^{SP_B}, \boldsymbol{\lambda}_2^{SP_A})$; in this case, Rater 1 specializes in Category B and Rater 2 specializes in Category A. This is because Rater 1 has a comparative advantage in producing information about category B. Taken together, our results suggest that even if the value-maximizing outcome is an equilibrium, the equilibrium may not be unique. This implies that some degree of coordination may be necessary for its implementation.

Alternative Allocations of Pricing-power. Our baseline model assumes that Rater 1 always sets its fee first, which gives Rater 1 stronger pricing power in the allocation of the surplus created by the ratings. In Internet Appendix IA.E, we relax this assumption by allowing Rater 1 to set its fee first with probability $p \in [0.5, 1]$. From Lemma 1, we know that when the combined value of the rating technologies exceeds the sum of the marginal values, both raters charge their marginal values regardless of being the first or second mover. Probability p parameterizes the allocation of surplus between the raters when the combined value of the rating technologies is smaller than the sum of their marginal values.

[40] Both $(\boldsymbol{\lambda}_1^{SP_A}, \boldsymbol{\lambda}_2^{SP_B})$ and $(\boldsymbol{\lambda}_1^{SP_B}, \boldsymbol{\lambda}_2^{SP_A})$ are equilibria when $u^{HL} \geq 0$. When $u^{HL} < 0$, $(\boldsymbol{\lambda}_1^{SP_B}, \boldsymbol{\lambda}_2^{SP_A})$ is an equilibrium when $\beta \leq \beta^*(\frac{b_2\bar{\lambda}}{4})$, while $(\boldsymbol{\lambda}_1^{SP_A}, \boldsymbol{\lambda}_2^{SP_B})$ is an equilibrium when β is lower than a threshold that is less than or equal to $\beta \leq \beta^*(\frac{b_2\bar{\lambda}}{4})$. The threshold value is provided in Proposition IA.4.

We find that the key insights of the baseline model are robust to this alternative fee-setting mechanism. Generalization by both raters remains the unique market equilibrium when β is sufficiently large, except for the knife-edge case of $p = 0.5$. The intuition is that the rater with a lower market power (Rater 2) assigns a higher weight to the stand-alone value of its ratings, and specialization generates a small stand-alone value when β is large. Hence, this rater benefits from producing information in both categories to improve the stand-alone value of its ratings, and consequently, its pricing power.

There is a large benefit to achieving pricing power in the model; the rater can extract much more of the total surplus. This may explain the seemingly continuous consolidation in the ESG ratings industry (ESMA, 2021).

When $p = 0.5$, specialization in different categories is always an equilibrium outcome; for large values of β, generalization by both raters is also an equilibrium. This is because in the specialization outcome, if a rater deviates by producing some information about the other category, the gain from the larger stand-alone value exactly offsets the loss from the lower marginal value. This knife-edge scenario is unlikely to capture the distribution of pricing power in the marketplace, given, as discussed earlier, that major ESG rating agencies hold disproportionate influence over the market.

Sequential Design of Rating Technologies. If Rater 1 in our model represents an incumbent provider of ESG ratings, the rater may also have a first-mover advantage in the design of ratings in addition to its first-mover advantage in fee-setting. We consider a variation of the baseline model where the choice of rating technologies is sequential in Internet Appendix IA.F, and Rater 1 moves first. After both raters choose their rating technologies, they set fees sequentially.

We find that the results are qualitatively similar to the baseline scenario. When $u^{HL} < 0$ and $\beta > \beta^*(\frac{1}{4})$, the unique equilibrium is generalization by both raters, and when $\beta < \beta^*(\frac{1}{4})$ or $u^{HL} > 0$, the unique equilibrium is specialization in different categories. The only difference compared to the results in Proposition 1 is that generalization is no longer an equilibrium when $u^{HL} < 0$ and $\beta \in [\beta^*(\frac{9}{16}), \beta^*(\frac{17}{32})]$. This is because Rater 1 effectively uses its first mover advantage in rating technology choice to select its preferred equilibrium, which is the specialization outcome.

Simultaneous Fee-setting. The sequential fee-setting in our model can be viewed as a device for equilibrium selection, as discussed in Footnote 14. In Appendix IA.G, we show that the equilibria characterized in Proposition 1 and Lemma 1 constitute a subgame-perfect Nash equilibrium when fee-setting is simultaneous. Therefore, our results continue to hold under simultaneous fee-setting.

C. Broader Implications

Choice of Measurement Methodologies. As discussed in Section V.A, Berg et al. (2022) attribute 38% of ratings disagreement to raters examining different subcategories, which is along the lines of what we call specialization. They further attribute 56% of disagreement to differences in measurement.

One possible interpretation of this is the precision choice in our main model. However, another possibility is that raters use different methodologies for subcategories that they have in common.

In Internet Appendix IA.H, we analyze the raters' measurement decisions when there are multiple methods available to measure performance in a single subcategory. Specifically, we depart from the baseline model by assuming that the investor is only concerned about the project's performance in a single subcategory, say subcategory X. Subcategory X could represent a subcategory of the major categories E, S, or G. However, there are two noisy variables that can be used to measure performance in X. For instance, to measure a firm's performance in providing occupational safety, one can examine the number of employee injuries, as well as the number or frequency of safety training sessions. Both measures inform investors about the firm's commitment to creating a safe working environment for its employees, but neither of these measures is perfect.

We show that specialization in different methodologies maximizes welfare. Therefore, the discrepancy documented in the measurements across raters is efficient to the extent that it captures specialization in independent methods of measurement. Interestingly, there is some specialization in the market in terms of measurement methods: Reprisk focuses on firm risk incidents.[41] We also demonstrate that generalization is the unique equilibrium when β is large and the analogue of u^{HL} in this model is negative.

[41] A discussion of Reprisk's approach is at https://www.reprisk.com/approach.

Endogenous ESG Performance. Our baseline model takes the project's ESG performance as given. However, if ESG performance were endogenous and could be improved (i.e., a higher value of η chosen) at a cost, it is intuitive that the incentives to improve would be stronger when ESG ratings are more informative, since the effort is more likely to be rewarded with higher ratings. Therefore, the outcome where both raters specialize in different categories, which perfectly reveals the project's ESG performance, would provide the strongest incentives to improve the project's ESG performance. As a result, specialization in different categories would remain the unique value-maximizing outcome in this case.

However, the decision to specialize or generalize is endogenous and depends on the equilibrium level of η. In fact, the threshold for the generalization outcome to be the unique equilibrium when $u^{HL} < 0$, $\beta^*(\frac{1}{4})$, increases with η. This implies that the raters are more likely to specialize when they expect greater effort toward improving ESG performance. Taken together, these observations point to a strategic complementarity between the firm's decision to improve ESG performance and raters' decision to specialize.

Therefore, multiple equilibria may arise. In one possible equilibrium, there is greater effort towards improving ESG performance and the raters specialize in different categories. In another possible equilibrium, less effort is made to improve ESG performance and the raters generalize, resulting in lower-quality ESG ratings. This implies that policies encouraging ESG investments by firms may be further reinforced through their positive impact on the equilibrium quality of ESG ratings.

Economies of Scale in Information Production. Our baseline model abstracts away from potential economies of scale in acquiring information on multiple categories. For example, in practice, ESG raters may be able to learn about multiple categories by collecting and analyzing one set of documents or surveys. Such economies of scale makes generalization by both raters an equilibrium for a larger range of β. Specifically, suppose the cost of generalization is less than any other rating technology, including specialization in either category, by $S > 0$. Then, the threshold value at which the unique equilibrium outcome switches from specialization to generalization decreases from $\beta^*(\frac{1}{4})$ — the point where the stand-alone values of generalization and specialization are equal— to a lower threshold, which we denote by β^*_S.

Suppose S is small enough such that specialization in different categories remains the value-maximizing outcome; that is, $V(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{\lambda}^{SP_B}) - V(\boldsymbol{\lambda}^{GN}, \boldsymbol{\lambda}^{GN}) > 2S$. In this case, the presence of economies of scale reduces the welfare loss associated with generalization. However, the overall welfare implications are ambiguous because the raters inefficiently generalize when $\beta \in [\beta^*_S, \beta^*\left(\frac{1}{4}\right)]$. This occurs because Rater 2 switches to generalization at these values due to its lower cost, but fails to fully internalize the value lost from the transition.

When S is large, the generalization outcome becomes value-maximizing and the equilibrium outcome is, in fact, efficient when $\beta \geq \beta^*_S$. However, when $\beta < \beta^*_S$, the specialization outcome becomes inefficient.

Overall, in the presence of economies of scale, the raters are more likely to generalize. The above discussion demonstrates that the effect on total surplus is ambiguous.

VII. Conclusion

ESG investing has become a key focus in financial markets. Investors need information on ESG factors in order to allocate their capital. ESG raters provide this information, but not without controversy. The media, regulators, and academics have attacked raters for providing inaccurate ratings. We construct a model of the market for ESG ratings. In the model, raters provide information about multiple (unrelated) categories and set fees. Specializing in different categories can maximize surplus. However, the competitive outcome may be for raters to generalize among categories. We also demonstrate that specialization maximizes disagreement among raters, and, hence, disagreement may be a poor measure of welfare. Greenwashing by firms may push the raters towards generalizing.

Our analysis suggests that attempts to decrease the divergence of ratings may reduce the amount of information conveyed to financial markets. Instead, regulations can improve market efficiency by mandating or improving disclosure standards for firms, which would reduce the scope for greenwashing. Price caps can implement the value maximizing outcome, but possibly at the cost of reducing incentives to acquire information.

Appendix A. Proofs

A. *Proof of Lemma 1 (Equilibrium fees)*

The following definition is helpful for analyzing the raters' fee-setting behavior.

DEFINITION A1: *The ratings generated by rating technologies* $\boldsymbol{\lambda}_1 = (\lambda_1^A, \lambda_1^B)$ *and* $\boldsymbol{\lambda}_2 = (\lambda_2^A, \lambda_2^B)$ *are called "**complements**" if*

$$V(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) - V(\boldsymbol{\lambda}_1, \boldsymbol{O}) \geq V(\boldsymbol{O}, \boldsymbol{\lambda}_2) - V(\boldsymbol{O}, \boldsymbol{O}). \tag{A1}$$

*Otherwise, we call the rating technologies "**substitutes**."*

First, we consider the case that the rating technologies are complements. Suppose Rater 1 sets ϕ_1 above the stand-alone value of $\boldsymbol{\lambda}_1$, i.e., $\phi_1 > V(\boldsymbol{\lambda}_1, \boldsymbol{O}) - V(\boldsymbol{O}, \boldsymbol{O})$. Therefore, the investor prefers purchasing no ratings to purchasing only Rater 1's ratings. To successfully sell its ratings, Rater 2 should set ϕ_2 such that the investor purchases its ratings with Rater 1's ratings, i.e., $\phi_2 = V(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) - V(\boldsymbol{O}, \boldsymbol{O}) - \phi_1$, or without Rater 1's ratings, i.e., $\phi_2 = V(\boldsymbol{O}, \boldsymbol{\lambda}_2) - V(\boldsymbol{O}, \boldsymbol{O})$. Thus, Rater 2 sets $\phi_2 = \max\{V(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) - V(\boldsymbol{O}, \boldsymbol{O}) - \phi_1, V(\boldsymbol{O}, \boldsymbol{\lambda}_2) - V(\boldsymbol{O}, \boldsymbol{O})\}$. Both raters' ratings are purchased when the first element is larger or when the elements are equal (according to the tie-breaking rule). When the second element is strictly larger, only Rater 2's ratings are purchased. Therefore, the maximum fee that Rater 1 can collect is:

$$\begin{aligned}\phi_1 &= \{V(\boldsymbol{\lambda}_1,\boldsymbol{\lambda}_2) - V(\boldsymbol{O},\boldsymbol{O})\} - \{V(\boldsymbol{O},\boldsymbol{\lambda}_2) - V(\boldsymbol{O},\boldsymbol{O})\} \\ &= V(\boldsymbol{\lambda}_1,\boldsymbol{\lambda}_2) - V(\boldsymbol{O},\boldsymbol{\lambda}_2) \geq V(\boldsymbol{\lambda}_1,\boldsymbol{O}) - V(\boldsymbol{O},\boldsymbol{O}).\end{aligned} \tag{A2}$$

As a result, $\phi_2 = V(\boldsymbol{O},\boldsymbol{\lambda}_2) - V(\boldsymbol{O},\boldsymbol{O})$, and both sets of ratings are purchased. Furthermore, it is suboptimal for Rater 1 to set ϕ_1 below $V(\boldsymbol{\lambda}_1,\boldsymbol{O}) - V(\boldsymbol{O},\boldsymbol{O})$. It confirms the equilibrium fees for the case that the rating technologies are complements.

Now, we analyze the case that the rating technologies are substitutes. That is, according to Definition A1, the following inequality holds:

$$V(\boldsymbol{\lambda}_1,\boldsymbol{\lambda}_2) - V(\boldsymbol{O},\boldsymbol{\lambda}_2) < V(\boldsymbol{\lambda}_1,\boldsymbol{O}) - V(\boldsymbol{O},\boldsymbol{O}). \tag{A3}$$

Similar to the previous case, suppose Rater 1 moves first and sets ϕ_1. There are three cases depending on the value of ϕ_1:

- If $\phi_1 > V(\boldsymbol{\lambda}_1,\boldsymbol{O}) - V(\boldsymbol{O},\boldsymbol{O})$, then the fee is larger than the stand-alone value of $\boldsymbol{\lambda}_1$ and its marginal value given Rater 2's ratings. Therefore, Rater 1's ratings are not purchased regardless of ϕ_2.

- If $V(\boldsymbol{\lambda}_1,\boldsymbol{\lambda}_2) - V(\boldsymbol{O},\boldsymbol{\lambda}_2) < \phi_1 \leq V(\boldsymbol{\lambda}_1,\boldsymbol{O}) - V(\boldsymbol{O},\boldsymbol{O})$, then Rater 1's fee is less than or equal to the stand-alone value of its ratings and more than their marginal value given Rater 2's ratings. Therefore, the ratings of both raters are not purchased together. To sell its ratings, Rater 2 sets ϕ_2 slightly below the value that makes the investor indifferent between $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2$, i.e., $\phi_2 = V(\boldsymbol{O},\boldsymbol{\lambda}_2) + \phi_1 - V(\boldsymbol{\lambda}_1,\boldsymbol{O}) - \varepsilon$ (for a sufficiently small value of $\varepsilon > 0$). With this choice, only $\boldsymbol{\lambda}_2$ is purchased, and Rater 1's payoff is zero.

- If $\phi_1 \leq V(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) - V(\boldsymbol{O}, \boldsymbol{\lambda}_2)$, then ϕ_1 is less than or equal to the marginal value of Rater 1's ratings given $\boldsymbol{\lambda}_2$. Rater 2 sets its fee equal to the marginal value of $\boldsymbol{\lambda}_2$ given $\boldsymbol{\lambda}_1$, i.e., $\phi_2 = V(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) - V(\boldsymbol{\lambda}_1, \boldsymbol{O})$, and both sets of ratings are purchased.

Therefore, the maximum fee that Rater 1 can set to successfully sell its ratings is $\phi_1 = V(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) - V(\boldsymbol{O}, \boldsymbol{\lambda}_2)$. In this case, $\phi_2 = V(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) - V(\boldsymbol{\lambda}_1, \boldsymbol{O})$, and both sets of ratings are purchased. We see that in both possibilities, Rater 1 sets ϕ_1 equal to the marginal value of its ratings, i.e., $V(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) - V(\boldsymbol{O}, \boldsymbol{\lambda}_2)$. Rater 2's equilibrium fee is either $V(\boldsymbol{O}, \boldsymbol{\lambda}_2) - V(\boldsymbol{O}, \boldsymbol{O})$, or $V(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) - V(\boldsymbol{\lambda}_1, \boldsymbol{O})$, whichever is smaller.

B. *Proof of Proposition 1 (Characterization of market equilibria)*

To prove the proposition, we show that rating technologies $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2$ can constitute an equilibrium only if the raters specialize in different categories (i.e., $(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) \in \{(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{\lambda}^{SP_B}), (\boldsymbol{\lambda}^{SP_B}, \boldsymbol{\lambda}^{SP_A})\}$), or both raters generalize (i.e., $(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) = (\boldsymbol{\lambda}^{GN}, \boldsymbol{\lambda}^{GN})$). After demonstrating that these outcomes are the only possible forms of equilibria in **Steps 1 and 2** , we complete the characterization in **Step 3**.

First, we prove two helpful Lemmas. Lemma A1 simplifies the value function. Lemma A2 provides a sufficient condition for a pair to form an equilibrium.

LEMMA A1: *Let* λ^* *be the solution to the following linear equation:*

$$\eta(1-\lambda^*)(\Delta+\beta u^{HH})+(1-\eta)(\Delta+\beta u^{HL})=0. \tag{A4}$$

Then, the combined value is:

$$V(\boldsymbol{\lambda}_1,\boldsymbol{\lambda}_2)=\Big\{\lambda^A\lambda^B+\lambda^A[\lambda^*-\lambda^B]^+ +\lambda^B[\lambda^*-\lambda^A]^+\Big\}\eta^2(\Delta+\beta u^{HH}) \tag{A5}$$

,where $[x]^+=\max\{x,0\}$*, and*

$$\lambda^i=\lambda_1^i+\lambda_2^i-\lambda_1^i\lambda_2^i, \quad i=A,B. \tag{A6}$$

Proof. Note that according to Assumption 1, the investment does not take place when all ratings are low, i.e., $(s^A,s^B)=(l,l)$. Therefore, we can expand equation 9 as below

$$\begin{aligned} V(\boldsymbol{\lambda}_1,\boldsymbol{\lambda}_2) &= (\Delta+\beta u^{HH})\eta^2\lambda^A\lambda^B \\ &+\frac{[(\Delta+\beta u^{HH})\eta^2\lambda^A(1-\lambda^B)+(\Delta+\beta u^{HL})\eta(1-\eta)\lambda^A]^+}{P(s^A=h,s^B=l)}P(s^A=h,s^B=l) \\ &+\frac{[(\Delta+\beta u^{HH})\eta^2(1-\lambda^A)\lambda^B+(\Delta+\beta u^{HL})(1-\eta)\eta\lambda^B]^+}{P(s^A=l,s^B=h)}P(s^A=l,s^B=h) \\ &=\eta^2(\Delta+\beta u^{HH})\lambda^A\lambda^B+\eta\lambda^A[\eta(1-\lambda^B)(\Delta+\beta u^{HH})+(1-\eta)(\Delta+\beta u^{HL})]^+ \\ &+\eta\lambda^B[\eta(1-\lambda^A)(\Delta+\beta u^{HH})+(1-\eta)(\Delta+\beta u^{HL})]^+. \end{aligned} \tag{A7}$$

By employing (A4), we can rewrite (A7) as below:

$$\begin{aligned} V(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) &= \eta^2(\Delta + \beta u^{HH})\lambda^A\lambda^B + \eta\lambda^A[\eta(\Delta + \beta u^{HH})(\lambda^* - \lambda^B)]^+ \\ &\quad + \eta\lambda^B[\eta(\Delta + \beta u^{HH})(\lambda^* - \lambda^A)]^+ \\ &= \Big\{\lambda^A\lambda^B + \lambda^A[\lambda^* - \lambda^B]^+ + \lambda^B[\lambda^* - \lambda^A]^+\Big\}\eta^2(\Delta + \beta u^{HH}). \end{aligned} \tag{A8}$$

□

Lemma A1 states that the value function is proportionate to $v(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2; \lambda^*)$, where

$$\begin{aligned} v(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2; \lambda^*) &= \lambda^A\lambda^B + \lambda^A[\lambda^* - \lambda^B]^+ + \lambda^B[\lambda^* - \lambda^A]^+ \\ &= \max\{\lambda^A\lambda^B, \lambda^A\lambda^*, \lambda^B\lambda^*, (\lambda^A + \lambda^B)\lambda^* - \lambda^A\lambda^B\}. \end{aligned} \tag{A9}$$

Therefore, according to Lemma 1, we can write the fees as below:

$$\begin{aligned} \phi_j(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) &= \eta^2(\Delta + \beta u^{HH})\hat{\phi}_j(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2), \qquad j = 1, 2 \\ \hat{\phi}_1(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) &= v(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2; \lambda^*) - v(\boldsymbol{O}, \boldsymbol{\lambda}_2; \lambda^*) \\ \hat{\phi}_2(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) &= \min\{v(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2; \lambda^*) - v(\boldsymbol{O}, \boldsymbol{\lambda}_2; \lambda^*), v(\boldsymbol{O}, \boldsymbol{\lambda}_2; \lambda^*) - v(\boldsymbol{O}, \boldsymbol{O}; \lambda^*)\}. \end{aligned} \tag{A10}$$

A key insight we obtain from Lemma A1 is that λ^* is a sufficient variable to analyze the raters' best response behavior. Note that λ^* is bigger than one when $\Delta + \beta u^{HL} > 0$, and it is less than one otherwise. Moreover, note that the mapping between λ^* and β is the inverse of $\beta^*(\lambda)$, introduced in equation 21. That is, $\beta = \beta^*(\lambda^*)$. Based on this observation, we characterize the equilibria in terms of λ^*, and use the mapping to describe them in terms of β.

LEMMA A2: *For rating technologies* $\boldsymbol{\lambda}_1$,$\boldsymbol{\lambda}_2$, *and* $\tilde{\boldsymbol{\lambda}}_2$, *suppose the following inequalities hold:*

$$V(\boldsymbol{\lambda}_1, \tilde{\boldsymbol{\lambda}}_2) > V(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2), \qquad V(\boldsymbol{O}, \tilde{\boldsymbol{\lambda}}_2) > V(\boldsymbol{O}, \boldsymbol{\lambda}_2). \tag{A11}$$

Then, $\phi_2(\boldsymbol{\lambda}_1, \tilde{\boldsymbol{\lambda}}_2) > \phi_2(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)$.

Proof. The inequality below shows that Rater 2's payoff is larger with $\tilde{\boldsymbol{\lambda}}_2$ than with $\boldsymbol{\lambda}_2$, given $\boldsymbol{\lambda}_1$:

$$\begin{aligned} \phi_2(\boldsymbol{\lambda}_1, \tilde{\boldsymbol{\lambda}}_2) &= \min\{V(\boldsymbol{\lambda}_1, \tilde{\boldsymbol{\lambda}}_2) - V(\boldsymbol{\lambda}_1, \boldsymbol{O}), V(\boldsymbol{O}, \tilde{\boldsymbol{\lambda}}_2) - V(\boldsymbol{O}, \boldsymbol{O})\} \\ &> \min\{V(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) - V(\boldsymbol{\lambda}_1, \boldsymbol{O}), V(\boldsymbol{O}, \boldsymbol{\lambda}_2) - V(\boldsymbol{O}, \boldsymbol{O})\} = \phi_2(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2). \end{aligned} \tag{A12}$$

□

An implication of Lemma A2 is that to prove $\boldsymbol{\lambda}_2$ is a best response to $\boldsymbol{\lambda}_1$ for Rater 2, it is sufficient (but not necessary) to show that $\boldsymbol{\lambda}_2$ has the largest stand-alone value and combined value given $\boldsymbol{\lambda}_1$. Moreover, we can also employ Lemma A2 to prove that the technological constraint should bind for both raters. Corollary A1 formalizes this point.

COROLLARY A1: *If pair* $(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)$ *forms an equilibrium, then the technological constraint 8 binds for both raters, i.e.,* $\lambda_j^A + \lambda_j^B = 1$, $j = 1, 2$.

Proof. Consider the contrary that the constraint does not bind for Rater $j \in \{1, 2\}$. That is, $\lambda_j^A + \lambda_j^B < 1$. $\boldsymbol{\lambda}_j$ should have a positive marginal value, as otherwise, this would imply that $\boldsymbol{\lambda}_j$ earns a zero payoff for Rater j (See

Lemma 1).[42]

Since $\boldsymbol{\lambda}_j$ has a positive marginal value, one can show that either $\frac{\partial v(\boldsymbol{\lambda}_1,\boldsymbol{\lambda}_2)}{\partial \lambda_j^A} > 0$ or $\frac{\partial v(\boldsymbol{\lambda}_1,\boldsymbol{\lambda}_2)}{\partial \lambda_j^B} > 0$. Moreover, if the rater increases the precision of its ratings in both categories, then the stand-alone value strictly increases. Therefore, Rater j can increase both the marginal value and stand-alone value of its rating technology by switching to $\boldsymbol{\lambda}'_j = (\lambda_j^{A\prime}, \lambda_j^{B\prime})$, where $\lambda_j^{i\prime} > \lambda_j^i$, $i = A, B$. According to Lemma A2, $\boldsymbol{\lambda}'_j$ obtains a higher payoff, which is a contradiction. Therefore, the technological constraint should be binding for both raters in any equilibrium.

□

With this result, it is without loss of generality to focus on pairs for which the technological constraint 8 holds with equality for both raters. Now, we solve for the market equilibria in three steps.

Step 1: Possible equilibria when $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2$ are substitutes

Now, we examine which pairs of $(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)$ can form an equilibrium if the rating technologies are substitutes. In this case, both raters charge their marginal values as their fees, according to Lemma 1. Particularly, we show that the only possible equilibrium outcomes are specialization in different categories and generalization by both raters. We divide cases based on the value of λ^*, as defined in equation A4.

- **Case 1:** $\lambda^* \geq \max\{\lambda^A, \lambda^B\}$

[42]It is straightforward to show that when $\Delta + \beta u^{HL} \neq 0$ (Assumption 2), given any $\boldsymbol{\lambda}_{-j}$, there is a choice of $\boldsymbol{\lambda}_j$ that earns a positive payoff for Rater j.

Suppose Rater 2 does not specialize. We show that by specialization, the rater can increase both the stand-alone value and the marginal value of its ratings, which demonstrates the contradiction, according to Lemma A2.

According to Lemma A1, we have:

$$v(\boldsymbol{O}, \boldsymbol{\lambda}_2) = \lambda_2^A \lambda_2^B + \lambda_2^A [\lambda^* - \lambda_2^B]^+ + \lambda_2^B [\lambda^* - \lambda_2^A]^+ = \lambda^* - \lambda_2^A \lambda_2^B$$
$$v(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) = \lambda^A \lambda^B + \lambda^A [\lambda^* - \lambda^B]^+ + \lambda^B [\lambda^* - \lambda^A]^+ = (\lambda^A + \lambda^B)\lambda^* - \lambda^A \lambda^B. \tag{A13}$$

In the first line of (A13), we use the fact that $\lambda_2^i \leq \lambda^i$, $i = A, B$. Since $\lambda_2^A, \lambda_2^B > 0$, the stand-alone value of $\boldsymbol{\lambda}_2$ is less than that of specialization:

$$v(\boldsymbol{O}, \boldsymbol{\lambda}^{SP_B}) = v(\boldsymbol{O}, \boldsymbol{\lambda}^{SP_A}) = \lambda^*. \tag{A14}$$

Moreover, $v(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2')$ is quasi-convex in $\lambda_2^{A\prime}$, where $\boldsymbol{\lambda}_2' = (\lambda_2^{A\prime}, 1 - \lambda_2^{A\prime})$.[43] Therefore, the expression obtains its maximum value in one of the extreme values, i.e., $\lambda_2^A = 0, 1$. Suppose the maximum value is obtained at $\lambda_2^A = 1$, implying $v(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}^{SP_A}) > v(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)$.

Therefore, according to Lemma A2, Rater 2 can increase its payoff by switching to $\boldsymbol{\lambda}^{SP_A}$. The argument for the case that Rater 1 does not specialize is similar. As a result, the only possible equilibrium outcome

[43] In particular, suppose $\lambda^{i\prime} = \lambda_1^i + \lambda_2^{i\prime} - \lambda_1^i \lambda_2^{i\prime}$, for $i = A, B$. Then, $v(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2') = \max\{\lambda^{A\prime}\lambda^*, \lambda^{B\prime}\lambda^*, (\lambda^{A\prime} + \lambda^{B\prime})\lambda^* - \lambda^{A\prime}\lambda^{B\prime}\}$, because $\min\{\lambda^{A\prime}, \lambda^{B\prime}\} \leq \max\{\lambda^A, \lambda^B\} \leq \lambda^*$ for any $\boldsymbol{\lambda}_2'$. It is straightforward to show that the function $\max\{\lambda^{A\prime}\lambda^*, \lambda^{B\prime}\lambda^*, (\lambda^{A\prime} + \lambda^{B\prime})\lambda^* - \lambda^{A\prime}\lambda^{B\prime}\}$ is quasi-convex in $\lambda_2^{A\prime}$.

in this case is specialization in different categories.

- **Case 2:** $\max\{\lambda^A, \lambda^B\} > \lambda^* \geq \min\{\lambda^A, \lambda^B\}$

  Note that λ^* should be less than one since λ^A and λ^B are capped at one. Without loss, assume $\lambda^A \geq \lambda^B$. By applying equation A9, we find $v(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) = \lambda^A \lambda^*$. If no rater specializes, Rater 1 can increase the combined value, and consequently, its marginal value, by specializing in category A. If Rater 1 specializes in category A, then $v(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{\lambda}_2) = \lambda^*$ since $\lambda_2^B \leq \lambda^B \leq \lambda^*$, implying $\phi_2 = 0$. Therefore, Rater 2 can increase its payoff by specializing in category B since it obtains a strictly positive payoff. With a similar argument, we can rule out the possibility that only Rater 2 specializes.

- **Case 3:** $\min\{\lambda^A, \lambda^B\} > \lambda^*$

  Note that according to Definition A1, if $(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)$ are substitutes, the rating technologies are substitutes in an open neighborhood of the pair. Therefore, the first-order conditions are necessary for a pair to constitute an equilibrium.

  Moreover, in this case, Lemma A1 implies that $v(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) = \lambda^A \lambda^B$. Without loss, assume that $\lambda_2^A \geq \lambda_2^B$. The first-order conditions imply:

$$
\begin{aligned}
\frac{\partial v(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)}{\partial \lambda_1^A} \leq 0 \Rightarrow \lambda_2^A - \lambda_2^B \geq (\lambda_1^B - \lambda_1^A)(1 - \lambda_2^A)(1 - \lambda_2^B), \\
\frac{\partial v(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)}{\partial \lambda_2^A} \geq 0 \Rightarrow \lambda_1^B - \lambda_1^A \geq (\lambda_2^A - \lambda_2^B)(1 - \lambda_1^A)(1 - \lambda_1^B),
\end{aligned}
\tag{A15}
$$

where the equality holds for the first (second) inequality when the first (second) rater does not specialize.[44] It is straightforward to show that the only pairs that can satisfy these conditions are specialization in different categories and generalization by both raters.

Step 2: Possible equilibria when $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2$ are complements

The goal of this step is to demonstrate when $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2$ are complements, then they cannot constitute an equilibrium unless the raters specialize in different categories, or they both generalize.

First, we prove that for two rating technologies to be complements, we need to have $\lambda^A, \lambda^B > \lambda^*$. To this end, we rule out the other possibilities:

- $\lambda^* \geq \lambda^A, \lambda^B$: In this case, $\lambda_j^i \leq \lambda^*$, for $i = A, B$ and $j = 1, 2$. Therefore, according to equation A9, we have:

$$v(\boldsymbol{\lambda}_1, \boldsymbol{O}) = \lambda^* - \lambda_1^A \lambda_1^B, \qquad v(\boldsymbol{O}, \boldsymbol{\lambda}_2) = \lambda^* - \lambda_2^A \lambda_2^B$$
$$v(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) = (\lambda^A + \lambda^B)\lambda^* - \lambda^A \lambda^B = (2 - \lambda_1^A \lambda_2^A - \lambda_1^B \lambda_2^B)\lambda^* - \lambda^A \lambda^B. \tag{A16}$$

To prove that $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2$ are substitutes, we need to show:[45]

$$\begin{aligned} & v(\boldsymbol{\lambda}_1, \boldsymbol{O}) + v(\boldsymbol{O}, \boldsymbol{\lambda}_2) > v(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) \\ \iff & 2\lambda^* - \lambda_1^A \lambda_1^B - \lambda_2^A \lambda_2^B > (2 - \lambda_1^A \lambda_2^A - \lambda_1^B \lambda_2^B)\lambda^* - \lambda^A \lambda^B \\ \iff & \lambda^A \lambda^B - \lambda_1^A \lambda_1^B - \lambda_2^A \lambda_2^B > -\lambda^*(\lambda_1^A \lambda_2^A + \lambda_1^B \lambda_2^B). \end{aligned} \tag{A17}$$

---

[44]Note that the second inequality in (A15) implies that $\lambda_1^A \leq \lambda_1^B$. Therefore, the only feasible corner case in the first condition is that $\lambda_1^A = 0$.

[45]Note that $v(\boldsymbol{O}, \boldsymbol{O}) = 0$.

Since $\lambda^*(\lambda_1^A\lambda_2^A + \lambda_1^B\lambda_2^B)$ is non-negative, we only need to show that $\lambda^A\lambda^B - \lambda_1^A\lambda_1^B - \lambda_2^A\lambda_2^B > 0$, which is demonstrated by the inequalities below:

$$\begin{aligned} &\lambda^A\lambda^B - \lambda_1^A\lambda_1^B - \lambda_2^A\lambda_2^B \geq 0 \\ \iff &\{1 - \lambda^A - \lambda^B + \lambda^A\lambda^B\} - \{1 - \lambda^A - \lambda^B + \lambda_1^A\lambda_1^B + \lambda_2^A\lambda_2^B\} \geq 0 \\ \iff &\underbrace{(1-\lambda_1^A)(1-\lambda_1^B)(1-\lambda_2^A)(1-\lambda_2^B)}_{\geq 0} + \underbrace{1 - (\lambda_1^A + \lambda_2^B)(2 - \lambda_1^A - \lambda_2^B)}_{\geq 0} \geq 0. \end{aligned} \tag{A18}$$

The inequality in (A17) is strict; because in (A18), equality holds only when the raters specialize in the same category, which are clearly substitutes. This implies that any pair of rating technologies with $\lambda^* \geq \lambda^A, \lambda^B$ are substitutes.

- $\max\{\lambda^A, \lambda^B\} > \lambda^* \geq \min\{\lambda^A, \lambda^B\}$: Without loss of generality, assume that $\lambda^A > \lambda^B$. This implies that $\lambda_1^B, \lambda_2^B \leq \lambda^*$. Therefore:

$$v(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) = \lambda^A\lambda^* \leq \lambda_1^A\lambda^* + \lambda_2^A\lambda^* \leq v(\boldsymbol{\lambda}_1, \boldsymbol{O}) + v(\boldsymbol{O}, \boldsymbol{\lambda}_2). \tag{A19}$$

The first inequality is obtained from the definition of λ^A in equation 7. The second inequality is resulted from equation A9. In (A19), equality is never obtained. The reason is that equality is obtained in the first inequality only when either λ_1^A or λ_2^A is zero, which contradicts with $\lambda^B < \lambda^A \leq 1$.

Therefore, $\lambda^A, \lambda^B > \lambda^*$ if $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2$ are complements, implying that $v(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) = \lambda^A\lambda^B$. Now, we consider two possibilities for λ^*.

- **Case 1:** $\lambda^* \geq \min\{\lambda_2^A, \lambda_2^B\}$

  Without loss, suppose $\lambda_2^A \geq \lambda_2^B$. Note that Rater 1 charges its marginal contribution; thus $\boldsymbol{\lambda}_1$ maximizes $\lambda^A\lambda^B$, given $\boldsymbol{\lambda}_2$. The first-order condition for Rater 1 (characterized in (A15)) implies that either Rater 1 specializes in category B, or sets $\boldsymbol{\lambda}_1$ such that $\lambda_1^B - \lambda_1^A = \frac{\lambda_2^A - \lambda_2^B}{(1-\lambda_2^A)(1-\lambda_2^B)} \geq \lambda_2^A - \lambda_2^B$. If Rater 1 specializes, Rater 2 can increase the combined value and the stand-alone value of its rating technology by specializing in category A, which means that the specialization would obtain a higher payoff, according to Lemma A2. If Rater 1 generalizes, Rater 2 should also be generalizing, according to the first-order condition, as it implies that generalization is optimal for Rater 1 only when Rater 2 generalizes.

  If Rater 1 neither specializes nor generalizes, Rater 2 can increase its payoff by increasing λ_2^A:

$$
\begin{aligned}
\lambda_1^B - \lambda_1^A = \frac{\lambda_2^A - \lambda_2^B}{(1-\lambda_2^A)(1-\lambda_2^B)} > \lambda_2^A - \lambda_2^B \Rightarrow \lambda_1^B > \lambda_2^A \\
\Rightarrow \frac{\partial v}{\partial \lambda_2^A} = \frac{\partial \lambda^A \lambda^B}{\partial \lambda_2^A} = (1-\lambda_1^A)\lambda^B - (1-\lambda_1^B)\lambda^A \\
= \underbrace{\lambda_1^B}_{1-\lambda_1^A}(\lambda_1^B + \underbrace{\lambda_2^B\lambda_1^A}_{\lambda_2^B - \lambda_1^B\lambda_2^B}) - \underbrace{\lambda_1^A}_{1-\lambda_1^B}(\lambda_2^A + \underbrace{\lambda_2^B\lambda_1^A}_{\lambda_1^A - \lambda_1^A\lambda_2^A}) > 0 \\
\frac{\partial v(\boldsymbol{O}, \boldsymbol{\lambda}_2)}{\partial \lambda_2^A} = \frac{\partial \max\{\lambda_2^A\lambda^*, \lambda^* - \lambda_2^A\lambda_2^B\}}{\partial \lambda_2^A} > 0.
\end{aligned}
\tag{A20}
$$

  Therefore, according to Lemma A2, an interior λ_2 cannot be the best response to λ_1.

- **Case 2:** $\min\{\lambda_2^A, \lambda_2^B\} > \lambda^*$

Note that λ^* should be less than $\frac{1}{2}$, implying that $v(\boldsymbol{O}, \boldsymbol{\lambda}^{GN}) = \frac{1}{4}$. If Rater 2 generalizes, Rater 1's best response is also to generalize. Suppose $\boldsymbol{\lambda}_2 = (\lambda_2^A, \lambda_2^B) \neq (\frac{1}{2}, \frac{1}{2})$. If $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}^{GN}$ are complements, then Rater 2 can increase its payoff by generalizing as $\frac{1}{4} > \lambda_2^A \lambda_2^B$. If $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}^{GN}$ are substitutes, then we must have $\min\{\lambda_1^A, \lambda_1^B\} < \lambda^*$; otherwise, the inequality in (A18) would imply that they are complements. Assume that $\lambda_1^B > \lambda^* > \lambda_1^A$. Moreover, given that $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}^{GN}$ are substitutes, the inequality below shows that $\lambda^* > \frac{1}{4}$:

$$\begin{aligned} &(\frac{1}{2} + \lambda_1^B - \frac{1}{2}\lambda_1^B)(\frac{1}{2} + \lambda_1^A - \frac{1}{2}\lambda_1^A) < \frac{1}{4} + \lambda_1^B \lambda^* \\ &\Rightarrow \lambda_1^B < 1 + \lambda_1^B - \lambda_1^{B2} < 4\lambda_1^B \lambda^* \Rightarrow \frac{1}{4} < \lambda^*. \end{aligned} \tag{A21}$$

Note that

$$\hat{\phi}_2(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}^{SP_A}) = \min\{\lambda^*, \lambda_1^B(1 - \lambda^*)\}. \tag{A22}$$

Since $\hat{\phi}_2(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) = v(\boldsymbol{O}, \boldsymbol{\lambda}_2) < \frac{1}{4}$, we only need to show $\hat{\phi}_2(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}^{SP_A}) > \frac{1}{4}$ to demonstrate that $\boldsymbol{\lambda}_2$ is not the best response. It can be shown by noting that $\lambda^* \in (\frac{1}{4}, \frac{1}{2})$ and $\lambda_1^B \geq \frac{1}{2}$, implying $\lambda_1^B(1 - \lambda^*) > \frac{1}{4}$.

Step 3: Completing the characterization of the equilibria

Thus far, we have found that the possible equilibrium outcomes, for any value of λ^*, are:

$$(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{\lambda}^{SP_B}), (\boldsymbol{\lambda}^{SP_B}, \boldsymbol{\lambda}^{SP_A}), (\boldsymbol{\lambda}^{GN}, \boldsymbol{\lambda}^{GN}). \tag{A23}$$

We just need to examine which one of these outcomes forms an equilibrium for each value of λ^*.

- **Case 1:** $\lambda^* > \frac{9}{16}$

  This case covers part (a) of the proposition since $\lambda^* > 1$ when $u^{HL} \geq 0$, as well as part (b.2) when $\beta \leq \beta^*(\frac{9}{16})$ and $u^{HL} < 0$.

  We first show that $(\boldsymbol{\lambda}^{GN}, \boldsymbol{\lambda}^{GN})$ cannot be an equilibrium in this case. To see this, we compare Rater 1's payoff from generalization and specialization in response to generalization by Rater 2. According to Lemma 1, Rater 1 always obtains the marginal value of its ratings, i.e., $\phi_1 = V(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) - V(\boldsymbol{O}, \boldsymbol{\lambda}_2)$. Therefore, $\boldsymbol{\lambda}_1$ should maximize the combined value given $\boldsymbol{\lambda}_2$. The inequalities below show that $v(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{\lambda}^{GN}) > v(\boldsymbol{\lambda}^{GN}, \boldsymbol{\lambda}^{GN})$ when $\lambda^* > \frac{9}{16}$, which implies that the generalization outcome cannot be an equilibrium:

$$\begin{aligned} \lambda^* \geq 1: \quad & v(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{\lambda}^{GN}) = \frac{3}{2}\lambda^* - \frac{1}{2}, \quad v(\boldsymbol{\lambda}^{GN}, \boldsymbol{\lambda}^{GN}) = \frac{3}{2}\lambda^* - \frac{9}{16} \\ \lambda^* \in (\frac{3}{4}, 1): \quad & v(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{\lambda}^{GN}) = \lambda^*, \quad v(\boldsymbol{\lambda}^{GN}, \boldsymbol{\lambda}^{GN}) = \frac{3}{2}\lambda^* - \frac{9}{16} \\ \lambda^* \in (\frac{9}{16}, \frac{3}{4}): \quad & v(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{\lambda}^{GN}) = \lambda^*, \quad v(\boldsymbol{\lambda}^{GN}, \boldsymbol{\lambda}^{GN}) = \frac{9}{16} \\ \Rightarrow v(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{\lambda}^{GN}) > v(\boldsymbol{\lambda}^{GN}, & \boldsymbol{\lambda}^{GN}) \Rightarrow V(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{\lambda}^{GN}) > V(\boldsymbol{\lambda}^{GN}, \boldsymbol{\lambda}^{GN}). \end{aligned} \tag{A24}$$

  It is straightforward to show that any rating technology is a substitute for $\boldsymbol{\lambda}^{SP_A}$ and $\boldsymbol{\lambda}^{SP_B}$ when $\lambda^* > \frac{9}{16}$. Therefore, both raters charge their marginal values as their fees, according to Lemma 1. Note that maximizing the marginal value for a rater is equivalent to maximizing the combined value. In fact, specialization in different categories achieves the highest combined value, as the project's type would be perfectly revealed to the investor (we formally demonstrate this point in Proposition 2). Therefore, specialization in different categories is the only equilibrium outcome in pure strategies when $\lambda^* > \frac{9}{16}$.

- **Case 2:** $\lambda^* \in [\frac{17}{32}, \frac{9}{16}]$

  This case corresponds to part (b.2) in the proposition when $\beta \in [\beta^*(\frac{9}{16}), \beta^*(\frac{17}{32})]$.

  In this case, one can show that any two rating technologies are substitutes.[46] Therefore, both raters maximize the combined value of the ratings. We use this fact to prove that the best response to specialization (generalization) is specialization in the other category (generalization) for both raters.

  In response to specialization, the raters' best response is to specialize in the other category since specialization in different categories achieves the highest combined value.

  To see Rater 1's best response to generalization is generalization (the same logic applies to Rater 2), suppose $\boldsymbol{\lambda}_2 = \boldsymbol{\lambda}^{GN}$ and $\lambda^i = \lambda_1^i + \frac{1}{2} - \frac{1}{2}\lambda_1^i = \frac{1}{2} + \frac{1}{2}\lambda_1^i$, $i = A, B$. Note that $\max\{\lambda^A, \lambda^B\} \geq \frac{3}{4} > \lambda^*$. Therefore,

$$v(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}^{GN}) = \max\{\lambda^A\lambda^*, \lambda^B\lambda^*, \lambda^A\lambda^B\} \leq \frac{9}{16} = v(\boldsymbol{\lambda}^{GN}, \boldsymbol{\lambda}^{GN}). \quad \text{(A25)}$$

  Case 3: $\lambda^* \in (\frac{1}{4}, \frac{17}{32})$

[46] The only exception is $(\boldsymbol{\lambda}^{GN}, \boldsymbol{\lambda}^{GN})$ that are at the borderline of being complements when $\lambda^* = \frac{17}{32}$, while being substitutes when $\lambda^* > \frac{17}{32}$.

This case corresponds to part (b.2) in the proposition when $\beta \in (\beta^*(\frac{17}{32}), \beta^*(\frac{1}{4})]$.

Similar to the previous case, one can show that Rater 1's best response to specialization is specialization in the other category. However, Rater 2's fee can be the stand-alone value or the marginal value of its ratings depending on λ^*. It is straightforward to show that $\boldsymbol{\lambda}^{SP_A}$ and $\boldsymbol{\lambda}^{SP_B}$ are substitutes when $\lambda^* > \frac{1}{2}$ and complements when $\lambda^* \leq \frac{1}{2}$. When the two specialized rating technologies are substitutes, the best response to specialization is specialization in the other category. When the two specialized technologies are complements, then Rater 2 obtains the stand-alone value of its ratings, which is λ^*. When $\lambda^* \geq \frac{1}{4}$, specialization has the largest stand-alone value. Therefore, the optimality of specialization for Rater 2 in this case can be verified as below:

$$\begin{aligned}\phi_2(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{\lambda}_2) = \min\{V(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{\lambda}_2) - V(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{O}), V(\boldsymbol{O}, \boldsymbol{\lambda}_2)\} \leq V(\boldsymbol{O}, \boldsymbol{\lambda}_2) \\ \leq V(\boldsymbol{O}, \boldsymbol{\lambda}^{SP_B}) = \phi_2(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{\lambda}^{SP_B}).\end{aligned} \tag{A26}$$

When $\lambda^* < \frac{17}{32}$, $\boldsymbol{\lambda}^{GN}$ is a complement for itself. We use this fact to demonstrate that Rater 2's best response to generalization is not generalization. Consider the following possibilities:

- $\lambda^* \in (\frac{1}{2}, \frac{17}{32})$: Rater 2's payoff is the stand-alone value of its ratings in a sufficiently small neighborhood of $\boldsymbol{\lambda}_2 = \boldsymbol{\lambda}^{GN}$ since $2V(\boldsymbol{O}, \boldsymbol{\lambda}^{GN}) < V(\boldsymbol{\lambda}^{GN}, \boldsymbol{\lambda}^{GN})$. In this neighborhood $v(\boldsymbol{O}, \boldsymbol{\lambda}_2) = \lambda^* - \lambda_2^A \lambda_2^B$, implying that $\boldsymbol{\lambda}_2 = \boldsymbol{\lambda}^{GN}$ locally minimizes Rater 2's payoff. Therefore, the generalization outcome cannot be an equilibrium.

– $\lambda^* \in (\frac{1}{4}, \frac{1}{2}]$: In this case, $\hat{\phi}_2(\boldsymbol{\lambda}^{GN}, \boldsymbol{\lambda}^{GN}) = \frac{1}{4}$ since the pair $(\boldsymbol{\lambda}^{GN}, \boldsymbol{\lambda}^{GN})$ are complements. Now, consider $\boldsymbol{\lambda}_2$ such that $\lambda_2^A = \frac{1}{4\lambda^*} + \varepsilon < 1$, for a sufficiently small value of $\varepsilon > 0$. Note that:

$$v(\boldsymbol{O}, \boldsymbol{\lambda}_2) \geq \lambda_2^A \lambda^* > \frac{1}{4}$$
$$v(\boldsymbol{\lambda}^{GN}, \boldsymbol{\lambda}_2) - v(\boldsymbol{\lambda}^{GN}, \boldsymbol{O}) = (\frac{1}{2} + \frac{1}{2}\lambda_2^A)(\frac{1}{2} + \frac{1}{2}\lambda_2^B) - \frac{1}{4} = \frac{1}{4} + \frac{1}{4}\lambda_2^A\lambda_2^B > \frac{1}{4}. \tag{A27}$$

Therefore, $\phi_2(\boldsymbol{\lambda}^{GN}, \boldsymbol{\lambda}_2) > \phi_2(\boldsymbol{\lambda}^{GN}, \boldsymbol{\lambda}^{GN})$.

- **Case 4:** $\lambda^* < \frac{1}{4}$

  This case corresponds to part (b.1) of the proposition. In this case, one can show that any rating technology is a complement for $\boldsymbol{\lambda}^{GN}$. Moreover, $\boldsymbol{\lambda}^{SP_A}$ and $\boldsymbol{\lambda}^{SP_B}$ are complements. Therefore,

$$\hat{\phi}_2(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{\lambda}^{SP_B}) = \lambda^* < \frac{1}{4} = \hat{\phi}_2(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{\lambda}^{GN}). \tag{A28}$$

  As such, specialization in different categories is not an equilibrium. Generalization by both raters is an equilibrium since generalization has the highest stand-alone value, thus it maximizes Rater 2's payoff. Moreover, Rater 1's best response to generalization is generalization.

C. Proof of Proposition 2 (Value-maximizing ratings)

We show that the value function $V(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)$ attains its maximum only when

$$(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) \in \{(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{\lambda}^{SP_B}), (\boldsymbol{\lambda}^{SP_B}, \boldsymbol{\lambda}^{SP_A})\}. \tag{A29}$$

Note that specialization in different categories achieves the highest possible investment value since the project's type is fully revealed, so the investment always takes place efficiently. We show that no other pair of rating technologies can achieve the investment value created by these specialization outcomes.

For this purpose, we use equation A9. Recall that equation A9 defines function $v(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)$ that is proportionate to $V(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)$. We divide the cases into two:

$\underline{\Delta + \beta u^{HL} > 0}$: In this case, $\lambda^* > 1$, where λ^* is defined in equation A4. We find the following from equation A9:

$$v(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{\lambda}^{SP_B}) = v(\boldsymbol{\lambda}^{SP_B}, \boldsymbol{\lambda}^{SP_A}) = 2\lambda^* - 1. \tag{A30}$$

Since $\lambda^* > 1 \geq \lambda^A, \lambda^B$, we have:

$$v(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) = (\lambda^A + \lambda^B)\lambda^* - \lambda^A \lambda^B. \tag{A31}$$

Note that $\lambda^A, \lambda^B \in [0,1]$. Therefore,

$$\begin{aligned}
&(1-\lambda^A)(1-\lambda^B) \geq 0 \\
\Rightarrow\ & 1-\lambda^A\lambda^B \leq 2-\lambda^A-\lambda^B \leq (2-\lambda^A-\lambda^B)\lambda^* \\
\Rightarrow\ & \underbrace{(\lambda^A+\lambda^B)\lambda^* - \lambda^A\lambda^B}_{v(\boldsymbol{\lambda}_1,\boldsymbol{\lambda}_2)} \leq \underbrace{2\lambda^*-1}_{v(\boldsymbol{\lambda}^{SP_A},\boldsymbol{\lambda}^{SP_B})}
\end{aligned} \tag{A32}$$

Equality holds only when $\lambda^A = \lambda^B = 1$, corresponding to specialization in different categories.

$\underline{\Delta + \beta u^{HL} < 0}$: In this case, only projects with type (H,H) receive investment in the optimal investment outcome. Note that if the second or third term in (A7) are positive, a project with type (H,L) or (L,H) receives investment with a positive probability, which is inefficient. Therefore, for a value-maximizing pair, we should have $V = \lambda^A\lambda^B\eta^2(\Delta+\beta u^{HH})$. The maximum value is obtained when $\lambda^A = \lambda^B = 1$, which corresponds to specialization in different categories.

D. *Proof of Proposition 3 (Market equilibrium with greenwashing)*

Define u_M^{HH} and u_M^{HL} as below:

$$u_M^{HH} = \mathbb{E}[u|w_M^A = w_M^B = H] = \frac{\eta^2 u^{HH} + 2\eta(1-\eta)\alpha u^{HL} + (1-\eta)^2\alpha^2 u^{LL}}{\eta^2 + 2\eta(1-\eta)\alpha + (1-\eta)^2\alpha^2}$$

$$u_M^{HL} = \mathbb{E}[u|w_M^A = H, w_M^B = L] = \frac{\eta u^{HL} + (1-\eta)\alpha u^{LL}}{\eta + (1-\eta)\alpha}. \tag{A33}$$

According to the condition in equation 25, $\Delta + \beta u_M^{HH} > 0$. Note that Assumption 1 implies that $\Delta + \beta u_M^{HH}$ converges to a negative number as $\alpha \to 1$. Therefore, there exists $\bar{\alpha} \in (0,1)$ such that $\Delta + \beta u_M^{HH} = 0$ when $\alpha = \bar{\alpha}$.[47] Therefore, the condition in (25) implies that $\alpha < \bar{\alpha}$. It is also clear that $u_M^{HL} > u^{LL}$.

Therefore, we can use Proposition 1 for the preference parameters $(u^{LL}, u_M^{HL}, u_M^{HH})$ to characterize the equilibrium outcomes. According to Proposition 1, the unique equilibrium is generalization by both raters when $\beta > \beta^*(\frac{1}{4})$, which corresponds to $\lambda^* < \frac{1}{4}$. We have (See equation A4):

$$\lambda^* = 1 + \frac{(1-\eta)(\Delta + \beta u_M^{HL})}{\eta(\Delta + \beta u_M^{HH})}. \tag{A35}$$

Note that in the neighborhood of $\bar{\alpha}$, $u_M^{HL} < 0$. Therefore, λ^* converges to $-\infty$ as α goes to $\bar{\alpha}$ from the left (the denominator goes to zero and the numerator converges to a negative number), implying the existence of $\alpha^* \in (0, \bar{\alpha})$ such that $\lambda^* < \frac{1}{4}$ for all $\alpha \in (\alpha^*, \bar{\alpha})$.

[47] There is a unique $\bar{\alpha} > 0$ for which $\Delta + \beta u_M^{HH} = 0$ since u_M^{HH} is decreasing in α. This can be verified by noting that:

$$\begin{aligned} u_M^{HH} &= u^{LL} + \frac{\eta^2(u^{HH} - u^{LL}) + 2\eta(1-\eta)\alpha(u^{HL} - u^{LL})}{\eta^2 + 2\eta(1-\eta)\alpha + (1-\eta)^2\alpha^2} \\ &= u^{LL} + \frac{\eta^2(u^{HH} - u^{LL}) + 2\eta(1-\eta)\alpha(u^{HL} - u^{LL})}{\eta^2 + 2\eta(1-\eta)\alpha} \frac{\eta^2 + 2\eta(1-\eta)\alpha}{\eta^2 + 2\eta(1-\eta)\alpha + (1-\eta)^2\alpha^2} \\ &= u^{LL} + (u^{HL} - u^{LL} + \frac{\eta^2(u^{HH} - u^{HL})}{\eta^2 + 2\eta(1-\eta)\alpha})(1 - \{\frac{(1-\eta)\alpha}{\eta + (1-\eta)\alpha}\}^2). \end{aligned} \tag{A34}$$

The expressions inside the parentheses in the last line of (A34) is decreasing in α. Therefore, u_M^{HH}, as a function of α, crosses zero only once.

E. *Proof of Proposition 4 (Disagreement)*

When the raters specialize in different categories, they provide information about two independent categories. Therefore, their ratings are independent, and as a result, have a zero correlation, i.e., $Agg = 0$. However, when the raters specialize in the same category, they are perfectly revealing the project's type in that category. Therefore, their ratings are perfectly correlated, i.e., $Agg = 1$. Since $Agg \in [0, 1]$, we observe that these outcomes obtain the extreme possible values of Agg.

REFERENCES

Admati, Anat R, and Paul Pfleiderer, 1986, A monopolistic market for information, *Journal of Economic Theory* 39, 400–438.

Agrawal, Sonakshi, Lisa Yao Liu, Shiva Rajgopal, Suhas A Sridharan, Yifan Yan, and Teri Lombardi Yohn, 2024, ESG ratings of ESG index providers, *Available at SSRN 4468531* .

Amel-Zadeh, Amir, and George Serafeim, 2018, Why and how investors use ESG information: Evidence from a global survey, *Financial Analysts Journal* 74, 87–103.

Ardia, David, Keven Bluteau, Kris Boudt, and Koen Inghelbrecht, 2023, Climate change concerns and the performance of green vs. brown stocks, *Management Science* 69, 7607–7632.

Avramov, Doron, Si Cheng, Abraham Lioui, and Andrea Tarelli, 2022, Sustainable investing with ESG rating uncertainty, *Journal of Financial Economics* 145, 642–664.

Baker, Andrew C, David F Larcker, Charles G McClure, Durgesh Saraph, and Edward M Watts, 2024a, Diversity washing, *Journal of Accounting Research* 62, 1661–1709.

Baker, Malcolm P, Mark Egan, and Suproteem K Sarkar, 2024b, Demand for ESG, *Available at SSRN 4284023* .

Bar-Isaac, Heski, and Joel Shapiro, 2013, Ratings quality over the business cycle, *Journal of Financial Economics* 108, 62–78.

Berg, Florian, Julian F Kölbel, Anna Pavlova, and Roberto Rigobon, 2024, ESG confusion and stock returns: Tackling the problem of noise, *Available at SSRN 3941514* .

Berg, Florian, Julian F Kölbel, and Roberto Rigobon, 2022, Aggregate confusion: The divergence of ESG ratings, *Review of Finance* 26, 1315–1344.

Bergemann, Dirk, and Alessandro Bonatti, 2019, Markets for information: An introduction, *Annual Review of Economics* 11, 85–107.

Blickle, Kristian, Zhiguo He, Jing Huang, and Cecilia Parlatore, 2025, Information-based pricing in specialized lending, *Journal of Financial Economics* 172, 104135.

Bolton, Patrick, Xavier Freixas, and Joel Shapiro, 2012, The credit ratings game, *The Journal of Finance* 67, 85–111.

Bouvard, Matthieu, and Raphael Levy, 2018, Two-sided reputation in certification markets, *Management Science* 64, 4755–4774.

Chatterji, Aaron K, Rodolphe Durand, David I Levine, and Samuel Touboul, 2016, Do ratings of firms converge? implications for managers, investors and strategy researchers, *Strategic Management Journal* 37, 1597–1614.

Chen, Hui, and Jian Sun, 2024, Market for manipulable information, *Available at SSRN 4712430* .

Christensen, Dane M, George Serafeim, and Anywhere Sikochi, 2022, Why is corporate virtue in the eye of the beholder? the case of ESG ratings, *The Accounting Review* 97, 147–175.

Cornaggia, Jess, and Kimberly Cornaggia, 2025, ESG ratings management, *Available at SSRN 4520688* .

Döttling, Robin, and Magdalena Rola-Janicka, 2025, Too levered for Pigou: carbon pricing, financial constraints, and leverage regulation, *Journal of Financial Economics* 172, 104105.

Eccles, Robert G, Linda-Eling Lee, and Judith C Stroehle, 2020, The social origins of ESG: An analysis of innovest and KLD, *Organization & Environment* 33, 575–596.

ESMA, 2021, ESG ratings: Status and key issues ahead.

ESMA, 2022, Outcome of ESMA Call for Evidence on Market Characteristics of ESG Rating and Data Providers in the EU.

Gibson Brandon, Rajna, Philipp Krueger, and Peter Steffen Schmidt, 2021, ESG rating disagreement and stock returns, *Financial Analysts Journal* 77, 104–127.

Giglio, Stefano, Theresa Kuchler, Johannes Stroebel, and Xuran Zeng, 2025, Biodiversity risk, *Review of Finance* rfaf063.

Goldstein, Itay, Alexandr Kopytov, Lin Shen, and Haotian Xiang, 2025, On ESG investing: Heterogeneous preferences, information, and asset prices, *Available at SSRN 4080653* .

Goldstein, Itay, and Liyan Yang, 2015, Information diversity and complementarities in trading and information acquisition, *The Journal of Finance* 70, 1723–1765.

GSIA, 2022, The Global Sustainable Investment Review, November , 2022.

Gupta, Deeksha, Alexandr Kopytov, and Jan Starmans, 2022, The pace of change: Socially responsible investing in private markets, *Available at SSRN 3896511* .

Gupta, Deeksha, and Jan Starmans, 2024, Dynamic green disclosure requirements, *Available at SSRN 4557187* .

Hartzmark, Samuel M, and Abigail B Sussman, 2019, Do investors value sustainability? A natural experiment examining ranking and fund flows, *The Journal of Finance* 74, 2789–2837.

Hong, Harrison, Neng Wang, and Jinqiang Yang, 2023, Welfare consequences of sustainable finance, *The Review of Financial Studies* 36, 4864–4918.

Huang, Shiyang, and Alexandr Kopytov, 2023, Sustainable finance under regulation, *Available at SSRN 4231723* .

Huang, Shiyang, Liyan Yang, and Yan Xiong, 2018, Clientele, information sales, and asset prices, *Working Paper* .

Inderst, Roman, and Marcus M Opp, 2025, Sustainable finance versus environmental policy: Does greenwashing justify a taxonomy for sustainable investments?, *Journal of financial economics* 163, 103954.

Larcker, David F, Lukasz Pomorski, Brian Tayan, and Edward M Watts, 2022, ESG ratings: A compass without direction, *Rock Center for Corporate Governance at Stanford University Working Paper Forthcoming* .

Lovo, Stefano, and Jacques Olivier, 2025, Who should pay for ESG ratings?, *Review of Finance* rfaf060.

Oehmke, Martin, and Marcus M Opp, 2022, Green capital requirements, *Swedish House of Finance Research Paper* .

Oehmke, Martin, and Marcus M Opp, 2024, A theory of socially responsible investment, *Review of Economic Studies* rdae048.

Pástor, Luboš, Robert F Stambaugh, and Lucian A Taylor, 2021, Sustainable investing in equilibrium, *Journal of Financial Economics* 142, 550–571.

Pástor, L'uboš, Robert F Stambaugh, and Lucian A Taylor, 2022, Dissecting green returns, *Journal of Financial Economics* 146, 403–424.

Piatti, Ilaria, Joel Andrew Shapiro, and Xuan Wang, 2023, *Sustainable investing and public goods provision* (School of Economics and Finance, Queen Mary University of London).

Piccolo, Alessio, 2021, Credit ratings and competition, *Available at SSRN 3233964* .

Piccolo, Alessio, Jan Schneemeier, and Michele Bisceglia, 2022, Externalities of responsible investments, *Available at SSRN 4183855* .

Raghunandan, Aneesh, and Shiva Rajgopal, 2022, Do ESG funds make stakeholder-friendly investments?, *Review of Accounting Studies* 27, 822–863.

Sangiorgi, Francesco, and Chester Spatt, 2017, Opacity, credit rating shopping, and bias, *Management Science* 63, 4016–4036.

Sauzet, Maxime, and Olivier David Zerbib, 2022, When green investors are green consumers, in *Proceedings of the EUROFIDAI-ESSEC Paris December Finance Meeting*.

SustainAbility, 2020, Rate the raters 2020: Investor Survey and Interview Results.

SustainAbility, 2023, Rate the Raters 2023: ESG Ratings at a Crossroads.

Internet Appendix for
"The Market for ESG Ratings"

Ehsna Azarmsa and Joel Shapiro[1]

ABSTRACT

This Internet Appendix contains supplementary results demonstrating that the insights delivered by our model are robust to the key assumptions and modeling choices of the baseline model. The extensions analyzed are listed below, and the proofs are provided in Section IA.L. The timing of events in all sections is similar to that in the baseline model, unless stated otherwise.

- Section IA.A considers a setting with heterogeneous investors who differ in their ESG preferences.
- Section IA.B considers the case in which the investor values performance in the two categories differently.
- Section IA.C considers heterogeneity in access to ESG raters, resulting in market segmentation, and analyzes the equilibrium outcomes when the raters set their fees simultaneously.
- Section IA.D relaxes the symmetry between the two raters and the two categories by allowing for heterogeneous technological capacities

[1]Citation format: Azarmsa, Ehsan, and Joel Shapiro, Internet Appendix for "The Market for ESG Ratings," *Journal of Finance* [DOI String]. Please note: Wiley-Blackwell is not responsible for the content or functionality of any additional information provided by the authors. Any queries (other than missing material) should be directed to the authors of the article.

across raters and heterogeneous information acquisition costs across categories.

- Section IA.E considers an extension of the baseline model that allows the second rater to move first with a positive probability.
- Section IA.F considers the case in which raters choose their rating technologies sequentially rather than simultaneously.
- Section IA.G considers the case in which raters set fees simultaneously rather than sequentially.
- Section IA.H analyzes the raters' equilibrium behavior when they choose among different methods of measuring performance in a subcategory.
- Section IA.I examines the equilibria in mixed strategies.
- Section IA.J considers a more general signal structure for the ratings.

IA.A. Heterogeneous ESG Preferences

Consider the following departure from the baseline model presented in Section I. Suppose there are two types of investors: type-A and type-B investors. Type-i investors, $i = A, B$, are relatively more concerned about the project's low performance in category i (their primary category; we call the other category their secondary category). The net investment payoff for type-i investors is $\Delta + \beta u_i$, where $u_i \in \{u_i^{LL}, u_i^{LH}, u_i^{HL}, u_i^{HH}\}$, and the superscripts represent the project's type (w^A, w^B). To simplify the exposition, suppose type-A and type-B investors have equal mass, where the overall population is normalized to one. Furthermore, they only differ in the values of u_i^{LH} and u_i^{HL}, i.e., $u_A^{HH} = u_B^{HH} = u^{HH}$, and $u_A^{LL} = u_B^{LL} = u^{LL}$. In particular, let

$$u^{LL} < u_A^{LH} = u_B^{HL} < u_A^{HL} = u_B^{LH} < -\frac{\Delta}{\beta} < 0 < u^{HH}. \qquad \text{(IA.1)}$$

Note that Assumption 2 follows from the parametric restriction above. Furthermore, we assume that the project can be scaled flexibly depending on the investors' demand.

To sharpen the economic mechanism, we assume that type-i investors are so averse to low performance in their primary category, $i = A, B$, so that they do not invest without receiving a positive information about category i. Assumption IA.1 formalizes this. Note that Assumption IA.1 implies Assumption 1 for both investor groups.

ASSUMPTION IA.1:

$$\Delta + \beta\mathbb{E}[u_i|w^{-i} = H] < 0, \qquad i = A, B. \qquad \text{(IA.2)}$$

Similar to the baseline case, the raters simultaneously design rating technologies $\boldsymbol{\lambda}_1 = (\lambda_1^A, \lambda_1^B)$ and $\boldsymbol{\lambda}_2 = (\lambda_2^A, \lambda_2^B)$, and then set fees sequentially. The raters cannot set different fees for type-A and type-B investors. All investors then decide which ratings to adopt. The time-line is similar to Figure 2, except for the presence of multiple investor groups. Proposition IA.1 describes when the specialization and generalization outcomes are equilibriua.

PROPOSITION IA.1: *Define:*

$$\lambda^* = 1 + \frac{(1-\eta)(\Delta + \beta u_A^{HL})}{\eta(\Delta + \beta u^{HH})} = 1 + \frac{(1-\eta)(\Delta + \beta u_B^{LH})}{\eta(\Delta + \beta u^{HH})}. \tag{IA.3}$$

Given Assumption IA.1 and the parametric restriction specified in (IA.1):
a) *If $\lambda^* < \frac{1}{4}$, generalization by both raters is the unique equilibrium outcome. In this case, both investor groups buy ratings from both raters.*
b) *If $\lambda^* \geq \frac{1}{4}$, specialization in different categories is an equilibrium, while the generalization outcome is not, except when $\lambda^* \in \{\frac{1}{4}, \frac{9}{16}\}$. If $\lambda^* \in [\frac{1}{4}, \frac{2}{3}]$, all investors buy ratings from both raters. If $\lambda^* > \frac{2}{3}$, investors only purchase ratings from the rater specializing in their primary category.*[2]

Part (a) in Proposition IA.1 states that when all investors are sufficiently concerned about the project's low performance in both categories, the unique equilibrium is generalization by both raters, in line with our baseline result. To understand this result, note that Rater 2's fee is still capped by the stand-alone value of its ratings for the two investor types (in particular, whichever is larger), and generalization achieves the highest stand-alone value for both investor types. To prove the uniqueness, we show that generalization is Rater 2's best response to any choice of $\boldsymbol{\lambda}_1$; because, Rater 2 can indeed collect its stand-alone value for any $\boldsymbol{\lambda}_1$. The equilibrium uniqueness for $\lambda^* < \frac{1}{4}$ implies that the outcome where the raters specialize in different categories (namely, the specialization outcome) is not an equilibrium. This is because all investors are so concerned about the project's low performance in both

[2]Note that for this case, we do not discuss the uniqueness of the equilibrium due to its complexity.

categories that specialization yields little stand-alone value for Rater 2.

Part (b) of Proposition IA.1 reveals that when $\lambda^* \geq \frac{1}{4}$, investors assign a higher value to a rating technology specializing in their primary category than a generalized rating technology. The specialization outcome is thus an equilibrium. Depending on the importance of the secondary category to investors, captured by the values of u_A^{HL} and u_B^{LH}, and correspondingly λ^*, investors may buy ratings from both raters or only one. For intermediate values of u_A^{HL} and u_B^{LH}, investors purchase ratings from both raters. However, if u_A^{HL} and u_B^{LH} are sufficiently large, type-i investors have little willingness to pay for information about category $-i$. As a result, fees are set such that type-i investors purchase ratings only from the rater that specializes in their primary category i, indicating a segmented market outcome.[3]

Overall, Proposition IA.1 demonstrates that the key insights from the baseline model remain robust when accounting for heterogeneity in investor preferences. In fact, as long as investors sufficiently value information in all categories, the equilibrium behavior of the raters is similar to the baseline setup.

[3] Generalization by both raters is also an equilibrium when $\lambda^* = \frac{9}{16}$. This is primarily due to the strategic complementarity motive discussed for the intermediate region in Part (b.2) of Proposition 1.

IA.B. Asymmetric Preferences Over Categories

Our baseline model considers symmetric preferences across the two categories. Even in Appendix IA.A with multiple investor types, while each investor has asymmetric preferences, the aggregate preferences remain symmetric. This section relaxes this symmetry, and analyzes how it impacts the raters' equilibrium behavior.

Specifically, building on the baseline model with a single investor, suppose $u = u^{LH}$ when the project's type is (L, H), and $u = u^{HL}$ continues to represent the ESG performance when the project's type is (H, L). u^{HH} and u^{LL} are defined similarly to the baseline framework. Without loss of generality, let $u^{LL} < u^{LH} \leq u^{HL}$, implying that the investor is more concerned about low performance in category A than category B.

We continue to assume that the investor is sufficiently averse to the project's negative performance by imposing Assumption 1. Additionally, we focus on the case where $u^{LH} \leq u^{HL} < -\frac{\Delta}{\beta}$, which implies that low performance in any category leads to a negative net investment payoff, as the other scenarios are straightforward.[4] Proposition IA.2 characterizes the conditions under which specialization or generalization may emerge in equilibrium.[5]

[4] If $-\frac{\Delta}{\beta} < u^{LH} \leq u^{HL}$, then the unique equilibrium is specialization in different categories, and it is also the value-maximizing outcome. If $u^{LH} < -\frac{\Delta}{\beta} \leq u^{HL}$, the marginal value of information about category B is zero when the performance in category A is revealed. Therefore, in equilibrium, Rater 1 sets $\boldsymbol{\lambda}_1 = \boldsymbol{\lambda}^{SP_A}$, and Rater 2 generates no marginal value and charges a zero fee for any choice of $\boldsymbol{\lambda}_2$.

[5] The proposition does not characterize all possible equilibria due to its complexity.

PROPOSITION IA.2: *Define:*

$$\lambda_A^* = 1 + \frac{(1-\eta)(\Delta + \beta u^{LH})}{\eta(\Delta + \beta u^{HH})}, \qquad \lambda_B^* = 1 + \frac{(1-\eta)(\Delta + \beta u^{HL})}{\eta(\Delta + \beta u^{HH})} \tag{IA.4}$$

a) *If $\frac{1}{4} > \lambda_A^*, \lambda_B^*$, then generalization by both raters is the unique equilibrium outcome.*

b) *If $\lambda_B^* \geq \frac{1}{4}$, then $(\boldsymbol{\lambda}^{SP_B}, \boldsymbol{\lambda}^{SP_A})$ is an equilibrium. $(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{\lambda}^{SP_B})$ is also an equilibrium iff $\lambda_A^* \geq \sqrt{\lambda_B^*} - \lambda_B^*$. If $\lambda_A^* < \sqrt{\lambda_B^*} - \lambda_B^*$ and $\lambda_B^* \geq (\frac{\sqrt{5}-1}{2})^2 \simeq 0.38$, then there is an equilibrium in which Rater 1 chooses $\boldsymbol{\lambda}^{SP_A}$ and Rater 2 chooses an interior rating technology (i.e., does not specialize).*

Part (a) in Proposition IA.2 states that when the investor is sufficiently concerned about low performance in both categories, generalization is the unique equilibrium. This demonstrates that the generalization outcome being the unique equilibrium for large values of β is robust to asymmetries in the preferences across the two categories.

Part (b) describes the raters' behavior when the investor is less concerned about low performance in category B, reflected by a larger value of λ_B^* (Note that there is a one-to-one mapping between λ_B^* and u^{HL}). When u^{LH}, and consequently λ_A^*, is large, then both specialization outcomes constitute an equilibrium. However, as the investor becomes more concerned about low performance in category A, the stand-alone value of $\boldsymbol{\lambda}^{SP_B}$ declines. As a result, when u^{LH} falls below a certain threshold, $(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{\lambda}^{SP_B})$ is no longer an equilibrium, as $\boldsymbol{\lambda}^{SP_B}$ is not rater 2's best response to rater 1 specializing in category A. Instead, Rater 2's best response is to choose an interior rating technology. This outcome constitutes an equilibrium when u^{HL} is sufficiently large.

IA.C. Heterogeneous Preferences over Raters

Our baseline model assumes that the raters compete in the same market; there is one investor who can purchase ratings from both raters. This section relaxes that assumption by having multiple investors who have different preferences over raters. Specifically, a mass of infinitesimal investors with measure Z_1 can only purchase ratings from Rater 1; a mass with measure Z_2 can only purchase ratings from Rater 2; and a mass with measure Z_{12} can purchase ratings from both raters. Similar to Section IA.A, we assume that the project can be flexibly scaled with demand. Furthermore, suppose $Z_1 = Z_2$ to simplify the analysis.

As such, this section examines the case where raters have some market power due to the presence of captive investors. The timing is the same as in the baseline model.[6] Proposition IA.3 describes when specialization in different categories and generalization by both raters are equilibria for different values of u^{HL} and β.

PROPOSITION IA.3: *Let λ^* be such that $\beta = \beta^*(\lambda^*)$. Given Assumptions 1 and 2, we have:*

a) *If $u^{HL} \geq 0$, then specialization in different categories is an equilibrium, whereas generalization by both raters is not.*

[6] As in the baseline model, we continue to assume sequential fee-setting, which helps avoid the emergence of multiple equilibrium fees. See Footnote 14 for further discussion.

Under the specialization outcome, both raters charge their marginal values and serve both their captive investors and the shared group if $Z_{12} > \frac{1}{\lambda^*-1} Z_1$*, (when* $u^{HL} \geq 0$*,* λ^* *is always greater than one). Otherwise, Rater 2 sets its fee equal to its stand-alone value and only sells to its captive investors. Rater 1 sells to both its captive investors and the shared group, and sets its fee such that its payoff equates with Rater 2's payoff.*

b) *Suppose* $u^{HL} < 0$*.*

b.1) *Generalization by both raters is an equilibrium if and only if* $\beta \geq \beta^*(\frac{1}{4})$*. Moreover, Rater 2's fee is equal to its stand-alone value and it serves both its captive investors and the shared group. If* $Z_{12} \geq \frac{5}{4} Z_1$*, then Rater 1 sets its fee equal to its marginal value and serves only the shared group. Otherwise, Rater 1 also charges its stand-alone value and serves both its captive group and the shared group.*

b.2) *Specialization in different categories is an equilibrium if and only if* $\beta \leq \beta^*(\frac{1}{4})$*. The equilibrium fees are as follows:*

- *If* $\lambda^* \in [\frac{1}{4}, \frac{1}{2})$*, then Rater 2's fee is equal to its stand-alone value and it serves both its captive group and the shared group. Rater 1 does the same if* $(\frac{1}{\lambda^*} - 2) Z_{12} < Z_1$*. Otherwise, Rater 1 sets its fee equal to its marginal value and serves only the shared group.*

- *Suppose* $\lambda^* \in [\frac{1}{2}, 1]$*. If* $(1 - \lambda^*) Z_{12} > (2\lambda^* - 1) Z_1$*, then both raters charge their marginal values and serve both their captive groups and the shared group. Otherwise, Rater 2 charges its stand-alone value and serves only its captive investors. Rater 1 serves both its captive group and the shared group, and sets* ϕ_1 *such that its payoff equals Rater 2's* *payoff.*

Part (a) in Proposition IA.3 states that specialization in different categories is an equilibrium when $u^{HL} \geq 0$. This result generalizes Part (a) of Proposition 1 to a setting with captive investors. Furthermore, when the captive groups are sufficiently small, raters charge their marginal values, as in our baseline result. Otherwise, the market is segmented: Rater 2 serves only its captive investors at its stand-alone value, while Rater 1 serves both its captive investors and the shared group, but sets a low price to deter Rater 2 from entering the shared segment.

Part (b) of the proposition discusses the equilibrium outcomes when $u^{HL} < 0$. Consistent with our baseline result, Part (b.1) demonstrates that generalization by both raters constitutes an equilibrium when β is sufficiently large (i.e., $\beta \geq \beta^*(\frac{1}{4})$). When the shared segment is large relative to the captive groups, Rater 2 charges its stand-alone value and Rater 1 charges its marginal value, as in the baseline model. However, since the marginal value exceeds the stand-alone value in this case, the equilibrium fees price out Rater 1's captive group. When the captive groups are sufficiently large, both raters charge their stand-alone values and serve both their captive and the shared group.

Generalization by both raters is not an equilibrium when $\beta < \beta^*(\frac{1}{4})$. Recall from Proposition 1 that generalization is an equilibrium in the intermediate region $\beta \in [\beta^*(\frac{9}{16}), \beta^*(\frac{17}{32})]$ when there are no captive groups, i.e., $Z_1 = Z_2 = 0$. As a result, an interesting implication of Proposition IA.3 is that this equilibrium in the intermediate region is not robust to the presence of captive investors. The intuition is that the marginal value of generalization is zero if the other rater specializes, i.e., $V(\boldsymbol{\lambda}^{GN}, \boldsymbol{\lambda}^{SP_B}) = V(\boldsymbol{O}, \boldsymbol{\lambda}^{SP_B})$. In the baseline model, under the generalization outcome, no rater finds it beneficial to deviate to specialization, as it lowers both raters' payoffs. However, in the presence of captive investors, deviating to specialization forces the other rater to serve only its captive group, as the other rater is generalizing, which has no marginal value. This allows the deviating rater to set a greater fee for its captive and the shared group.

Part (b.2) shows that specialization in different categories is an equilibrium only when $\beta < \beta^*(\frac{1}{4})$, given $u^{HL} < 0$. When the captive groups are large, the equilibrium market structure depends on whether the stand-alone value of specialization exceeds its marginal value, given that the other rater specializes in the other category.

If the marginal value exceeds the stand-alone value, then Rater 1 sets its fee equal to its marginal value if the size of its captive group is sufficiently small, as this fee exceeds the captive investors' willingness to pay. Otherwise, Rater 1's fee equals its stand-alone value. Regardless of the size of its captive group, Rater 2's fee equals its stand-alone value, and it serves both its captive group and the shared group.

Conversely, if the stand-alone value exceeds the marginal value, then both raters charge their marginal values when the captive groups are sufficiently small. Otherwise, the market becomes segmented: Rater 2 serves only its captive group, while Rater 1 serves both its own captive group and the shared group. However, Rater 1 sets a sufficiently low fee so that Rater 2 has no incentive to enter the shared segment.

IA.D. Unequal Information Acquisition Costs Across Raters and Categories

In this section, we consider the case where the raters face different costs of information acquisition. Specifically,

$$\lambda_1^A + \frac{\lambda_1^B}{b_1} \leq 1, \qquad \lambda_2^A + \frac{\lambda_2^B}{b_2} \leq \bar{\lambda}, \qquad \bar{\lambda} \leq 1, \;\; b_2 \leq b_1 < 1. \tag{IA.5}$$

With these technological constraints, the specialized rating technologies are denoted by $\boldsymbol{\lambda}_j^{SP_i}$, $j = 1, 2$ and $i = A, B$, where:

$$\boldsymbol{\lambda}_1^{SP_A} \equiv (1, 0), \quad \boldsymbol{\lambda}_1^{SP_B} \equiv (0, b_1), \quad \boldsymbol{\lambda}_2^{SP_A} \equiv (\bar{\lambda}, 0), \text{ and } \quad \boldsymbol{\lambda}_2^{SP_B} \equiv (0, b_2\bar{\lambda}). \tag{IA.6}$$

In line with Assumption 2, we assume that the parameter values are such that specialization in either category has a positive marginal value for the investor when the othe rater specializes in the other category. Assumption 2′ formalizes this point.

ASSUMPTION 2′:

$$\begin{aligned} &V(\boldsymbol{\lambda}_1^{SP_A}, \boldsymbol{\lambda}_2^{SP_B}) > V(\boldsymbol{O}, \boldsymbol{\lambda}_2^{SP_B}), \qquad V(\boldsymbol{\lambda}_1^{SP_B}, \boldsymbol{\lambda}_2^{SP_A}) > V(\boldsymbol{O}, \boldsymbol{\lambda}_2^{SP_A}), \\ &V(\boldsymbol{\lambda}_1^{SP_A}, \boldsymbol{\lambda}_2^{SP_B}) > V(\boldsymbol{\lambda}_1^{SP_A}, \boldsymbol{O}), \qquad V(\boldsymbol{\lambda}_1^{SP_B}, \boldsymbol{\lambda}_2^{SP_A}) > V(\boldsymbol{\lambda}_1^{SP_B}, \boldsymbol{O}). \end{aligned} \tag{IA.7}$$

In Proposition IA.4, we see that the raters choose interior rating technologies in equilibrium when $u^{HL} < 0$ and β is sufficiently large, mirroring the generalization equilibrium outcome in our baseline model. The proposition also characterizes when the specialization outcomes $(\boldsymbol{\lambda}_1^{SP_A}, \boldsymbol{\lambda}_2^{SP_B})$ and

$(\boldsymbol{\lambda}_1^{SP_B}, \boldsymbol{\lambda}_2^{SP_A})$ are equilibria.[7]

PROPOSITION IA.4: *Suppose Assumptions 1 and 2′ hold.*

a) *If $u^{HL} \geq 0$, both $(\boldsymbol{\lambda}_1^{SP_A}, \boldsymbol{\lambda}_2^{SP_B})$ and $(\boldsymbol{\lambda}_1^{SP_B}, \boldsymbol{\lambda}_2^{SP_A})$ are equilibria.*

b) *If $u^{HL} < 0$:*

b.1) *When $\beta > \beta^*(\frac{b_2\bar{\lambda}}{4})$, there is a unique equilibrium outcome in which Rater 2 sets $\boldsymbol{\lambda}_2 = (\frac{\bar{\lambda}}{2}, \frac{b_2\bar{\lambda}}{2})$ and Rater 1 does not specialize.*

b.2) *When $\beta \leq \beta^*(\frac{b_2\bar{\lambda}}{4})$, $(\boldsymbol{\lambda}_1^{SP_B}, \boldsymbol{\lambda}_2^{SP_A})$ is always an equilibrium. $(\boldsymbol{\lambda}_1^{SP_A}, \boldsymbol{\lambda}_2^{SP_B})$ is also an equilibrium unless when the parameter values are such that:*[8]

$$\beta^*(\frac{b_2\bar{\lambda}}{4}) > \beta > \beta^*(\frac{b_2}{(1+b_2\bar{\lambda})^2} - \frac{b_2(1-\bar{\lambda})}{1+b_2\bar{\lambda}}). \tag{IA.8}$$

An implication of Proposition IA.4 (Part b.2) is that due to the heterogeneity in the information acquisition costs, the indeterminacy between the two specialization outcomes could be broken. In particular, when $u^{HL} < 0$, $(\boldsymbol{\lambda}_1^{SP_A}, \boldsymbol{\lambda}_2^{SP_B})$ is not an equilibrium for values of β that satisfy the condition in equation IA.8, while the other specialization outcome is an equilibrium. The intuition is that for Rater 2, $\boldsymbol{\lambda}_2^{SP_B}$ has a lower stand-alone value than $\boldsymbol{\lambda}_2^{SP_A}$. Therefore, specialization in category B by Rater 2 is less likely to be supported in an equilibrium.

Now, we discuss the value-maximizing outcomes in Proposition IA.5.

PROPOSITION IA.5: *Under Assumptions 1 and 2′, if $b_2 < b_1$, then the unique value-maximizing pair is $(\boldsymbol{\lambda}_1^{SP_B}, \boldsymbol{\lambda}_2^{SP_A})$. When $b_1 = b_2$, the value-*

[7]We continue to assume that fee-setting is sequential. As in the baseline model, this assumption helps rule out multiple equilibrium fees. See Footnote 14 for further discussion.

[8]This range of β does not exist for all values of b_1, b_2, and $\bar{\lambda}$.

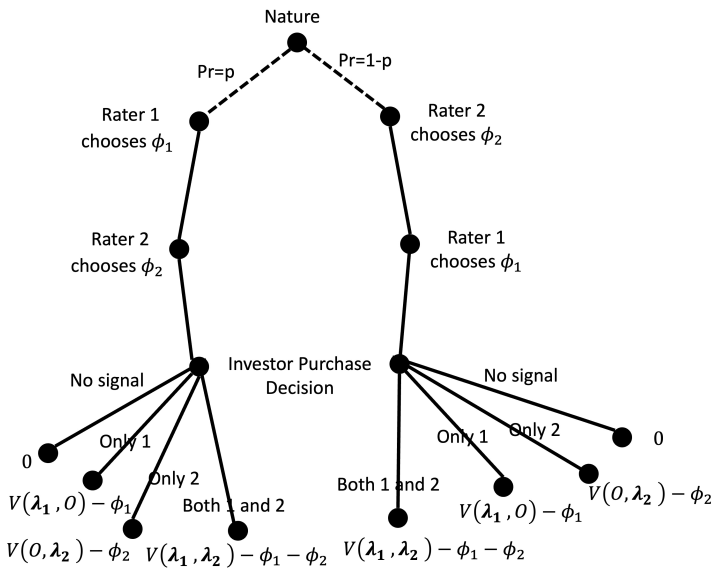


Figure IA.1. Sequence of actions and raters' payoff in the ratings market stage when the first mover is randomly determined.

maximizing outcomes are $(\boldsymbol{\lambda}_1^{SP_B}, \boldsymbol{\lambda}_2^{SP_A})$ *and* $(\boldsymbol{\lambda}_1^{SP_A}, \boldsymbol{\lambda}_2^{SP_B})$.

This result is intuitive since Rater 1 has a comparative advantage in producing information about category B when $b_1 > b_2$, and this specialization outcome avoids any overlap in information production.

IA.E. Alternative Allocations of Pricing-power

Suppose Rater 1 sets its fee first with probability $p \in [0.5, 1]$. Figure IA.1 presents the moves and the investor's payoff considered in this section.

Note that this modification does not impact the investment value function (i.e., $V(\cdot,\cdot)$). As a result, specialization in different categories is the value-maximizing outcome. Moreover, the expected payoff of each rater is a linear combination of the first-mover and second-mover's payoffs in equation 18:

$$\begin{aligned}\pi_1(\boldsymbol{\lambda}_1,\boldsymbol{\lambda}_2) &= p[V(\boldsymbol{\lambda}_1,\boldsymbol{\lambda}_2) - V(\boldsymbol{O},\boldsymbol{\lambda}_2)] \\ &\quad + (1-p)\min\{V(\boldsymbol{\lambda}_1,\boldsymbol{O}) - V(\boldsymbol{O},\boldsymbol{O}), V(\boldsymbol{\lambda}_1,\boldsymbol{\lambda}_2) - V(\boldsymbol{O},\boldsymbol{\lambda}_2)\} \\ \pi_2(\boldsymbol{\lambda}_1,\boldsymbol{\lambda}_2) &= (1-p)[V(\boldsymbol{\lambda}_1,\boldsymbol{\lambda}_2) - V(\boldsymbol{\lambda}_1,\boldsymbol{O})] \\ &\quad + p\min\{V(\boldsymbol{O},\boldsymbol{\lambda}_2) - V(\boldsymbol{O},\boldsymbol{O}), V(\boldsymbol{\lambda}_1,\boldsymbol{\lambda}_2) - V(\boldsymbol{\lambda}_1,\boldsymbol{O})\}.\end{aligned} \tag{IA.9}$$

Depending on whether the sum of the raters' marginal values exceeds their combined value or not (i.e., whether equation 19 holds or not), each rater's payoff is either its marginal value or a linear combination of its marginal value and its stand-alone value. Specifically, when the condition in equation 19 holds, we have:

$$\begin{aligned}\pi_1(\boldsymbol{\lambda}_1,\boldsymbol{\lambda}_2) &= p[V(\boldsymbol{\lambda}_1,\boldsymbol{\lambda}_2) - V(\boldsymbol{O},\boldsymbol{\lambda}_2)] + (1-p)(V(\boldsymbol{\lambda}_1,\boldsymbol{O}) - V(\boldsymbol{O},\boldsymbol{O})) \\ \pi_2(\boldsymbol{\lambda}_1,\boldsymbol{\lambda}_2) &= (1-p)[V(\boldsymbol{\lambda}_1,\boldsymbol{\lambda}_2) - V(\boldsymbol{\lambda}_1,\boldsymbol{O})] + p(V(\boldsymbol{O},\boldsymbol{\lambda}_2) - V(\boldsymbol{O},\boldsymbol{O}))\end{aligned} \tag{IA.10}$$

We see that p determines the allocation of surplus between the raters in this case. Both raters receive their stand-alone values, plus a fraction of their marginal values. As such, p can be thought of as a parameter capturing the allocation of market power in pricing. In Proposition IA.6, we characterize the equilibrium outcomes. The diagrams in Figure IA.2 illustrate how the set of equilibrium outcomes varies with β.

PROPOSITION IA.6: *Given Assumptions 1 and 2, the only possible equilibrium outcomes in pure strategies are generalization and specialization in different categories:*

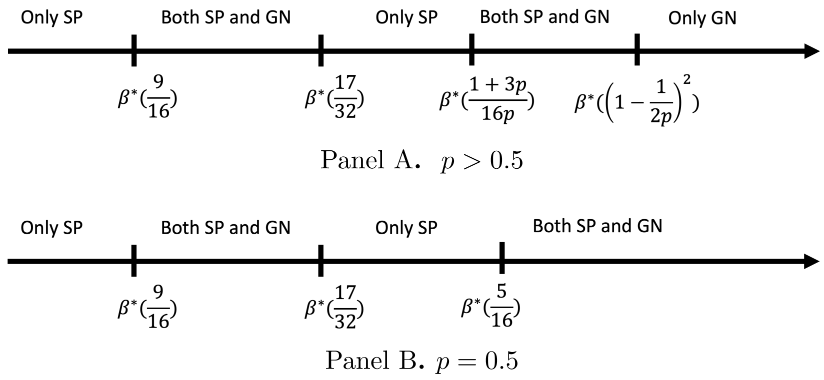


Figure IA.2. This diagram illustrates the set of market equilibria for different values of β when $u^{HL} < 0$ and the probability that Rater 1 sets its fee first $p \geq 0.5$. "GN" and "SP" represent generalization by both raters and specialization in different categories, respectively.

a) *If $u^{HL} \geq 0$, the only equilibrium outcomes are specialization in different categories.*

b) *Suppose $u^{HL} < 0$:*

b.1) *If $p > 0.5$, generalization by both raters is the unique market equilibrium when $\beta > \beta^*((1 - \frac{1}{2p})^2)$.*

b.2) *If $p > 0.5$, specialization in different categories is always an equilibrium when $\beta \leq \beta^*((1 - \frac{1}{2p})^2)$. Moreover, generalization by both raters is an equilibrium when $\beta \in [\beta^*(\frac{9}{16}), \beta^*(\frac{17}{32})] \cup [\beta^*(\frac{1+3p}{16p}), \beta^*((1 - \frac{1}{2p})^2)]$.*

b.3) *If $p = 0.5$, specialization in different categories is always an equilibrium. Generalization by both raters is also an equilibrium when $\beta \in [\beta^*(\frac{9}{16}), \beta^*(\frac{17}{32})]$ and $\beta \geq \beta^*(\frac{5}{16})$.*

The set of market equilibria remains the same for some parameter values, namely when $u^{HL} \geq 0$, and when $u^{HL} \leq 0$ and $\beta < \beta^*(\frac{17}{32})$; In these cases, both raters charge the marginal value of their ratings, thus their payoffs are the same as the baseline case.

Both generalization by both raters and specialization in different categories can be equilibria for some intermediate values of β due to the strategic complementarity motive discussed earlier. However, Proposition IA.6 suggests that there is an additional intermediate region with both specialization and generalization as equilibrium outcomes, i.e., $\beta \in [\beta^*(\frac{1+3p}{16p}), \beta^*((1-\frac{1}{2p})^2)]$ when $p \neq 0.5$ and $\beta \geq \beta^*(\frac{5}{16})$ when $p = 0.5$. The intuition is that both raters always care about their marginal values when $p < 1$, strengthening the strategic complementarity motive. However, multiple equilibria appear when the stand-alone values of both specialization and generalization are not too low compared to other options, as these also impact the raters' decisions. Generalization by both raters is the unique market equilibrium when β is sufficiently large, except when $p = 0.5$.

IA.F. Sequential Design of Rating Technologies

This section considers a scenario in which the raters design their rating technologies sequentially. Specifically, Rater 1 chooses $\boldsymbol{\lambda}_1$ first, and given this choice, Rater 2 chooses $\boldsymbol{\lambda}_2$. Then, similar to the baseline model, the raters set their fees sequentially given $(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)$, with Rater 1 moving first. The remainder of the game follows the structure of the baseline model. Proposition IA.7 characterizes the market equilibria under this configuration.

PROPOSITION IA.7: *Given Assumptions 1 and 2, and considering a sequential design of rating technologies, the equilibrium outcomes are as follows:*

a) *If* $u^{HL} \geq 0$*, the only equilibrium is that the raters specialize in different categories.*

b) *Suppose* $u^{HL} < 0$*.*

b.1) *If* $\beta^*(\frac{1}{4})$ *is finite and* $\beta > \beta^*(\frac{1}{4})$*, then the unique equilibrium is generalization by both raters.*

b.2) *If* $\beta < \beta^*(\frac{1}{4})$*, specialization in different categories is the only equilibrium. When* $\beta = \beta^*(\frac{1}{4})$*, both the specialization and generalization outcomes can be equilibria, depending on Rater 2's tie-breaking rule between specialization and generalization.*

To understand the intuition, note that when β is sufficiently large (i.e., $\beta > \beta^*(\frac{1}{4})$), Rater 2's best response for any choice of $\boldsymbol{\lambda}_1$ is generalization. Anticipating this, Rater 1 chooses $\boldsymbol{\lambda}_1 = \boldsymbol{\lambda}^{GN}$. When $\beta < \beta^*(\frac{1}{4})$ or $u^{HL} \geq 0$, Rater 2's best response to specialization is to specialize in the other category. Given this, Rater 1's payoff from specialization exceeds any other choice. Note that the generalization outcome does not emerge as equilibrium for any value of $\beta < \beta^*(\frac{1}{4})$, as opposed to the baseline model with simultaneous rating design, because the specialization outcome yields a higher payoff for Rater 1. Therefore, due to the sequential nature of the moves, Rater 1 is able to select its preferred outcome.

IA.G. Simultaneous Fee-setting

Our baseline model considers sequential fee-setting, following the raters' simultaneous choice of rating technologies. In this section, we show that the equilibria characterized in Proposition 1 are subgame-perfect when fee-setting occurs simultaneously; that is, at $t = 0$, the raters simultaneously set fees ϕ_1 and ϕ_2, after observing $(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)$. The remainder of the timeline is identical to the baseline model, as depicted in Figure 2.

To this end, we show that (ϕ_1, ϕ_2) specified in Equation 18 is a Nash equilibrium if the fee-setting is simultaneous. Note that ϕ_2 is Rater 2's best response to ϕ_1 by the equilibrium definition in Lemma 1. Given that $\phi_2 = \min\{V(\boldsymbol{O}, \boldsymbol{\lambda}_2), V(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) - V(\boldsymbol{\lambda}_1, \boldsymbol{O})\}$, the following inequalities show that the investor does not buy ratings from Rater 1 if Rater 1 sets $\hat{\phi}_1 > V(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) - V(\boldsymbol{O}, \boldsymbol{\lambda}_2)$:

$$\begin{aligned} V(\boldsymbol{O}, \boldsymbol{\lambda}_2) - \phi_2 &> V(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) - \hat{\phi}_1 - \phi_2 \\ V(\boldsymbol{O}, \boldsymbol{\lambda}_2) - \phi_2 \geq V(\boldsymbol{O}, \boldsymbol{\lambda}_2) + V(\boldsymbol{\lambda}_1, \boldsymbol{O}) - V(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) &> V(\boldsymbol{\lambda}_1, \boldsymbol{O}) - \hat{\phi}_1 \\ \Rightarrow V(\boldsymbol{O}, \boldsymbol{\lambda}_2) - \phi_2 &> \max\{V(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) - \hat{\phi}_1 - \phi_2, V(\boldsymbol{\lambda}_1, \boldsymbol{O}) - \hat{\phi}_1\}. \end{aligned} \tag{IA.11}$$

Therefore, $\phi_1 = V(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) - V(\boldsymbol{O}, \boldsymbol{\lambda}_2)$ is Rater 1's best response to ϕ_2. Given this result, it is straightforward to construct a subgame-perfect Nash equilibrium for any β in which: (i) the raters simultaneously choose $(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)$ from the equilibrium set in Proposition 1, and (ii) Rater 1 sets $\phi_1(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)$ and Rater 2 sets $\phi_2(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)$.

IA.H. Choice of Measurement Methodologies

We consider an investor that is only concerned about the project's performance in a single subcategory, say subcategory X. The performance in subcategory X is binary and represented by $w^X \in \{H, L\}$, where $\eta_X \in (0, 1)$

is the probability of $w^X = H$. The investor's net payoff if she invests is $\Delta + \beta u$, where:

$$u = \begin{cases} \Omega^H > 0 & w^X = H \\ \Omega^L < 0 & w^X = L. \end{cases} \tag{IA.12}$$

As in our main model, we assume that no investment takes place in the absence of ESG information. Thus, we modify Assumption 1 as follows:

ASSUMPTION 1′:

$$\Delta + \beta \mathbb{E}[u] = \eta_X(\Delta + \beta\Omega^H) + (1 - \eta_X)(\Delta + \beta\Omega^L) < 0. \tag{IA.13}$$

Suppose there are two noisy variables, m^a and m^b, that can be used to measure the performance. Measurement variables m^a and m^b are independent and have the following conditional distributions:

$$Prob(m^i = H|w^X = H) = z_H, \qquad Prob(m^i = H|w^X = L) = 1 - z_L, \qquad i = a, b, \tag{IA.14}$$

where $z_H, z_L \in (0, 1)$ capture the precision of these variables. We assume that the measurement variables are precise enough that the investor would invest if both variables indicate high performance:

$$\Delta + \beta \mathbb{E}[u|m^A = m^B = H] > 0. \tag{IA.15}$$

Furthermore, as in Assumption 2, we rule out the possibility that the investor becomes indifferent between investing and not investing when the two measures contradict:

ASSUMPTION 2″:

$$\Delta + \beta\mathbb{E}[u|m^A = H, m^B = L] \neq 0. \tag{IA.16}$$

Similar to the baseline model, the raters design rating technologies $(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)$ that map the performance measures into ratings s_j^a and s_j^b, $j = 1, 2$:

$$P(s_j^i = h|m^i = H) = \lambda_j^i, \quad P(s_j^i = h|m^i = L) = 0, \quad j = 1, 2, \quad i = a, b. \tag{IA.17}$$

Note that the raters can either perfectly measure the project's performance in one of the two methods, or generate some noisy ratings using both measurement methods. The raters are subject to the same technological constraint (equation 8) and follow the same fee-setting mechanism as in the baseline model. In Proposition IA.8, we describe the raters' equilibrium behavior:

PROPOSITION IA.8: *Under Assumptions 1′ and 2′, the only possible pure strategy equilibrium outcomes are that the raters specialize in different measurement methods, or generalize across the two methods. Specifically, define:*

$$\begin{aligned}
P_m^{HH} &= \eta_X z_H^2 + (1-\eta_X)(1-z_L)^2,\\
P_m^{HL} &= \eta_X z_H(1-z_H) + (1-\eta_X)z_L(1-z_L),\\
u_m^{HH} &= \mathbb{E}[u|m^A = m^B = H] = \frac{\eta_X z_H^2 m^H + (1-\eta_X)(1-z_L)^2 m^L}{\eta_X z_H^2 + (1-\eta_X)(1-z_L)^2}, \text{ and}\\
u_m^{HL} &= \mathbb{E}[u|m^A = H, m^B = L] = \frac{\eta_X z_H(1-z_H)m^H + (1-\eta_X)z_L(1-z_L)m^L}{\eta_X z_H(1-z_H) + (1-\eta_X)z_L(1-z_L)}.
\end{aligned} \tag{IA.18}$$

a) *If $u_m^{HL} \geq 0$, then the unique equilibrium is that the raters specialize in different measurement methods.*

b) *If* $u_m^{HL} < 0$*, define:*

$$\beta_m^*(\lambda) = \sup\{\beta | \frac{P_m^{HL}(\Delta + \beta u_m^{HL})}{P_m^{HH}(\Delta + \beta u_m^{HH})} \geq \lambda - 1\}. \tag{IA.19}$$

b.1) *If* $\beta_m^*(\frac{1}{4})$ *is finite and* $\beta > \beta_m^*(\frac{1}{4})$*, then the unique equilibrium is generalization by both raters.*

b.2) *If* $\beta \leq \beta_m^*(\frac{1}{4})$*, specialization in different measurement methods is an equilibrium. If* $\beta \in [\beta_m^*(\frac{9}{16}), \beta_m^*(\frac{17}{32})] \cup \{\beta_m^*(\frac{1}{4})\}$*, then generalization in different measurement methods is also an equilibrium.*

Proposition IA.8 states that both specializing in different measurement methods and generalizing in those are possible equilibria, depending on the parameter values. The specialization outcome is the unique equilibrium outcome when high performance in a single measurement variable is sufficient to indicate positive ESG performance in the corresponding subcategory ($u_m^{HL} \geq 0$). However, when β is sufficiently large and u_m^{HL} is sufficiently small, which happens when the measurements have a large false-positive error (small z_L), for example due to greater susceptibility to greenwashing, the raters generalize across the two methods.

With a logic similar to Proposition 2, one can show that the combined value is maximized when the raters specialize in different measurement methods.

IA.I. Mixed Strategy Equilibria

In order to simplify the characterization of the set of mixed strategy equilibria, we focus on equilibria that are robust to small perturbations to β (the ESG preference parameter). We formalize this refinement in Definition 1.

DEFINITION 1: *[Robust Equilibrium] Let σ_1 and σ_2 be some probability density functions over the set of rating technologies. We say mixed strategies σ_1 and σ_2 constitute a "robust equilibrium" if for $i = 1, 2$, σ_i is a best response to σ_{-i} in a neighborhood of β.*

Intuitively, this refinement ensures that the raters' choices of rating technology are robust to some uncertainty about the investor's preference parameter. One can verify that the pure strategy equilibria characterized in Proposition 1 are *robust*, except at the threshold values (i.e., $\beta^*(\frac{9}{16})$, $\beta^*(\frac{17}{32})$, and $\beta^*(\frac{1}{4})$). In Proposition IA.9, we characterize the set of robust equilibria in mixed strategies.

PROPOSITION IA.9: *Under Assumptions 1 and 2, the only outcome in mixed strategies that is a robust equilibrium for some values of β is that both raters randomize between $\boldsymbol{\lambda}^{SP_A}$ and $\boldsymbol{\lambda}^{SP_B}$ with equal probabilities. The following provides the details:*

a) *If $u^{HL} \geq 0$, the outcome above is an equilibrium. More generally, this is the only mixed strategy equilibrium outcome when not restricting to robust equilibria.*

b) *If $u^{HL} < 0$:*

b.1) *When* $\beta \geq \beta^*(\frac{1}{3})$, *there is no robust equilibrium in mixed strategies. If* $\beta > \beta^*(\frac{1}{4})$, *there is no equilibrium in mixed strategies even when not restricting to robust equilibria.*

b.2) *When* $\beta < \beta^*(\frac{1}{3})$, *the only robust equilibrium in mixed strategies is the outcome specified above.*

Proposition IA.9 demonstrates that both raters might mix between specialization in the two categories when $u^{HL} > 0$ or when $u^{HL} < 0$ and β is below a threshold value. This is intuitive since the pure specialization outcomes, i.e., $(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{\lambda}^{SP_B})$ and $(\boldsymbol{\lambda}^{SP_B}, \boldsymbol{\lambda}^{SP_A})$, require coordination among the raters. In the absence of this coordination, the raters might randomly specialize in a category. Hence, with a positive probability, this leads to the inefficient outcome that both raters specialize in the same category. For this outcome to be an equilibrium, the raters should randomize with equal probabilities, as otherwise, the raters would specialize in the category that the other rater specializes in with a lower probability and it would be profitable to deviate. Proposition IA.9 further states that this randomization between $\boldsymbol{\lambda}^{SP_A}$ and $\boldsymbol{\lambda}^{SP_B}$ is the only possible robust equilibrium in mixed strategies.

However, this outcome is not an equilibrium when $u^{HL} < 0$ and β is large. This is because as the stand-alone value of specialization decreases with β, so does Rater 2's payoff from specialization. In particular, when $\beta > \beta^*(\frac{1}{3})$ and Rater 1 randomizes between specialization in the two categories with equal probabilities, Rater 2 prefers to generalize instead of randomizing between $\boldsymbol{\lambda}^{SP_A}$ and $\boldsymbol{\lambda}^{SP_B}$.

Furthermore, Part (b.1) of Proposition IA.9 verifies the robustness of the key insight that when β is sufficiently large, the unique equilibrium is generalization by both raters. The intuition is that, in this case, generalization is the unique best response of Rater 2 to any choice of rating technology by Rater 1. This arises from the fact that generalization achieves the highest stand-alone value when $\beta \geq \beta^*(\frac{1}{4})$, and, according to Lemma 1, Rater 2's payoff is capped by the stand-alone value of its ratings. Because Rater 1's unique best response to generalization is also generalization, the generalization outcome is the only equilibrium outcome.

Overall, we see that allowing for mixed strategies does not affect the key insight developed by the baseline model: the raters generalize when the investor assigns a large weight to ESG performance, and may specialize otherwise.

IA.J. General Information Structure

Suppose each rater chooses conditional probabilities $\boldsymbol{\lambda}_j = (\lambda_j^{A_H}, \lambda_j^{A_L}, \lambda_j^{B_H}, \lambda_j^{B_L}) \in [0,1]^4$, where:

$$\lambda_j^{i_H} = Prob(s_j^i = h|w^i = H), \quad \lambda_j^{i_L} = Prob(s_j^i = l|w^i = L), \quad j = 1,2, \quad i = A, B, \tag{IA.20}$$

and the technological constraint is:

$$\lambda_j^{A_H} + \lambda_j^{A_L} + \lambda_j^{B_H} + \lambda_j^{B_L} \leq 3. \tag{IA.21}$$

Note that a higher value of $\lambda_j^{i_H}$ or $\lambda_j^{i_L}$ corresponds to a more precise rating in category i for rater j. Equation IA.21 implies that the raters face two types of trade-offs. The first is the trade-off between the precision of the ratings in categories A and B. The second trade-off, which is absent in our baseline model, is between the level of false-negative and false-positive errors in the ratings for each category. For instance, they can make their ratings more tilted toward high ratings by equally increasing the conditional probability of a high rating for both performance levels, i.e., $\lambda_j^{i_H}, \lambda_j^{i_L} \to \lambda_j^{i_H} + \varepsilon, \lambda_j^{i_L} - \varepsilon$, for category $i \in \{A, B\}$ and rater $j \in \{1, 2\}$.

Without loss of generality, we impose the conditions below to ensure that a high-performing project is not less likely to receive a high rating in a category than a low-performing one:

$$\begin{aligned} &Prob(s_j^i = h|w^i = H) \geq Prob(s_j^i = h|w^i = L) \\ &\Rightarrow \lambda_j^{A_H} \geq 1 - \lambda_j^{A_L}, \quad \lambda_j^{B_H} \geq 1 - \lambda_j^{B_L}, \quad j = 1, 2. \end{aligned} \tag{IA.22}$$

The baseline model corresponds to the case with $\lambda_j^{A_L} = \lambda_j^{B_L} = 1$, $j = 1, 2$, implying no false-positive errors in the ratings. The conditions above nest the technological constraint in the baseline model (equation 8), implying that all feasible rating technologies in the baseline model remain available. For instance, the specialized rating technologies correspond to $\boldsymbol{\lambda}^{SP_A} = (1, 1, 0, 1)$ and $\boldsymbol{\lambda}^{SP_B} = (0, 1, 1, 1)$,[9] and generalization corresponds to $\boldsymbol{\lambda}^{GN} = (\frac{1}{2}, 1, \frac{1}{2}, 1)$. To make the results comparable to those for the baseline setup, we set the

[9] $\boldsymbol{\lambda}^{SP_A}$ is informationally equivalent to any other rating technology that perfectly reveals the project's performance in category A and provides an uninformative rating for category B, i.e., $\boldsymbol{\lambda}^{SP'_A} = (1, 1, x, 1 - x)$, for $x \in [0, 1]$. Likewise, $\boldsymbol{\lambda}^{SP_B}$ is equivalent to $\boldsymbol{\lambda}^{SP'_B} = (y, 1 - y, 1, 1)$ for any $y \in [0, 1]$. For brevity, we do not report these equivalent choices in our characterizations.

right-hand side of the constraint in equation IA.21 to three. This choice keeps these special rating technologies at the frontier of the rating technologies available to raters.

Similar to the baseline model, the raters first simultaneously decide on their rating technologies, and then they sequentially set fees. The investor then decides which ratings to purchase and, lastly, decides whether or not to invest in the project given the realized ratings.

To simplify the characterization of the equilibria, we assume that the investor is so averse to investing in (L, L)-projects that she does not invest if, given her information, there is a positive probability that the project has low performance in both categories. This can be thought of as assigning a very low value to u^{LL}.

In Proposition IA.10, we describe the market equilibria in pure strategies. We once again employ the notion of robust equilibria, which is defined in Definition 1, to refine the set of equilibria.

PROPOSITION IA.10: *Under Assumptions 1 and 2, generalization by both raters and specialization in different categories are the only outcomes that can form a robust equilibrium in pure strategies for some values of β. The following provides the characterization in detail:*

a) *If $u^{HL} \geq 0$, then the only robust equilibrium outcome is that the raters specialize in different categories. This equilibrium outcome remains unique even when not restricting to robust equilibria.*

b) *Suppose $u^{HL} < 0$:*

b.1) *If $\beta > \beta^*(0)$, the unique robust equilibrium is that both raters gen-*

eralize. This uniqueness holds even when not restricting to robust equilibria.

b.2) *If $\beta \in (\beta^*(\frac{1}{2}), \beta^*(0)]$, there is no robust equilibrium.*

b.3) *If $\beta \leq \beta^*(\frac{1}{2})$, the unique robust equilibrium is that the raters specialize in different categories.*

The proof is provided in Section IA.L.IA.L.10. Proposition IA.10 demonstrates that the key insights developed in the baseline model continue to hold if we allow for a more flexible set of rating technologies. Thus, they are robust to our earlier assumption about the signal structure available to the raters.

As stated in Part (a) of Proposition IA.10, the unique equilibrium is that the raters specialize in different categories when $u^{HL} \geq 0$. The intuition is similar to the baseline case: In this case, a single rating ensuring a high performance in a category is enough for the investor to invest. As a result, the stand-alone value of specialization is large enough that specialization in different categories, the value-maximizing outcome, is an equilibrium. We show this equilibrium outcome is unique as well.

When $u^{HL} < 0$, the unique equilibrium is generalization when the investor is highly concerned about the project's ESG performance (i.e., high β). To understand the intuition, consider the extreme case where $|\frac{u^{HL}}{u^{HH}}|$ is very large; that is, the loss from investing in a (L, H)- or (H, L)-project is substantially larger than the gain from investing in a (H, H)-project. In this case, the investor is more concerned about the false-positive error in the ratings than the false-negative error because the former might lead to investment in projects with low performance in a category, resulting in a very low payoff, while the latter just inefficiently screens out some good projects, which is not as costly. As a result, the raters have a strong incentive to eliminate the false-positive error in their ratings, reducing the model to the baseline setup, where we find that generalization is the unique equilibrium outcome when β exceeds a threshold value.

Furthermore, there is no robust equilibrium for some intermediate values of β, as stated in Part (b.2) of Proposition IA.10. This is because neither specialization nor generalization maximizes the stand-alone value of the ratings when $\beta \in (\beta^*(\frac{1}{2}), \beta^*(0))$, given that we consider a more flexible set of rating technologies. In fact, Rater 2 chooses a rating technology that depends on β, implying that the equilibrium is not robust to β.

Specialization achieves the largest stand-alone value when $\beta \leq \beta^*(\frac{1}{2})$. Furthermore, we show that the best response to specialization is to specialize in the other category for both raters. Therefore, specialization in different categories forms an equilibrium. We also show that the specialization outcome is the only robust equilibrium.

IA.K. Unbundling the Ratings

Our baseline model assumes that raters are not allowed to sell their ratings for the two categories separately. We now show that if the option to unbundle is available, the raters are not able to charge more by unbundling their ratings.

To see this, consider the decision of Rater 1. Suppose the rater unbundles the ratings and sets fees ϕ_1^A and ϕ_1^B for its rating in categories A and B (i.e.,

s_1^A and s_1^B), respectively. Rater 2 continues to bundle its ratings and sells at ϕ_2. If $\phi_1^A + \phi_1^B$ is set above the marginal value of $\boldsymbol{\lambda}_1$ (i.e., $V(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) - V(\boldsymbol{O}, \boldsymbol{\lambda}_2)$), then Rater 2 sets ϕ_2 slightly below $\min\{V(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) - V(\boldsymbol{\lambda}_1, \boldsymbol{O}), V(\boldsymbol{O}, \boldsymbol{\lambda}_2)\}$, and the investor does not buy s_1^A and s_1^B together, because:[10]

$$V(\boldsymbol{O}, \boldsymbol{\lambda}_2) - \phi_2 > \max\{V(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) - \phi_1^A - \phi_1^B - \phi_2, V(\boldsymbol{\lambda}_1, \boldsymbol{O}) - \phi_1^A - \phi_1^B\}. \quad \text{(IA.25)}$$

Therefore, Rater 1 cannot set $\phi_1^A + \phi_1^B$ above its marginal value and sell its ratings in both categories. Also, Rater 1 cannot sell any individual rating at a fee above $V(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) - V(\boldsymbol{O}, \boldsymbol{\lambda}_2)$.[11] Therefore, the possibility to unbundle does not increase the fee that Rater 1 can charge. Using the same logic, the option to unbundle does not benefit Rater 2 either.[12]

[10]The inequalities below present the reasoning behind the inequality:

$$\begin{aligned} & V(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) - V(\boldsymbol{O}, \boldsymbol{\lambda}_2) < \phi_1^A + \phi_1^B \\ \Rightarrow\ & V(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) - (\phi_1^A + \phi_1^B) < V(\boldsymbol{O}, \boldsymbol{\lambda}_2) \\ \Rightarrow\ & V(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) - (\phi_1^A + \phi_1^B) - \phi_2 < V(\boldsymbol{O}, \boldsymbol{\lambda}_2) - \phi_2. \end{aligned} \quad \text{(IA.23)}$$

$$\begin{aligned} & V(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) - V(\boldsymbol{\lambda}_1, \boldsymbol{O}) > \phi_2 \\ \Rightarrow\ & V(\boldsymbol{\lambda}_1, \boldsymbol{O}) - (\phi_1^A + \phi_1^B) < V(\boldsymbol{\lambda}_1, \boldsymbol{O}) + V(\boldsymbol{O}, \boldsymbol{\lambda}_2) - V(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) < V(\boldsymbol{O}, \boldsymbol{\lambda}_2) - \phi_2 \end{aligned} \quad \text{(IA.24)}$$

[11]The maximum amount Rater 1 can charge by unbundling and selling only its rating in category A, s_1^A, is $V((\lambda_1^A, 0), \boldsymbol{\lambda}_2) - V(\boldsymbol{O}, \boldsymbol{\lambda}_2)$, which can be shown using an argument similar to the case where both ratings are sold. Likewise, the rating in category B can be sold for at most $V((0, \lambda_1^B), \boldsymbol{\lambda}_2) - V(\boldsymbol{O}, \boldsymbol{\lambda}_2)$. Both of these amounts are less than Rater 1's marginal value.

[12]The problem of optimal bundling in monopolistic and competitive environments has

IA.L. Proofs of the Results in the Internet Appendix

IA.L.1. Proof of Proposition IA.1 (Market equilibria with heterogeneous preferences over categories)

Similar to Lemma A1 and equation A9, and given Assumption IA.1, one can show that the combined value for type-i investors is $V_i(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) = \eta^2(\Delta + \beta u^{HH})v_i(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)$, where

$$v_A(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) = \lambda^A \max\{\lambda^B, \lambda^*\}, \quad v_B(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) = \lambda^B \max\{\lambda^A, \lambda^*\}. \qquad \text{(IA.26)}$$

The equations in (IA.26) imply that the ratings produce no value for type-i investors if they produce no information in category i. With no loss of generality, we can normalize the payoffs such that $\eta^2(\Delta + \beta u^{HH}) = 1$. This simplifies the notations.

To prove the proposition, we first characterize the equilibrium fees in Lemma 1, and then examine when the specialization and generalization outcomes are equilibria for different values of λ^*.

LEMMA 1 (Equilibrium fees with investor heterogeneity): *Suppose Raters 1 and 2 choose rating technologies $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2$, respectively. To simplify the*

been studied in the industrial organization literature (e.g., Adams and Yellen (1976); Zhou (2017)). However, these studies typically consider consumers with heterogeneous preferences, and their results are, therefore, not directly applicable to our baseline model.

exposition, define

$$V_i^{12} \equiv V_i(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2), \quad V_i^1 \equiv V_i(\boldsymbol{\lambda}_1, \boldsymbol{O}), \quad V_i^2 \equiv V_i(\boldsymbol{O}, \boldsymbol{\lambda}_2), \qquad i = A, B. \tag{IA.27}$$

The characterization of the equilibrium fees ϕ_1 and ϕ_2 depends on the types of investors purchasing ratings from each rater. Accordingly, we characterize the fees for each possible configuration. The equilibrium values and assignment are determined by the case that maximizes Rater 1's revenue, that is, the fraction of investors purchasing its ratings multiplied by ϕ_1. Among all outcomes that maximize Rater 1's revenue, Rater 2 picks the one that maximizes its revenue.

a) *If both investor types buy ratings from both raters in equilibrium, then the fees are:*

$$\begin{aligned} \phi_1^{AB,AB} &= \min\{V_A^{12} - V_A^2, V_B^{12} - V_B^2\} \\ \phi_2^{AB,AB} &= \min\{V_A^{12} - V_A^1, V_B^{12} - V_B^1, V_A^{12} - \phi_1^{AB,AB}, V_B^{12} - \phi_1^{AB,AB}\}. \end{aligned} \tag{IA.28}$$

b) *If both raters buy ratings from Rater 1 and only type-A investors buy from Rater 2 in equilibrium, then*

$$\begin{aligned} \phi_1^{AB,A} &= \min\{V_B^1, V_A^{12} - V_A^2\} \\ \phi_2^{AB,A} &= \min\{V_A^{12} - V_A^1, V_A^{12} - \phi_1^{AB,A}\} > V_B^{12} - V_B^1. \end{aligned} \tag{IA.29}$$

Similarly, one can find $\phi_1^{AB,B}$ and $\phi_2^{AB,B}$ when only type-B investors buy from Rater 2.

c) *If both investors buy from Rater 2, but only type-A investors buy from*

Rater 1, then

$$\begin{aligned} \phi_1^{A,AB} &= V_A^{12} - V_A^2 > V_B^{12} - V_B^2 \\ \phi_2^{A,AB} &= \min\{V_A^{12} - V_A^1, V_A^2, V_B^2\}. \end{aligned} \tag{IA.30}$$

Similarly, one can find $\phi_1^{B,AB}$ *and* $\phi_2^{B,AB}$ *when only type-B investors buy from Rater 1.*

d) *If only type-A investors buy from Rater 1 and only type-B investors buy from Rater 2, then*

$$\begin{aligned} \phi_1^{A,B} &= V_A^1 > V_B^{12} - V_B^2 \\ \phi_2^{A,B} &= V_B^2 > V_A^{12} - V_A^2. \end{aligned} \tag{IA.31}$$

$\phi_1^{B,A}$ and $\phi_2^{B,A}$ are determined similarly.

e) If type-A investors buy from both raters, and type-B investors do not purchase ratings, then

$$\begin{aligned} \phi_1^{A,A} &= V_A^{12} - V_A^2 > V_B^1 \\ \phi_2^{A,A} &= \min\{V_A^2, V_A^{12} - V_A^1\} > V_B^2. \end{aligned} \tag{IA.32}$$

$\phi_1^{B,B}$ and $\phi_2^{B,B}$ are determined similarly.

The proof strategy is similar to that for Lemma 1 and is omitted for brevity. Now, we proceed to prove the proposition, part by part.

a) To prove that $(\boldsymbol{\lambda}^{GN}, \boldsymbol{\lambda}^{GN})$ is the unique equilibrium, we show that (1) Rater 2's best response to any choice of $\boldsymbol{\lambda}_1$ is $\boldsymbol{\lambda}^{GN}$. (2) Rater 1's best response to Rater 2 generalizing is generalization.

To prove (1), first, it is straightforward to show that $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}^{GN}$ are complements for both investor types, for any $\boldsymbol{\lambda}_1$. This verifies that the equilibrium configuration is either case (a) or (c) in Lemma 1. In both cases, ϕ_2

cannot exceed both V_A^2 and V_B^2, and both are strictly less than $\frac{1}{4}$ if $\boldsymbol{\lambda}_2 \neq \boldsymbol{\lambda}^{GN}$. Moreover, one can show that if $\boldsymbol{\lambda}_2 = \lambda^{GN}$, then $V_A^2 = V_B^2 = \frac{1}{4}$, and therefore, $\phi_2 = \frac{1}{4}$, thereby maximizing Rater 2's revenue.

To prove (2), note that given $\boldsymbol{\lambda}_2 = \boldsymbol{\lambda}^{GN}$ and $\lambda^* < \frac{1}{4} < \frac{1}{2}$, $V_A^{12} = V_B^{12} = \lambda^A \lambda^B$, for any $\boldsymbol{\lambda}_1$. It is straightforward to show that $\boldsymbol{\lambda}_1 = \boldsymbol{\lambda}^{GN}$ maximizes $\lambda^A \lambda^B$. Therefore, both raters generalize and case (a) prevails; that is, both investor types purchase ratings from both raters.

b) We first consider the case where $\lambda^* \in [\frac{1}{4}, \frac{2}{3}]$. When the two raters specialize in different categories, say $\boldsymbol{\lambda}_1 = \boldsymbol{\lambda}^{SP_A}$ and $\boldsymbol{\lambda}_2 = \boldsymbol{\lambda}^{SP_B}$, one can show that

$$V_A^{12} = V_B^{12} = 1, \qquad V_A^1 = V_B^2 = \lambda^*, \qquad V_B^1 = V_A^2 = 0. \tag{IA.33}$$

In this case, both investor types purchase ratings from both raters, and the equilibrium fees are

$$\phi_1 = 1 - \lambda^*, \qquad \phi_2 = \min\{\lambda^*, 1 - \lambda^*\}. \tag{IA.34}$$

First, we show that Rater 2's best response to $\boldsymbol{\lambda}_1 = \boldsymbol{\lambda}^{SP_A}$ is $\boldsymbol{\lambda}_2 = \boldsymbol{\lambda}^{SP_B}$. Note that if Rater 2 sells ratings to only one group of investors, its revenue is at most $\frac{1}{2}\max\{\lambda^*, 1 - \lambda^*\}$, which is less than $\min\{\lambda^*, 1 - \lambda^*\}$ when $\lambda^* \leq \frac{2}{3}$. Therefore, Rater 2 chooses $\boldsymbol{\lambda}_2$ such that it sells ratings to both groups. Given this, Rater 2's revenue cannot exceed $V_A^{12} - V_A^1 \leq 1 - \lambda^*$, and $\max\{V_A^2, V_B^2\} \leq \lambda^*$. Therefore, Rater 2's revenue is bounded by $\min\{1 - \lambda^*, \lambda^*\}$, which is attained when $\boldsymbol{\lambda}_2 = \boldsymbol{\lambda}^{SP_B}$. One can show that Rater 1 can maximize its revenue by maximizing the combined value, which is achieved by setting

$\boldsymbol{\lambda}_1 = \boldsymbol{\lambda}^{SP_A}$ when to $\boldsymbol{\lambda}_2 = \boldsymbol{\lambda}^{SP_B}$. This confirms specialization in different categories as an equilibrium outcome.

Now, we consider the case where $\lambda^* \in (\frac{2}{3}, 1)$. The combined value and stand-alone values when the two raters specialize in different categories are the same as those in equation IA.33. If Rater 1 wishes to sell ratings to type-B investors, the equation implies that it can charge at most $1 - \lambda^*$. This is less than $\frac{1}{2}\lambda^*$, the revenue it earns by selling only to type-A investors. In fact, among the scenarios in Lemma 1, the scenario in case (d) is selected, and the revenue for both raters is $\frac{1}{2}\lambda^*$. No rater has the incentive to deviate. This is because $\frac{1}{2}\lambda^*$ is the maximum revenue that can be earned by selling to only one investor group. For instance, if Rater 1 wishes to choose $\boldsymbol{\lambda}_1$ so that it sells ratings to both investor types, its fee cannot exceed $V_B^{12} - V_B^2 \leq 1 - \lambda^*$, according to cases (a) and (b) in Lemma 1.

To prove that $(\boldsymbol{\lambda}^{GN}, \boldsymbol{\lambda}^{GN})$ cannot be an equilibrium when $\lambda^* > \frac{1}{4}$, except when $\lambda^* = \frac{9}{16}$, we consider the following cases separately:

$\boldsymbol{\lambda^* \in [\frac{1}{4}, \frac{1}{2}]}$: In this case, the combined and stand-alone values associated with $(\boldsymbol{\lambda}^{GN}, \boldsymbol{\lambda}^{GN})$ are:

$$V_A^{12} = V_B^{12} = \frac{9}{16}, \qquad V_A^1 = V_A^2 = V_B^1 = V_B^2 = \frac{1}{4}. \tag{IA.35}$$

Therefore, the revenues are (according to part (a) in Lemma 1):

$$\phi_1 = \frac{9}{16} - \frac{1}{4} = \frac{5}{16}, \qquad \phi_2 = \frac{1}{4}. \tag{IA.36}$$

One can show that Rater 2 can increase its revenue by deviating to $\boldsymbol{\lambda}_2 =$

$(\lambda_2^A, \lambda_2^B)$, where $\lambda_2^A = \frac{1}{4\lambda^*} + \varepsilon > \lambda^*$, and $\varepsilon > 0$ is sufficiently small. Note that $\lambda_2^B \leq 1 - \lambda_2^A < 1 - \frac{1}{4\lambda^*} \leq \lambda^*$. In this case,

$$V_A^{12} = V_B^{12} = \frac{1}{2} + \frac{1}{4}\lambda_2^A\lambda_2^B, \quad V_A^1 = V_B^1 = \frac{1}{4}, \quad V_2^A = \frac{1}{4} + \lambda^*\varepsilon, \quad V_2^B = \lambda_2^A\lambda_2^B < \frac{1}{4}, \tag{IA.37}$$

and the fees are

$$\phi_1 = V_A^{12} - V_A^2, \quad \phi_2 = \min\{\frac{1}{4} + \frac{1}{4}\lambda_2^A\lambda_2^B, \frac{1}{4} + \lambda^*\varepsilon\} > \frac{1}{4}. \tag{IA.38}$$

$\boldsymbol{\lambda^* \in (\frac{1}{2}, \frac{9}{16})}$: In this case, the combined and stand-alone values associated with $(\boldsymbol{\lambda}^{GN}, \boldsymbol{\lambda}^{GN})$ are:

$$V_A^{12} = V_B^{12} = \frac{9}{16}, \qquad V_A^1 = V_A^2 = V_B^1 = V_B^2 = \frac{1}{2}\lambda^*, \tag{IA.39}$$

and the fees are

$$\phi_1 = \frac{9}{16} - \frac{1}{2}\lambda^*, \qquad \phi_2 = \frac{1}{2}\lambda^*. \tag{IA.40}$$

Note that Rater 2's fee its stand-alone value for the two investor groups. Rater 2 can increase its fee, and thereby its revenue, by slightly increasing its effort in either direction. For instance, by deviating to $(\frac{1}{2} + \varepsilon, \frac{1}{2} - \varepsilon)$ for a sufficiently small $\varepsilon > 0$, its fee increases to $(\frac{1}{2} + \varepsilon)\lambda^*$.

$\boldsymbol{\lambda^* = \frac{9}{16}}$**:** Under the generalization outcome, both raters set $\phi_1 = \phi_2 = \frac{9}{16} - \frac{1}{2}\lambda^* = \frac{1}{2}\lambda^*$. No rater can increase its revenue by deviating to other rating technologies. This is because when the raters wish to sell ratings to both investor groups, their fee is constrained by their marginal values, and given $\boldsymbol{\lambda}_{-j} = \boldsymbol{\lambda}^{GN}$, $\boldsymbol{\lambda}_j = \boldsymbol{\lambda}^{GN}$ maximizes the combined value. Moreover, if they wish to sell ratings only to one group, their revenue cannot exceed $\frac{1}{2}\lambda^*$; because their fee is not greater than their stand-alone value, which is at most λ^*. Therefore, $(\boldsymbol{\lambda}^{GN}, \boldsymbol{\lambda}^{GN})$ is an equilibrium in this case.

$\boldsymbol{\lambda^* > \frac{9}{16}}$: When $\lambda^* \in (\frac{9}{16}, \frac{2}{3}]$, the combined value and stand-alone values when both raters generalize are similar to the previous case. The fess are $\phi_1 = \phi_2 = \frac{9}{16} - \frac{1}{2}\lambda^*$. Since $\phi_1 < \frac{1}{2}\lambda^*$, Rater 1 can increase its revenue by deviating to specialization in either category, say category A. In this case, $\phi_1 = \lambda^*$ and $\phi_2 = \frac{1}{2}\lambda^*$, type-A (type-B) investors purchase rating only from Rater 1 (Rater 2). Rater 1's revenue is $\frac{1}{2}\lambda^* > \frac{9}{16} - \frac{1}{2}\lambda^*$. The proof strategy for the case $\lambda^* > \frac{2}{3}$ is similar.

IA.L.2. Proof of Proposition IA.2 (Market equilibrium with asymmetric preferences)

Similar to Lemma A1, one can show that the combined value of the ratings is:

$$V(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) = \{\lambda^A\lambda^B + \lambda^A[\lambda^*_B - \lambda^B]^+ + \lambda^B[\lambda^*_A - \lambda^A]^+\}\eta^2(\Delta + \beta u^{HH}), \quad \text{(IA.41)}$$

where λ^A and λ^B are defined in (7). Without loss of generality, we normalize the payoffs such that $\eta^2(\Delta + \beta u^{HH}) = 1$ to simplify the exposition. The equilibrium fees continue to follow the characterization in Lemma 1.

a) $\boldsymbol{\lambda}^{GN}$ is Rater 2's best response to any choice of $\boldsymbol{\lambda}_1$ when $\lambda^*_A, \lambda^*_B \leq \frac{1}{4}$. This is because Rater 2's payoff is capped at its stand-alone value, according to Lemma 1. Moreover, $\boldsymbol{\lambda}^{GN}$ has the largest stand-alone value, and it is a complement for any $\boldsymbol{\lambda}_1$. Therefore, $\phi_2(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}^{GN}) = V(\boldsymbol{O}, \boldsymbol{\lambda}^{GN})$, confirming that generalization obtains the largest fee for Rater 2. It is straightforward to show that Rater 1's best response to generalization is also generalization.

b) First, we show that $(\boldsymbol{\lambda}^{SP_B}, \boldsymbol{\lambda}^{SP_A})$ is always an equilibrium when $\lambda_B^* \geq \frac{1}{4}$ and $\lambda_B^* \geq \lambda_A^*$. Since Rater 1 effectively maximizes the combined value of the ratings by choosing $\boldsymbol{\lambda}_1$ that maximizes its marginal value, its best response to $\boldsymbol{\lambda}^{SP_A}$ is $\boldsymbol{\lambda}^{SP_B}$ (Note that the specialization outcome is value-maximizing). Moreover, it is straightforward to show that $\boldsymbol{\lambda}^{SP_A}$ has the largest stand-alone value. Therefore, by choosing $\boldsymbol{\lambda}^{SP_A}$, Rater 2 maximizing its stand-alone value and marginal value given $\boldsymbol{\lambda}_1 = \boldsymbol{\lambda}^{SP_B}$. Therefore, $\boldsymbol{\lambda}^{SP_A}$ is Rater 2's best response, according to Lemma A2.

The other specialization outcome, $(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{\lambda}^{SP_B})$, is an equilibrium when λ_A^* is sufficiently large. With an argument similar to the previous case, $\boldsymbol{\lambda}^{SP_A}$ is Rater 1's best response to $\boldsymbol{\lambda}^{SP_B}$. However, when λ_A^* is sufficiently small, $\boldsymbol{\lambda}^{SP_B}$ is not Rater 2's best response to $\boldsymbol{\lambda}^{SP_A}$. Specifically, the necessary and sufficient condition for $\boldsymbol{\lambda}^{SP_B}$ not being the best response is that there exists $\boldsymbol{\lambda}_2 = (\lambda_2^A, \lambda_2^B)$ such that

$$\min\{[\lambda_2^B - \lambda_B^*]^+, \lambda_2^A \lambda_2^B\} > \lambda_A^*. \tag{IA.42}$$

This condition is equivalent to $\sqrt{\lambda_B^*} - \lambda_B^* > \lambda_A^*$. Therefore, $(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{\lambda}^{SP_B})$ is an equilibrium iff $\lambda_A^* \geq \sqrt{\lambda_B^*} - \lambda_B^*$.

In particular, when $\lambda_A^* < \sqrt{\lambda_B^*} - \lambda_B^*$, Rater 2's best response to $\boldsymbol{\lambda}^{SP_A}$ is $\boldsymbol{\lambda}_2^* = (1 - \sqrt{\lambda_B^*}, \sqrt{\lambda_B^*})$. By analyzing the first-order conditions, one can show that the best response to $\boldsymbol{\lambda}_2 = (\lambda_2^A, \lambda_2^B)$ is $\boldsymbol{\lambda}^{SP_A}$ when $\lambda_2^{B2} - \lambda_2^A \geq 0$.

Therefore, $(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{\lambda}_2^*)$ constitutes an equilibrium when

$$\lambda_B^* + \sqrt{\lambda_B^*} \geq 1 \Rightarrow \lambda_B^* \geq (\frac{\sqrt{5}-1}{2})^2. \tag{IA.43}$$

IA.L.3. *Proof of Proposition IA.3 (Market equilibria with heterogeneous preferences over raters)*

The following Lemma characterizes the equilibrium fees and payoffs for a given pair $(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)$, and general values of Z_1, Z_2, and Z_{12}. The proof follows similar steps to those in the proof of Lemma 1.

LEMMA 2: *Let $\pi_j(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)$ denote the equilibrium payoff of Rater j, $j = 1, 2$. The fees are represented by ϕ_j, $j = 1, 2$, as in the baseline model. Furthermore, let V_1, V_2, and V_{12} respectively represent the stand-alone value of Rater 1, the stand-alone value of Rater 2, and the combined value of the rating technologies.*

a) *If the rating technologies in $(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)$ are complements (i.e., the condition in equation 19 holds), then:*

$$\pi_1(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) = \max\{(Z_1 + Z_{12})V_1, Z_{12}(V_{12} - V_2)\}. \tag{IA.44}$$

$\phi_1 = V_1$ if the first element in (IA.44) is larger; otherwise, $\phi_1 = V_{12} - V_2$.

If $\phi_1 = V_{12} - V_2$, then $\phi_2 = V_2$, implying:

$$\pi_2 = (Z_2 + Z_{12})V_2. \tag{IA.45}$$

Otherwise,

$$\pi_2 = \max\{(Z_2 + Z_{12})V_2, Z_{12}(V_{12} - V_1)\}. \tag{IA.46}$$

In this case, $\phi_2 = V_2$ *if the first term is larger; and,* $\phi_2 = V_{12} - V_1$*, otherwise.*

b) *Suppose the rating technologies in the pair* $(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)$ *are substitutes (i.e., the condition in equation 19 does not hold).*

b.1) *If*

$$\begin{aligned} (Z_1 + Z_{12})(V_{12} - V_2) &\geq Z_1 V_1 \\ (Z_2 + Z_{12})(V_{12} - V_1) &\geq Z_2 V_2, \end{aligned} \tag{IA.47}$$

then:

$$\begin{aligned} \phi_1 = V_{12} - V_2, &\quad \phi_2 = V_{12} - V_1, \\ \pi_1 &= (Z_1 + Z_{12})(V_{12} - V_2) \\ \pi_2 &= (Z_2 + Z_{12})(V_{12} - V_1). \end{aligned} \tag{IA.48}$$

b.2) *If*

$$\begin{aligned} (Z_1 + Z_{12})(V_{12} - V_2) &< Z_1 V_1 \\ (Z_2 + Z_{12})(V_{12} - V_1) &\geq Z_2 V_2, \end{aligned} \tag{IA.49}$$

then:

$$\begin{aligned} \phi_1 = V_1, &\quad \phi_2 = V_2, \\ \pi_1 &= Z_1 V_1 \\ \pi_2 &= (Z_2 + Z_{12})V_2. \end{aligned} \tag{IA.50}$$

b.3) *If*

$$\begin{aligned} (Z_1 + Z_{12})(V_{12} - V_2) &< Z_1 V_1 \\ (Z_2 + Z_{12})(V_{12} - V_1) &< Z_2 V_2, \end{aligned} \tag{IA.51}$$

then ϕ_1 is the solution to:

$$(Z_2 + Z_{12})(V_2 - V_1 + \phi_1) = Z_2 V_2, \tag{IA.52}$$

and $\phi_2 = V_2$. Moreover,

$$\pi_1 = (Z_1 + Z_{12})\phi_1, \pi_2 = Z_2 V_2. \tag{IA.53}$$

We analyze the specialization and generalization outcomes separately. For each case, we divide the analysis into cases based on λ^*. For this proof, we assume that $Z_1 = Z_2$, following the statement of the proposition. The characterization of the fees when these outcomes are equilibria directly follows from Lemma 2.

Specialization in different categories

$\boldsymbol{\lambda^* \geq \frac{1}{2}}$: In this case, it is straightforward to show that all rating technologies are substitutes with the specialized rating technologies, i.e., $\boldsymbol{\lambda}^{SP_A}$ and $\boldsymbol{\lambda}^{SP_B}$. As a result, the payoffs are determined based on Part (b) in Lemma 2 when at least one rater specializes. There are two possibilities regarding the payoffs associated with the specialization outcome (Without loss of generality, we assume that $\boldsymbol{\lambda}_1 = \boldsymbol{\lambda}^{SP_A}$ and $\boldsymbol{\lambda}_2 = \boldsymbol{\lambda}^{SP_B}$).

First, suppose

$$(Z_1 + Z_{12})(V(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{\lambda}^{SP_B}) - V(\boldsymbol{O}, \boldsymbol{\lambda}^{SP_B})) \geq Z_1 V(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{O}). \tag{IA.54}$$

According to Lemma 2, the raters' payoffs are equal to $(Z_1 +$

$Z_{12})(V(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{\lambda}^{SP_B}) - V(\boldsymbol{O}, \boldsymbol{\lambda}^{SP_B}))$ when they specialize in different categories. Note that no rater can increase their marginal values. Also, they cannot increase their stand-alone values. Therefore, the inequality in (IA.54) implies that the raters cannot increase their payoffs by choosing a different rating technology and only selling to their captive investors. The only remaining possibility is that Rater 1 chooses $\boldsymbol{\lambda}_1'$ such that the condition in part (b.3) of Lemma 2 holds. Even in this case, one can use equation IA.52 to show that Rater 1's payoff is less than Rater 2's, who specializes and sells ratings only to its captive investors.

Second, suppose

$$(Z_1 + Z_{12})(V(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{\lambda}^{SP_B}) - V(\boldsymbol{O}, \boldsymbol{\lambda}^{SP_B})) < Z_1 V(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{O}). \quad \text{(IA.55)}$$

In this case, part (b.3) in Lemma 2 implies that under the specialization outcome, ϕ_1 is set such that:

$$(Z_1 + Z_{12})\phi_1 = Z_2 V(\boldsymbol{O}, \boldsymbol{\lambda}^{SP_B}) = Z_1 V(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{O}). \quad \text{(IA.56)}$$

Using an argument similar to that in the first case, it is straightforward to show that no rater has a profitable deviation.

$\boldsymbol{\lambda^* \in [\frac{1}{4}, \frac{1}{2})}$: In this case $\boldsymbol{\lambda}^{SP_A}$ and $\boldsymbol{\lambda}^{SP_B}$ are complements. Therefore, according to Part (a) in Lemma 2, Rater 1's payoff is:

$$\pi_1 = \max\{(Z_1 + Z_{12})V(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{O}), Z_{12}(V(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{\lambda}^{SP_B}) - V(\boldsymbol{O}, \boldsymbol{\lambda}^{SP_B}))\}. \quad \text{(IA.57)}$$

Rater 1 cannot increase its payoff by selling only to the common group, as the maximum amount they are willing to pay is $V(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{\lambda}^{SP_B}) - V(\boldsymbol{O}, \boldsymbol{\lambda}^{SP_B})$. This implies that the second term in the right-hand side of equation IA.57 is the maximum amount Rater 1 can collect by just selling to the common group, given $\boldsymbol{\lambda}_2 = \boldsymbol{\lambda}^{SP_B}$. Moreover, Rater 1 cannot charge the captive investors more than its stand-alone value, which is maximized under specialization. Therefore, the first term in (IA.57) represents the maximum amount Rater 1 can collect by selling to both the captive and common groups. As a result, Rater 1 has no incentive to deviate from specialization in category A, given $\boldsymbol{\lambda}_2 = \boldsymbol{\lambda}^{SP_B}$. A similar argument shows that Rater 2 also does not benefit from any deviation.

$\boldsymbol{\lambda^* < \frac{1}{4}}$: In this case, $\boldsymbol{\lambda}^{GN}$ is a complement for all rating technologies, including the specialized rating technologies. According to part (a) in Lemma 2, Rater 2's payoff from generalization is at least $(Z_2 + Z_{12})V(\boldsymbol{O}, \boldsymbol{\lambda}^{GN})$. It is straightforward to show that Rater 2's payoff under the specialization outcome is $(Z_2 + Z_{12})V(\boldsymbol{O}, \boldsymbol{\lambda}^{SP_B})$, which is lower than the payoff from switching to $\boldsymbol{\lambda}^{GN}$. This shows that the specialization outcome cannot be an equilibrium in this case.

Generalization by both raters

Now, we analyze when generalization by both raters is an equilibrium.

$\boldsymbol{\lambda^* > \frac{9}{16}}$: In this case, generalization has a lower stand-alone value than specialization. Moreover, given the other rater generalizing, the marginal value of generalization is also smaller than specialization. Given this, it is straightforward to show that both raters profit from deviating to specialization.

$\boldsymbol{\lambda^* \in [\frac{17}{32}, \frac{9}{16}]}$ **:** In this case $\boldsymbol{\lambda}^{GN}$ is a substitute for itself. We show that generalization by both raters cannot be an equilibrium in this case. First, consider the case where:

$$(Z_1 + Z_{12})(V(\boldsymbol{\lambda}^{GN}, \boldsymbol{\lambda}^{GN}) - V(\boldsymbol{O}, \boldsymbol{\lambda}^{GN})) < Z_1 V(\boldsymbol{\lambda}^{GN}, \boldsymbol{O}). \tag{IA.58}$$

In this case, according to part (b.3) in Lemma 2, Rater 2 only sells to its captive investors and it charges its stand-alone value. Therefore, the rater can increase its payoff by switching to specialization, as it has a higher stand-alone value.

Now, consider the case where:

$$(Z_1 + Z_{12})(V(\boldsymbol{\lambda}^{GN}, \boldsymbol{\lambda}^{GN}) - V(\boldsymbol{O}, \boldsymbol{\lambda}^{GN})) \geq Z_1 V(\boldsymbol{\lambda}^{GN}, \boldsymbol{O}). \tag{IA.59}$$

In this case, both raters get $\pi_1 = \pi_2 = (Z_1 + Z_{12})(V(\boldsymbol{\lambda}^{GN}, \boldsymbol{\lambda}^{GN}) - V(\boldsymbol{O}, \boldsymbol{\lambda}^{GN}))$. We show that either Rater 1 or Rater 2 can benefit from switching to specialization. First, consider the case where:

$$(Z_1 + Z_{12})(V(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{\lambda}^{GN}) - V(\boldsymbol{O}, \boldsymbol{\lambda}^{GN})) < Z_1 V(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{O}). \tag{IA.60}$$

Note that the marginal value of generalization is zero when the other rater specializes, i.e., $V(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{\lambda}^{GN}) = V(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{O})$, for this range of λ^*.[13] Given this and (IA.60), if Rater 1 deviates to $\boldsymbol{\lambda}^{SP_A}$, the payoffs are determined

[13]This directly follows from equation A9.

based on part (b.3) in Lemma 2. As a result, ϕ_1 is such that:

$$\begin{aligned}(Z_2 + Z_{12})(V(\boldsymbol{O}, \boldsymbol{\lambda}^{GN}) - V(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{O}) + \phi_1) = Z_2 V(\boldsymbol{O}, \boldsymbol{\lambda}^{GN}) \\ \Rightarrow \pi_1' = (Z_1 + Z_{12})\phi_1 = (Z_2 + Z_{12})\phi_1 \\ = Z_2 V(\boldsymbol{O}, \boldsymbol{\lambda}^{GN}) + (Z_1 + Z_{12})(V(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{O}) - V(\boldsymbol{O}, \boldsymbol{\lambda}^{GN})).\end{aligned} \tag{IA.61}$$

Given (IA.60), it is straightforward to show that $\pi_1' > \pi_1$.

Now, consider the case where:

$$(Z_1 + Z_{12})(V(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{\lambda}^{GN}) - V(\boldsymbol{O}, \boldsymbol{\lambda}^{GN})) \geq Z_1 V(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{O}). \tag{IA.62}$$

In this case, Rater 2 can benefit from deviating to specialization. In fact, the conditions of part (b.2) in Lemma 2 would hold after the deviation. This means that Rater 2's payoff would be $(Z_2 + Z_{12})V(\boldsymbol{O}, \boldsymbol{\lambda}^{SP_B})$. This clearly dominates Rater 2's payoff under the generalization outcome (π_2). Overall, we see that generalization by both raters cannot be an equilibrium.

$\boldsymbol{\lambda^* \in (\frac{1}{2}, \frac{17}{32})}$: In this case, $\boldsymbol{\lambda}^{GN}$ is a complement for itself. As a result, according to Lemma 2, Rater 2 receives $(Z_2 + Z_{12})V(\boldsymbol{O}, \boldsymbol{\lambda}^{GN})$ under the generalization outcome.[14] When $\lambda^* \in (\frac{1}{2}, \frac{17}{32})$, the rater can increase its stand-alone value, and thereby its payoff, by changing its rating technology from $\boldsymbol{\lambda}^{GN} = (\frac{1}{2}, \frac{1}{2})$ to $\boldsymbol{\lambda}_2' = (\frac{1}{2} - \varepsilon, \frac{1}{2} + \varepsilon)$ (ε is chosen such that $\boldsymbol{\lambda}_2'$ and $\boldsymbol{\lambda}^{GN}$ remain complements). Therefore, the generalization outcome is not an equilibrium.

[14]Note that if $Z_1 = Z_2$ and Z_{12} are such that Rater 1 prefers to charge its stand-alone value and sell to both its captive investors and the shared group, i.e., $(Z_1 + Z_{12})V(\boldsymbol{\lambda}^{GN}, \boldsymbol{O}) > Z_{12}(V(\boldsymbol{\lambda}^{GN}, \boldsymbol{\lambda}^{GN}) - V(\boldsymbol{O}, \boldsymbol{\lambda}^{GN}))$, then Rater 2 also prefers to charge its stand-alone value.

$\boldsymbol{\lambda^* \in (\frac{1}{4}, \frac{1}{2}]}$ **:** Similar to the previous case, $\boldsymbol{\lambda}^{GN}$ is a complement for itself, implying that Rater 2's payoff is $(Z_2+Z_{12})V(\boldsymbol{O}, \boldsymbol{\lambda}^{GN})$. Consider $\boldsymbol{\lambda}_2' = (\frac{1}{4\lambda^*}+\varepsilon, 1-\frac{1}{4\lambda^*}-\varepsilon)$. It is straightforward to show that $\boldsymbol{\lambda}_2'$ and $\boldsymbol{\lambda}^{GN}$ are complements, and $\boldsymbol{\lambda}_2'$ has a larger stand-alone value than $\boldsymbol{\lambda}^{GN}$. As such, Rater 2 profits from deviating to $\boldsymbol{\lambda}_2'$. This means that the generalization outcome cannot be an equilibrium in this case.

$\boldsymbol{\lambda^* \leq \frac{1}{4}}$ **:** In this case, any rating technology is a complement for $\boldsymbol{\lambda}^{GN}$. Therefore, given $\boldsymbol{\lambda}_2 = \boldsymbol{\lambda}^{GN}$, Rater 1's payoff is $\max\{(Z_1 + Z_{12})V(\boldsymbol{\lambda}_1, \boldsymbol{O}), Z_{12}(V(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}^{GN}) - V(\boldsymbol{O}, \boldsymbol{\lambda}^{GN}))\}$. This payoff is maximized at $\boldsymbol{\lambda}_1 = \boldsymbol{\lambda}^{GN}$, since $\boldsymbol{\lambda}^{GN}$ maximizes both the stand-alone value and marginal value given $\boldsymbol{\lambda}_2 = \boldsymbol{\lambda}^{GN}$.

Now, we show that Rater 2's best response to $\boldsymbol{\lambda}_1 = \boldsymbol{\lambda}^{GN}$ is to generalize. It is straightforward to show that Rater 2's payoff from generalization is:

$$\pi_2 = (Z_2 + Z_{12})V(\boldsymbol{O}, \boldsymbol{\lambda}^{GN}). \tag{IA.63}$$

Therefore, if $\boldsymbol{\lambda}_2' \neq \boldsymbol{\lambda}^{GN}$ obtains a higher payoff, Rater 2 should charge more than its stand-alone value. According to part (a) in Lemma 2, this implies that Rater 1 charges its stand-alone value, implying that:

$$(Z_1 + Z_{12})V(\boldsymbol{\lambda}^{GN}, \boldsymbol{O}) \geq Z_{12}(V(\boldsymbol{\lambda}^{GN}, \boldsymbol{\lambda}_2') - V(\boldsymbol{O}, \boldsymbol{\lambda}_2')). \tag{IA.64}$$

Note that Rater 2's payoff is

$$\begin{aligned} \pi_2' = Z_{12}(V(\boldsymbol{\lambda}^{GN}, \boldsymbol{\lambda}_2') - V(\boldsymbol{\lambda}^{GN}, \boldsymbol{O})) < Z_{12}(V(\boldsymbol{\lambda}^{GN}, \boldsymbol{\lambda}_2') - V(\boldsymbol{O}, \boldsymbol{\lambda}_2')) \\ \leq (Z_1 + Z_{12})V(\boldsymbol{\lambda}^{GN}, \boldsymbol{O}) = \pi_2, \end{aligned} \tag{IA.65}$$

where the last inequality follows from (IA.64). This is a contradiction that $\boldsymbol{\lambda}_2'$ obtains a higher payoff for Rater 2. As a result, generalization by both raters is an equilibrium.

IA.L.4. Proof of Proposition IA.4 (Market equilibria with heterogeneous information acquisition costs)

Recall from the proof of Proposition 1 that the raters' preferences across outcomes depend solely on λ^*, defined in equation A4. To prove the statements of the proposition, we first determine the values of λ^* for which the specialization outcomes (i.e., $(\boldsymbol{\lambda}_1^{SP_A}, \boldsymbol{\lambda}_2^{SP_B})$ and $(\boldsymbol{\lambda}_1^{SP_B}, \boldsymbol{\lambda}_2^{SP_A})$) constitute an equilibrium. This corresponds to parts (a) and (b.2) of the proposition. Then, we show that when $\lambda^* < \frac{b_2\bar{\lambda}}{4}$, the unique equilibrium is that Rater 2 chooses $\boldsymbol{\lambda}_2^{GN} \equiv (\frac{\bar{\lambda}}{2}, \frac{b_2\bar{\lambda}}{2})$, and Rater 1 chooses an interior rating technology. This corresponds to part (b.1) of the proposition.

Specialization outcomes

Let $BR_j(\boldsymbol{\lambda}_{-j})$ denote the set of best responses for rater $j \in \{1, 2\}$, when the other rater chooses $\boldsymbol{\lambda}_{-j}$. Since Rater 1 always maximizes its marginal value, it is straightforward to show that $\boldsymbol{\lambda}_1^{SP_A} = BR_1(\boldsymbol{\lambda}_2^{SP_B})$ and $\boldsymbol{\lambda}_1^{SP_B} = BR_1(\boldsymbol{\lambda}_2^{SP_A})$. Therefore, we only need to find values of λ^* for which specialization in a category is the best response for Rater 2 when Rater 1 specializes in the other category.

When $\lambda^* \geq \frac{b\bar{\lambda}}{4}$, $\boldsymbol{\lambda}_2^{SP_A}$ has the highest stand-alone value for Rater 2, and the highest marginal value given that $\boldsymbol{\lambda}_1 = \boldsymbol{\lambda}_1^{SP_B}$. Therefore, $\boldsymbol{\lambda}_2^{SP_A} = BR_2(\boldsymbol{\lambda}_1^{SP_B})$ if $\lambda^* \geq \frac{b_2\bar{\lambda}}{4}$, implying that $(\boldsymbol{\lambda}_1^{SP_B}, \boldsymbol{\lambda}_2^{SP_A})$ is an equilibrium outcome for this set of values of λ^*. To see that this outcome is not an equilibrium outcome when $\lambda^* < \frac{b_2\bar{\lambda}}{4}$, note that $\boldsymbol{\lambda}_2^{GN}$ has the largest stand-alone value, and it is straightforward to show that $\boldsymbol{\lambda}_2^{GN}$ is a complement for any choice of $\boldsymbol{\lambda}_1$ for this range of λ^*. Therefore, $\boldsymbol{\lambda}_2^{GN} = BR_2(\boldsymbol{\lambda}_1^{SP_B})$ when $\lambda^* < \frac{b_2\bar{\lambda}}{4}$.

Now, we analyze when $(\boldsymbol{\lambda}_1^{SP_A}, \boldsymbol{\lambda}_2^{SP_B})$ is an equilibrium. Note that $\boldsymbol{\lambda}_2 = \boldsymbol{\lambda}_2^{SP_B}$ has the highest marginal value when $\boldsymbol{\lambda}_1 = \boldsymbol{\lambda}_1^{SP_A}$. Therefore, the only possibility for $\boldsymbol{\lambda}_2^{SP_B} \neq BR_2(\boldsymbol{\lambda}^{SP_A})$ is the existence of $\boldsymbol{\lambda}_2'$ with a higher stand-alone value such that $\boldsymbol{\lambda}_2'$ and $\boldsymbol{\lambda}^{SP_A}$ are complements. More specifically, the conditions below should jointly hold for $\boldsymbol{\lambda}_2'$:

$$\begin{aligned} v(\boldsymbol{O}, \boldsymbol{\lambda}_2') &> v(\boldsymbol{O}, \boldsymbol{\lambda}_2^{SP_B}) \\ v(\boldsymbol{\lambda}_1^{SP_A}, \boldsymbol{\lambda}_2') &\geq v(\boldsymbol{\lambda}_1^{SP_A}, \boldsymbol{O}) + v(\boldsymbol{O}, \boldsymbol{\lambda}_2'). \end{aligned} \tag{IA.66}$$

It is straightforward to show that such $\boldsymbol{\lambda}_2'$ exists if and only if:

$$\lambda^* < \max\{\frac{b_2\bar{\lambda}}{4}, \frac{b_2}{(1+b_2\bar{\lambda})^2} - \frac{b_2(1-\bar{\lambda})}{1+b_2\bar{\lambda}}\}. \tag{IA.67}$$

Note that the right-hand side in the inequality above is less than $\frac{1}{4}$. Therefore, both specialization outcomes are equilibria when $\lambda^* \geq \frac{1}{4}$.

Finally, we analyze the case where $\lambda^* < \frac{b_2\bar{\lambda}}{4}$. In this case, $\boldsymbol{\lambda}_2^{GN}$ is Rater 2's best response to any choice of $\boldsymbol{\lambda}_1$. This is because $\boldsymbol{\lambda}_2^{GN}$ has the largest stand-alone value and it is a complement for any choice of $\boldsymbol{\lambda}_1$. Therefore, the unique equilibrium is $(BR_1(\boldsymbol{\lambda}_2^{GN}), \boldsymbol{\lambda}_2^{GN})$. As a result, it is sufficient to show that $BR_1(\boldsymbol{\lambda}_2^{GN})$ is interior. That is, $BR_1(\boldsymbol{\lambda}_2^{GN}) \notin \{\boldsymbol{\lambda}_1^{SP_A}, \boldsymbol{\lambda}_1^{SP_B}\}$.

Note that Rater 1 chooses $\boldsymbol{\lambda}_1$ to maximize its marginal value given Rater 2's choice of rating technology, implying that $\boldsymbol{\lambda}_1$ also maximizes the combined value. Therefore, we need to show that

$$\frac{d}{d\lambda_1^A} v(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2^{GN})|_{\lambda_1^A=0} > 0, \qquad \frac{d}{d\lambda_1^A} v(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2^{GN})|_{\lambda_1^A=1} < 0. \tag{IA.68}$$

Note that $v(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) = \lambda^A \lambda^B$ since $\lambda^A, \lambda^B \geq \frac{b_2 \bar{\lambda}}{2} > \lambda^*$. Therefore,

$$\frac{d}{d\lambda_1^A} v(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2^{GN})|_{\lambda_1^A=0} = \frac{d}{d\lambda_1^A} \lambda^A \lambda^B|_{\lambda_1^A=0} = (1-\frac{\bar{\lambda}}{2}) - b_1 \frac{\bar{\lambda}}{2}(1-\frac{b_2 \bar{\lambda}}{2}) > 0. \tag{IA.69}$$

Furthermore,

$$\begin{aligned} \frac{d}{d\lambda_1^A} v(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2^{GN})|_{\lambda_1^A=1} &= \frac{d}{d\lambda_1^A} \lambda^A \lambda^B|_{\lambda_1^A=1} = (1-\frac{\bar{\lambda}}{2})\frac{b_2 \bar{\lambda}}{2} - b_1(1-\frac{b_2 \bar{\lambda}}{2}) \\ \leq (1-\frac{\bar{\lambda}}{2})\frac{b_2 \bar{\lambda}}{2} - b_2(1-\frac{b_2 \bar{\lambda}}{2}) &= -b_2\{1 + \frac{\bar{\lambda}^2}{4} - \frac{\bar{\lambda}}{2}(1+b_2)\} < -b_2(1-\frac{\bar{\lambda}}{2})^2 < 0. \end{aligned} \tag{IA.70}$$

Therefore, $BR_1(\boldsymbol{\lambda}_2^{GN})$ is interior.

IA.L.5. Proof of Proposition IA.5 (Value-maximizing pairs with heterogeneous information acquisition costs)

From Lemma A1, recall that the value created by pair $(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)$ is:

$$\begin{aligned} V(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) &= (\Delta + \beta u^{HH}) v(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) \\ v(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) &= \lambda^A \lambda^B + \lambda^A [\lambda^* - \lambda^B]^+ + \lambda^B [\lambda^* - \lambda^A]^+. \end{aligned} \tag{IA.71}$$

We show that $v(\lambda_1, \lambda_2) \leq v(\boldsymbol{\lambda}_1^{SP_B}, \boldsymbol{\lambda}_2^{SP_A})$. There are three possibilities:

- $\lambda^* > \lambda^A, \lambda^B$: In this case, $v(\lambda_1, \lambda_2) = (\lambda^A + \lambda^B)\lambda^* - \lambda^A \lambda^B$. Since the expression is convex in λ_1^A and λ_2^A, the maximum value is attained when both raters specialize. Therefore, in this case, $v(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)$ is less than the maximum combined value attainable by a specialization outcome, which is $v(\boldsymbol{\lambda}_1^{SP_B}, \boldsymbol{\lambda}_2^{SP_A})$. There is no other maximizer except for $(\boldsymbol{\lambda}_1^{SP_A}, \boldsymbol{\lambda}_2^{SP_B})$ when $b_1 = b_2$.

- $\max\{\lambda^A, \lambda^B\} \geq \lambda^* \geq \min\{\lambda^A, \lambda^B\}$: In this case, we have:

$$v(\lambda_1, \lambda_2) = \max\{\lambda^A, \lambda^B\}\lambda^* \leq \lambda^* < b_2\bar{\lambda} \leq b_1\bar{\lambda} = v(\boldsymbol{\lambda}_1^{SP_B}, \boldsymbol{\lambda}_2^{SP_A}), \tag{IA.72}$$

  where $\lambda^* < b_2\bar{\lambda}$ is obtained from Assumption $2'$.

- $\lambda^A, \hat{\lambda}^B > \lambda^*$: In this case, $v(\lambda_1, \lambda_2) = \lambda^A\lambda^B$. One can show that the Hessian matrix for $\lambda^A\lambda^B$ has a positive eigenvalue, implying that the function does not have an interior global optimum. This implies that the global optimum is attained at a boundary point. Given this, it is straightforward to show that $v(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) \leq v(\boldsymbol{\lambda}_1^{SP_B}, \boldsymbol{\lambda}_2^{SP_A})$, with equality holding at $(\boldsymbol{\lambda}_1^{SP_B}, \boldsymbol{\lambda}_2^{SP_A})$, and $(\boldsymbol{\lambda}_1^{SP_A}, \boldsymbol{\lambda}_2^{SP_B})$ iff $b_1 = b_2$.

IA.L.6. Proof of Proposition IA.6 (Alternative Allocations of Pricing-power)

With steps similar to those in the proof of Proposition 1, one can show that the only equilibria are generalization by both raters and specialization in different categories. Given this, we only need to examine for what values of β these outcomes form an equilibrium.

In the proof of Proposition 1, we show that any two ratings are substitutes when $\lambda^* > \frac{17}{32}$, where λ^* is defined in equation A4. In this case, since the raters receive the marginal value of their ratings, the payoffs are similar to the baseline case ($p = 1$), and consequently, so are the equilibrium outcomes.

Now, we analyze for what values of $\lambda^* < \frac{17}{32}$, specialization in different categories is an equilibrium. Note that the specialization outcome is value-maximizing, and specialization in either category yields the largest stand-alone value when $\lambda^* \geq \frac{1}{4}$. As such, specialization in different categories remains an equilibrium when $\lambda^* \geq \frac{1}{4}$ for all values of $p \in [0.5, 1]$.

When $\lambda^* < \frac{1}{4}$, the specialized rating technologies are complements. In this case, Rater 2's expected payoffs is:

$$\pi_2(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{\lambda}^{SP_B}) = (1-p)(1-\lambda^*) + p\lambda^* = 1 - p + (2p-1)\lambda^*. \quad \text{(IA.73)}$$

Suppose Rater 2, in response to $\boldsymbol{\lambda}_1 = \boldsymbol{\lambda}^{SP_A}$, alternatively chooses $\boldsymbol{\lambda}_2 = (\lambda_2^A, \lambda_2^B)$. It is straightforward to show that $\boldsymbol{\lambda}^{SP_B}$ dominates $\boldsymbol{\lambda}_2$ if $\min\{\lambda_2^A, \lambda_2^B\} \leq \lambda^*$, and it is suboptimal to set $\lambda_2^A > \lambda_2^B$. Thus, suppose $\lambda_2^B \geq \lambda_2^A > \lambda^*$. Rater 2's expected payoff is:

$$\pi_2(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{\lambda}_2) = (1-p)(\lambda_2^B - \lambda^*) + p\lambda_2^A\lambda_2^B \Rightarrow \frac{d\pi_2}{d\lambda_2^B} = 1 - 2p\lambda_2^B. \quad \text{(IA.74)}$$

Note that if $p = 0.5$, the profit function is increasing in λ_2^B, which implies that $\boldsymbol{\lambda}^{SP_B}$ is indeed Rater 2's best response to $\boldsymbol{\lambda}^{SP_A}$. If $p > 0.5$, then the highest expected payoff Rater 2 can obtain from choices with $\lambda_2^B \geq \lambda_2^A > \lambda^*$ is when $\lambda_2^B = \frac{1}{2p}$. One can show that the expected payoff from this choice exceeds that from specializing in category B when $\lambda^* < (1 - \frac{1}{2p})^2$. In this case, specialization in different categories is not an equilibrium.

Generalization is a complement for itself when $\lambda^* \leq \frac{17}{32}$. Therefore, the

Rater 2's expected payoff from generalization is:

$$\pi_2(\boldsymbol{\lambda}^{GN}, \boldsymbol{\lambda}^{GN}) = (1-p)(\frac{9}{16} - \frac{1}{4}) + p\frac{1}{4} = \frac{1}{4} + \frac{1-p}{16}. \tag{IA.75}$$

Again, assume that $\lambda_2^B \geq \lambda_2^A$. It is straightforward to show that $\boldsymbol{\lambda}_2 = \boldsymbol{\lambda}^{GN}$ yields a higher expected payoff compared to all choices with $\lambda_2^A, \lambda_2^B \geq \lambda^*$ since $\lambda_2^A \lambda_2^B \leq \frac{1}{4}$ and generalization has the highest marginal value given the other rater generalizes when $\lambda^* < \frac{9}{16}$. Moreover, when $\lambda^* \leq \frac{1}{4}$, generalization has the highest stand-alone value among all rating technologies, implying that generalization is always an equilibrium when $\lambda^* \leq \frac{1}{4}$.

When $\lambda^* > \frac{1}{4}$, it is straightforward to show that among choices with $\lambda_2^A < \lambda^*$, specialization yields the highest expected payoff. Furthermore, note that Rater 2 has a stronger motive to deviate from the generalization outcome as it assigns a larger weight to its stand-alone value, given that generalization maximizes the marginal values. Therefore, we only need to find for what values of λ^*, $\pi_2(\boldsymbol{\lambda}^{GN}, \boldsymbol{\lambda}^{SP_B}) > \pi_2(\boldsymbol{\lambda}^{GN}, \boldsymbol{\lambda}^{GN})$. This translates into:

$$\frac{1-p}{4} + p\lambda^* > \frac{1}{4} + \frac{1-p}{16} \Rightarrow \lambda^* > \frac{1+3p}{16p}. \tag{IA.76}$$

IA.L.7. Proof of Proposition IA.7 (Market equilibria with sequential design of rating technologies)

We prove the proposition by separating the cases based on λ^*.

$\boldsymbol{\lambda^* \geq 1}$: This case corresponds to part (a) of the proposition, and part (b.2) when $\beta \leq \beta^*(1)$. In this case, Rater 2's best response to any choice of $\boldsymbol{\lambda}_1$ is to specialize in one of the two categories, as shown in the proof of

Proposition 1 (See Step 1 in the proof). Since Rater 1 chooses $\boldsymbol{\lambda}_1$ to effectively maximize the combined value of the rating technologies (See Lemma 1), Rater 1's optimal choice is to specialize; because, specialization in different categories create the largest combined value (See Proposition 2).

$\boldsymbol{\lambda^* \in [\frac{3}{4}, 1)}$: This case corresponds to part (b.2) when $\beta \in (\beta^*(1), \beta^*(\frac{3}{4})]$. Similar to the previous case, Rater 2's best response to any choice of $\boldsymbol{\lambda}_1$ is either $\boldsymbol{\lambda}^{SP_A}$ or $\boldsymbol{\lambda}^{SP_B}$. To see this, first note that specialization has the largest stand-alone value. Second, for any choice of $\boldsymbol{\lambda}_1$, the combined value is maximized when Rater 2 chooses $\boldsymbol{\lambda}^{SP_A}$ or $\boldsymbol{\lambda}^{SP_B}$. The proof of this statement is as follows.

Suppose, $\lambda_1^A \geq \frac{1}{2}$. Specifically, we show that

$$\hat{v}(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) \leq \max\{\lambda_1^A, \lambda^*\} = \hat{v}(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}^{SP_B}). \tag{IA.77}$$

Note that $\hat{v}(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) = \max\{\lambda^A\lambda^B, \lambda^A\lambda^*, \lambda^B\lambda^*, (\lambda^A + \lambda^B)\lambda^* - \lambda^A\lambda^B\}$ (See equation A9). Therefore, we only need to show that none of the four arguments of the maximization exceeds $\max\{\lambda_1^A, \lambda^*\}$. Since $\lambda^A, \lambda^B \leq 1$, this is straightforward for the second and third terms. For the last term:

$$\begin{aligned} (1-\lambda^A)(1-\lambda^B) \geq 0 \Rightarrow \lambda^*(\lambda^A + \lambda^B - 1) \leq \lambda^A + \lambda^B - 1 \leq \lambda^A\lambda^B \\ \Rightarrow \lambda^*(\lambda^A + \lambda^B) - \lambda^A\lambda^B \leq \lambda^* \leq \max\{\lambda_1^A, \lambda^*\}. \end{aligned} \tag{IA.78}$$

What remains to be shown is that, for any choice of $\boldsymbol{\lambda}_1$, $\lambda^A\lambda^B \leq \max\{\lambda_1^A, \lambda^*\}$. If $\lambda_1^A \geq \lambda^*$, by examining the first-order conditions, and noting that $\lambda^A\lambda^B$ is a concave function of λ_2^B, one can show that given $\boldsymbol{\lambda}_1$, $\lambda^A\lambda^B$ attains its maximum when $\lambda_2^B = 1$. In this case $\lambda^A\lambda^B = \lambda_1^A$. If $\lambda_1^A < \lambda^*$, it is straightforward to show that $\lambda^A\lambda^B \leq \lambda^*$ for any choice of $\boldsymbol{\lambda}_2$.[15]

Therefore, Rater 2's best response to any choice of $\boldsymbol{\lambda}_1$ is $\boldsymbol{\lambda}_2 = \boldsymbol{\lambda}^{SP_B}$, as this choice maximizes both stand-alone value and combined value of the rating technologies (See Lemma A2). Given this, Rater 1 also specializes.

$\boldsymbol{\lambda^* \in (\frac{1}{4}, \frac{3}{4})}$: This case corresponds to part (b.2) when $\beta \in (\beta^*(\frac{3}{4}), \beta^*(\frac{1}{4}))$. According to Proposition 1, Rater 2's best response to specialization is to specialize in the other category. Therefore, if Rater 1 specializes in either category, its payoff is $\hat{\phi}_1(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{\lambda}^{SP_B}) = \hat{\phi}_1(\boldsymbol{\lambda}^{SP_B}, \boldsymbol{\lambda}^{SP_A}) = 1 - \lambda^*$. As a result, we only need to show that for any pair of rating technologies $(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)$,

$$1 - \lambda^* \geq \hat{v}(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) - \hat{v}(\boldsymbol{O}, \boldsymbol{\lambda}_2). \tag{IA.79}$$

It is straightforward to show that $\max\{\lambda^A, \lambda^B\} \geq 0.75$ for any $(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)$. Therefore, we consider the following possibilities:

$\max\{\lambda^A, \lambda^B\} > \lambda^* \geq \min\{\lambda^A, \lambda^B\}$: Without loss of generality, suppose $\lambda^B > \lambda^A$. Furthermore, note that $\hat{v}(\boldsymbol{O}, \boldsymbol{\lambda}_2) \geq \lambda^* - \lambda_2^A\lambda_2^B$. Therefore,

$$\hat{v}(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) - \hat{v}(\boldsymbol{O}, \boldsymbol{\lambda}_2) \leq \lambda^B\lambda^* - \lambda^* + \lambda_2^A\lambda_2^B. \tag{IA.80}$$

It implies that to prove (IA.79) in this case, it is sufficient to show that $\lambda^B\lambda^* + \lambda_2^A\lambda_2^B \leq 1$. It is shown in the inequalities below:

$$\lambda^*\lambda^B + \lambda_2^A\lambda_2^B = \lambda^*\lambda_1^B + \lambda_2^B(1 + \lambda^*(1 - \lambda_1^B) - \lambda_2^B) \leq \lambda^*\lambda_1^B + (\frac{1 + \lambda^*(1 - \lambda_1^B)}{2})^2 \leq \frac{1}{4} + \lambda^* \leq 1. \tag{IA.81}$$

[15] Specifically, function $f(\boldsymbol{\lambda}_1) = \max_{\boldsymbol{\lambda}_2} \lambda^A\lambda^B$ is increasing in $\max\{\lambda_1^A, \lambda_1^B\}$.

$\lambda^A, \lambda^B \geq \lambda^*$: Similar to the previous case, one can show that:

$$\hat{v}(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) - \hat{v}(\boldsymbol{O}, \boldsymbol{\lambda}_2) \leq \lambda^A \lambda^B - \lambda^* + \lambda_2^A \lambda_2^B. \tag{IA.82}$$

Therefore, we only need to show that $\lambda^A \lambda^B + \lambda_2^A \lambda_2^B \leq 1$. This is shown in the inequalities below:[16]

$$\lambda^A \lambda^B + \lambda_2^A \lambda_2^B = (1 - (1 - \lambda_1^A)(1 - \lambda_2^A))(1 - (1 - \lambda_1^B)(1 - \lambda_2^B)) + \lambda_2^A \lambda_2^B$$
$$= (1 - \lambda_1^B \lambda_2^B)(1 - \lambda_1^A \lambda_2^A) + \lambda_2^A \lambda_2^B = 1 - (\lambda_1^A \lambda_1^B - \lambda_1^A \lambda_1^B \lambda_2^A \lambda_2^B) \underbrace{+ (\lambda_2^A - \lambda_1^B)(\lambda_2^B - \lambda_1^A)}_{-(\lambda_2^A - \lambda_1^B)^2} \leq 1. \tag{IA.83}$$

$\lambda^* \leq \frac{1}{4}$: This case corresponds to part (b.1) of the proposition, i.e., $\beta \geq \beta^*(\frac{1}{4})$ and $u^{HL} < 0$. Note that when $\lambda^* < \frac{1}{4}$, generalization is Rater 2's unique best response to any choice of λ_1. This is because, λ^{GN} has the largest stand-alone value, and it is a complement for itself.[17] Therefore, Rater 1 chooses $\boldsymbol{\lambda}^{GN}$ as it is the best response to generalization by Rater 2 (See Step 3 in the proof of Proposition 1). When $\lambda^* = \frac{1}{4}$, the equilibrium is determined by Rater 2's tie-breaking rule between specialization and generalization.

[16] In the inequalities, we note that the raters choose rating technologies on the frontier of the technological constraint in (8), which can be shown with a logic similar to that in Corollary A1.

[17] Note that according to Lemma 1, Rater 2's payoff is capped by its stand-alone value.

IA.L.8. Proof of Proposition IA.8 (Divergence in measurement)

Define $V_m(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)$ as the combined value of the ratings. We have:

$$\begin{aligned} V_m(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) &= P_m^{HH}\lambda^A\lambda^B(\Delta + \beta u_m^{HH}) \\ &+[P_m^{HH}\lambda^A(1-\lambda^B)(\Delta + \beta u_m^{HH}) + P_m^{HL}\lambda^A(\Delta + \beta u_m^{HL})]^+ \\ &+[P_m^{HH}\lambda^B(1-\lambda^A)(\Delta + \beta u_m^{HH}) + P_m^{HL}\lambda^B(\Delta + \beta u_m^{HL})]^+ \\ &= P_m^{HH}(\Delta + \beta u^{HH})\Big\{\lambda^A\lambda^B + \lambda^A[\lambda_m^* - \lambda^B]^+ + \lambda^B[\lambda_m^* - \lambda^A]^+\Big\}, \end{aligned} \tag{IA.84}$$

where λ_m^* is such that

$$P_m^{HH}(1-\lambda_m^*)(\Delta + \beta u_m^{HH}) + P_m^{HL}(\Delta + \beta u_m^{HL}) = 0. \tag{IA.85}$$

Therefore, the payoffs, and consequently the equilibrium outcomes, depend on λ_m^*, analogously to how they depend on λ^* in Proposition 1.

IA.L.9. Proof of Proposition IA.9 (Mixed strategy equilibria)

a) According to equation A4, $\lambda^* \geq 1$ when $u^{HL} > 0$. Therefore, Lemma A1 implies that the value functions are proportional to:

$$v(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) = (\lambda^A + \lambda^B)\lambda^* - \lambda^A\lambda^B. \tag{IA.86}$$

Moreover, any two rating technologies are substitutes when $\lambda^* \geq 1$ (shown in the proof of Proposition 1). Therefore, both raters receive their marginal contribution as their payoff. To prove the statement in part (a) of the proposition, we show that the best response to any mixed strategy is specialization or randomizing between specializing in the two categories.

The expression above is convex in λ_1^A and λ_2^A. Since a linear combination of convex functions is also convex, the expected payoffs are also convex when the other rater follows a mixed strategy. Therefore, the best response to any mixed strategy is to choose among the extreme values, i.e., $\lambda_i^A \in \{0, 1\}$, corresponding to $\boldsymbol{\lambda}^{SP_B}$ and $\boldsymbol{\lambda}^{SP_A}$.

It is straightforward to show that if rater $-j$ specializes in a category with a higher probability, the unique best response for rater j is to specialize in the other category. Therefore, the only possible mixed strategy equilibrium is that both raters randomize between specializing in the two categories with equal probabilities.

Part b.1) When $\lambda^* < \frac{1}{4}$, Rater 2's best response is always generalization, and it is unique; because,

$$\phi_2(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) \leq V(\boldsymbol{O}, \boldsymbol{\lambda}_2) \leq V(\boldsymbol{O}, \boldsymbol{\lambda}^{GN}) = \phi_2(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}^{GN}). \tag{IA.87}$$

In the inequalities above, the first inequality is obtained from Lemma 1. The second inequality reflects that generalization has the highest stand-alone value when $\lambda^* \leq \frac{1}{4}$, and the last inequality is resulted from $\boldsymbol{\lambda}^{GN}$ being a complement for any other rating technology when $\lambda^* \leq \frac{1}{4}$. Therefore, the only equilibrium is that both raters generalize, even when considering mixed strategy equilibria.

b.2) Suppose (σ_1, σ_2) is a pair of mixed strategies that constitute a robust equilibrium. Let Λ_j, $j = 1, 2$, be the support of σ_j; that is, the set of rating technologies that are selected with a positive probability by rater j. Since

Rater 1's payoff is always the marginal value of its ratings, any $\boldsymbol{\lambda}_1 \in \Lambda_1$ should maximize the expected combined value given σ_2 in a neighborhood of β. Therefore, any rating technology in Λ_1 should maximize the expression below in a neighborhood of λ^*

$$\boldsymbol{\lambda}_1 \in argmax\; v(\hat{\boldsymbol{\lambda}}, \sigma_2; \tilde{\lambda}^*), \quad \forall\, \boldsymbol{\lambda}_1 \in \Lambda_1, \tilde{\lambda}^* \in (\lambda^* - \varepsilon, \lambda^* + \varepsilon), \tag{IA.88}$$

where

$$v(\hat{\boldsymbol{\lambda}}, \sigma_2; \tilde{\lambda}^*) = \int_0^1 v(\hat{\boldsymbol{\lambda}}, \boldsymbol{\lambda}_2; \tilde{\lambda}^*)\sigma_2(\boldsymbol{\lambda}_2) d\lambda_2^A. \tag{IA.89}$$

Now, we show that the robustness of (σ_1, σ_2) implies that if there is an interior solution, then the solution is the unique best response to σ_2.

If $\boldsymbol{\lambda}_1$ is an interior solution, it implies that the right derivative of $v(\hat{\boldsymbol{\lambda}}, \sigma_2; \tilde{\lambda}^*)$ is zero at $\hat{\lambda}^A = \lambda_1^A$ in a neighborhood of λ^*.[18] Therefore,

$$\frac{\partial^2 v(\boldsymbol{\lambda}_1, \sigma_2; \lambda^*)}{\partial_+ \lambda_1^A \partial \lambda^*} = 0. \tag{IA.90}$$

Note that $\frac{\partial}{\partial_+ \lambda_1^A} v(\cdot, \sigma_2; \lambda^*)$ is linear in λ^*:

$$\begin{aligned}\frac{\partial v(\boldsymbol{\lambda}_1, \sigma_2; \lambda^*)}{\partial_+ \lambda_1^A} &= \sum_{\boldsymbol{\lambda}_2 \in \Lambda_2} \frac{\partial}{\partial \lambda_1^A} \lambda^A \lambda^B \mathbb{I}_{\{\lambda^* \leq \lambda^A, \lambda^B\}} \sigma_2(\boldsymbol{\lambda}_2) \\ +\lambda^* \big[\sum_{\boldsymbol{\lambda}_2 \in \Lambda_2} (1 - \lambda_2^A) &\mathbb{I}_{\{\lambda^* \geq \lambda^B\}} \sigma_2(\boldsymbol{\lambda}_2) - \sum_{\boldsymbol{\lambda}_2 \in \Lambda_2} (1 - \lambda_2^B) \mathbb{I}_{\{\lambda^* \geq \lambda^A\}} \sigma_2(\boldsymbol{\lambda}_2) \big].\end{aligned} \tag{IA.91}$$

[18] Since function v is the maximum of multiple differentiable functions, the right derivative is not smaller than the left derivative at any point. Therefore, if the right derivative is negative and there is a kink at $\boldsymbol{\lambda}_1$ for some $\boldsymbol{\lambda}_2 \in \Lambda_2$, then the left derivative is also negative, meaning that $\boldsymbol{\lambda}_1$ could not be a local optimum.

Equation IA.90 implies that the second line in (IA.91) is zero. Moreover, the expression in the bracket is weakly increasing in λ_1^A. Consider the contrary that there is another interior solution, say $\boldsymbol{\lambda}_1'$. Therefore, the signs of $\lambda^A - \lambda^*$ and $\lambda^B - \lambda^*$ should be the same for any pair of $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_1'$ with any rating technology in Λ_2. Therefore, the first term should be zero for both interior solutions. However, it is not possible since the derivative is linear in λ_1^A, implying that the derivative cannot have two roots.

Now, we show that $\boldsymbol{\lambda}^{SP_A}$ and $\boldsymbol{\lambda}^{SP_B}$ cannot be optimum. To see this, note that the expression in (IA.91) implies that $V(\boldsymbol{\lambda}^{SP_A}, \sigma_2) - V(\boldsymbol{\lambda}_1, \sigma_2)$ increases in λ^* or stays unchanged since the second term in (IA.91) is weakly increasing in λ_1^A. If it increases, it means that $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}^{SP_A}$ cannot be jointly optimal in a neighborhood of λ^*. Even if the second term in (IA.91) is zero for both $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}^{SP_A}$, the linearity of the derivative of the first term implies that $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}^{SP_A}$ cannot be jointly optimal. A similar argument also applies to $\boldsymbol{\lambda}^{SP_B}$. Therefore, if σ_2 has an interior best response, the best response is unique.

Therefore, the only possibility for a robust mixed strategy equilibrium is that both raters randomize between specializing in the two categories. With a logic similar to part (a), the raters should randomize with equal probabilities. One can show that this mixed strategies is dominated by generalization when $\lambda^* < \frac{1}{3}$.

IA.L.10. Proof of Proposition IA.10 (Market equilibria with the general structure of rating technologies)

By expanding equation 9, one can show that the combined value of any two rating technologies can be written as below:

$$
\begin{aligned}
V(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) &= \eta^2(\Delta + \beta u^{HH}) v(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) \\
v(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) &= l^A M^B + l^B M^A - l^A l^B,
\end{aligned}
\tag{IA.92}
$$

where:

$$
\begin{aligned}
l^i &= 1 - (1 - \lambda_1^{i_H} \mathbb{I}\{\lambda_1^{i_L} = 1\})(1 - \lambda_2^{i_H} \mathbb{I}\{\lambda_2^{i_L} = 1\}) \\
M^i &= [\lambda_1^{i_H} \lambda_2^{i_H} + (\lambda^* - 1)(1 - \lambda_1^{i_L})(1 - \lambda_2^{i_L})]^+ \\
&\quad + [(1 - \lambda_1^{i_H})\lambda_2^{i_H} + (\lambda^* - 1)\lambda_1^{i_L}(1 - \lambda_2^{i_L})]^+ \\
&\quad + [\lambda_1^{i_H}(1 - \lambda_2^{i_H}) + (\lambda^* - 1)(1 - \lambda_1^{i_L})\lambda_2^{i_L}]^+ \\
&\quad + [(1 - \lambda_1^{i_H})(1 - \lambda_2^{i_H}) + (\lambda^* - 1)\lambda_1^{i_L}\lambda_2^{i_L}]^+, \quad i = A, B,
\end{aligned}
\tag{IA.93}
$$

and λ^* is defined in equation A4. To simplify the exposition, we can scale the investment payoffs, without loss of generality, such that $\eta^2(\Delta + \beta u^{HH}) = 1$, implying that $V(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) = v(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)$.

The intuition for equation IA.92 is as follows: the investment takes place only if the possibility of type (L, L) is ruled out by the realized ratings. For instance, if $\lambda_1^{A_L} = 1$, s_1^A has no false-positive error, meaning that $Prob(w^A = L | s_1^A = h) = 0$. If $w^A = H$, with probability l^A, the project receives a high rating in category A that is free of a false-positive error. Given this high rating is in category A, the investor might invest depending on the realizations of s_1^B and s_2^B, which lead to an expected payoff of M^B, explaining the first term in the second equation in (IA.92). The intuition for the second term is similar. The last term corrects for double-counting. It is straightforward to verify that in the absence of false-positive errors, i.e., $\lambda_j^{i_L} = 1$, $i = A, B$, $j = 1, 2$, the expression for $v(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)$ boils down to equation A9. Now, we analyze the equilibria for different values of λ^*:

$\lambda^* \geq 1$

This case corresponds to $u^{HL} \geq 0$ (See equation A4). Since $\lambda^* \geq 1$, all terms in M^i in equation IA.93 are positive. Therefore, $M^A = M^B = \lambda^*$, and consequently

$$v(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) = (l^A + l^B)\lambda^* - l^A l^B. \tag{IA.94}$$

Since $\lambda^* \geq 1 \geq l^A, l^B$, the combined value is increasing in l^A and l^B. As a result, it is suboptimal to set $\lambda_j^{i_L} < 1$ for any $i = A, B$. The same logic applies to stand-alone values. In other words, it is suboptimal for the raters to introduce false-positive errors in their ratings since it would reduce l^i. Therefore, $\lambda_j^{A_L} = \lambda_j^{B_L} = 1$, $j = 1, 2$, which simplifies the game to the baseline case. According to part (a) in Proposition 1, the only equilibrium in pure strategies are specialization in different categories.

$\lambda^* < 0$

This case corresponds to Part b.1. We prove the following statements:

Statement 1: $(\boldsymbol{\lambda}^{GN}, \boldsymbol{\lambda}^{GN})$ is an equilibrium for any $\lambda^* < 0$.

Statement 2: There is no other equilibrium for these values of λ^*.

Proof of Statement 1: According to Lemma 1, Rater 2 receives the minimum of the stand-alone value of its ratings and their marginal value. With this observation, we only need to show: (1) $\boldsymbol{\lambda}^{GN}$ continues to uniquely maximize the stand-alone value when $\lambda^* < 0$. In the generalization outcome, Rater 2 obtains $V(\boldsymbol{O}, \boldsymbol{\lambda}^{GN})$, which exceeds its payoff from other choices of rating technology, as they have a lower stand-alone value[19]. (2) Rater 1's best response to $\boldsymbol{\lambda}^{GN}$ is also $\boldsymbol{\lambda}^{GN}$.

The stand-alone value of rating technology $\boldsymbol{\lambda}_2 = (\lambda_2^{A_H}, \lambda_2^{A_L}, \lambda_2^{B_H}, \lambda_2^{B_L})$ can be obtained from equation IA.92:

$$v(\boldsymbol{O}, \boldsymbol{\lambda}_2) = \lambda_2^{A_H}\mathbb{I}\{\lambda_2^{A_L} = 1\}[\lambda_2^{B_H} + (\lambda^* - 1)(1 - \lambda_2^{B_L})]^+$$
$$+\lambda_2^{B_H}\mathbb{I}\{\lambda_2^{B_L} = 1\}[\lambda_2^{A_H} + (\lambda^* - 1)(1 - \lambda_2^{A_L})]^+ - \lambda_2^{A_H}\lambda_2^{B_H}\mathbb{I}\{\lambda_2^{A_L} = 1\}\mathbb{I}\{\lambda_2^{B_L} = 1\}. \quad \text{(IA.95)}$$

According to equation IA.95, the stand-alone value would be zero if both $\lambda_2^{A_L}$ and $\lambda_2^{B_L}$ are less than one. This is intuitive since if $\lambda_2^{A_L}, \lambda_2^{B_L} < 0$ and the investor observes a high rating in both categories from Rater 2 (i.e., $s_2^A = s_2^B = h$), there is still a positive probability that the project has low performance in both categories. Hence, the investor never invests if she only purchases Rater 2's rating technology, resulting in a payoff of zero. Therefore, it is suboptimal for Rater 2 to choose a rating technology with $\lambda_2^{A_L}, \lambda_2^{B_L} < 1$, as it would lead to a payoff of zero.

If $\lambda_2^{A_L} = 1$ and $\lambda_2^{B_L} < 1$, it is straightforward to employ equation IA.95 to show that the stand-alone value of $\boldsymbol{\lambda}_2 = (\lambda_2^{A_H}, 1, \lambda_2^{B_H}, \lambda_2^{B_L})$ is strictly

[19] Note that we earlier showed that generalization is a complement for itself when $\lambda^* \leq \frac{17}{32}$.

less than that of $\boldsymbol{\lambda}_2' = (\lambda_2^{A_H}, 1, \lambda_2^{B_H} + \lambda_2^{B_L} - 1, 1)$. We earlier proved that generalization has the highest stand-alone value among rating technologies with $\lambda_2^{A_L} = \lambda_2^{B_L} = 1$ (i.e., the set of rating technologies with no false-positive error). Therefore, generalization maximizes the stand-alone value under the more general information structure when $\lambda^* < 0$, which completes step (1).

Step (2) is to prove that Rater 1's best response to generalization is generalization. To this end, we only need to show that it is suboptimal to set $\lambda_1^{A_L} < 1$ or $\lambda_1^{B_L} < 1$ when $\boldsymbol{\lambda}_2 = \boldsymbol{\lambda}^{GN}$ and $\lambda^* < 0$; because, Proposition 1 implies that generalization is the best response in the set of rating technologies with no false-positive error, i.e., $\lambda_1^{A_L} = \lambda_1^{B_L} = 1$.

Specifically, define $\boldsymbol{\lambda}_1' = (\lambda_1^{A_H} + \lambda_1^{A_L} - 1, 1, \lambda_1^{B_H} + \lambda_1^{B_L} - 1, 1)$. We prove that $V(\boldsymbol{\lambda}_1', \boldsymbol{\lambda}^{GN}) > V(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}^{GN})$ if $\lambda_1^{A_L} < 1$ or $\lambda_1^{B_L} < 1$. Consider the case that both $\lambda_1^{A_L} < 1$ and $\lambda_1^{B_L} < 1$ hold. The proof strategy for the other possibilities is similar:

$$v(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}^{GN}) = \frac{1}{2}\Big[\frac{1}{2} + \underbrace{\{\frac{1}{2}\lambda_1^{B_H} + (\lambda^* - 1)(1 - \lambda_1^{B_L})\}^+}_{<\frac{1}{2}(\lambda_1^{B_H} + \lambda_1^{B_L} - 1)} + \frac{1}{2} + \underbrace{\{\frac{1}{2}\lambda_1^{A_H} + (\lambda^* - 1)(1 - \lambda_1^{A_L})\}^+}_{<\frac{1}{2}(\lambda_1^{A_H} + \lambda_1^{A_L} - 1)}\Big] - \frac{1}{4}$$
$$< (\frac{1}{2} + \frac{1}{2}(\lambda_1^{B_H} + \lambda_1^{B_L} - 1))(\frac{1}{2} + \frac{1}{2}(\lambda_1^{A_H} + \lambda_1^{A_L} - 1)) = v(\boldsymbol{\lambda}_1', \boldsymbol{\lambda}^{GN}) \tag{IA.96}$$

As such, any rating technology with $\lambda_1^{A_L} < 1$ or $\lambda_1^{B_L} < 1$ is dominated by a rating technology with $\lambda_1^{A_L} = \lambda_1^{B_L} = 1$, which is the set of available rating technologies in the baseline case. Therefore, generalization is the unique best response for Rater 1 in response to $\boldsymbol{\lambda}_2 = \boldsymbol{\lambda}^{GN}$.

Proof of Statement 2:

Recall that in the baseline model, where we consider the restricted case with $\lambda_j^{A_L} = \lambda_j^{B_L} = 1$, $j = 1, 2$, Rater 2's best response to any rating technology is to generalize. The intuition is that any two rating technologies with no false-positive error are complements, and generalization yields the highest stand-alone value for Rater 2. However, this argument does not work in the more general case considered here since $\boldsymbol{\lambda}^{GN}$ is a substitute for some rating technologies. To prove that $(\boldsymbol{\lambda}^{GN}, \boldsymbol{\lambda}^{GN})$ remains the unique equilibrium in pure strategies, we rule out other possibilities of $(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)$ in several steps. In particular, we divide the cases based on the values of $\lambda_1^{A_L}, \lambda_1^{B_L}, \lambda_2^{A_L}, \lambda_2^{B_L}$.

$\boldsymbol{\lambda_1^{A_L}, \lambda_1^{B_L} < 1}$**:** The combined value in this case is,

$$v(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) = \lambda_2^{A_H}\mathbb{I}\{\lambda_2^{A_L} = 1\}M^B + \lambda_2^{B_H}\mathbb{I}\{\lambda_2^{B_L} = 1\}M^A - \lambda_2^{A_H}\lambda_2^{B_H}\mathbb{I}\{\lambda_2^{A_L} = 1\}\mathbb{I}\{\lambda_2^{B_L} = 1\}, \tag{IA.97}$$

where M^A and M^B are defined in (IA.93). If $\lambda_2^{A_L} < 1$, then Rater 1's rating for category B has no impact on the investor's decision, and consequently, the combined value. Therefore, specializing in category A would increase the combined value, which generates no false-positive error in category A. A similar argument applies when $\lambda_2^{B_L} < 1$. If $\lambda_2^{A_L} = \lambda_2^{B_L} = 1$, then one can show that the combined value increases if Rater 1 switches to $\boldsymbol{\lambda}_1' = (\lambda_1^{A_H} + \lambda_1^{A_L} - 1, 1, \lambda_1^{B_H} + \lambda_1^{B_L} - 1, 1)$. Therefore, there is no equilibrium in which $\lambda_1^{A_L}, \lambda_1^{B_L} < 1$.

$\boldsymbol{\lambda_1^{A_L}, \lambda_1^{B_L} = 1}$**:** In this case, $\boldsymbol{\lambda}_1$ is a complement for $\boldsymbol{\lambda}^{GN}$. Therefore, Rater 2 can obtain $\phi_2 = V(\boldsymbol{O}, \boldsymbol{\lambda}^{GN})$ by generalizing, which is the largest possible payoff. As such, $\boldsymbol{\lambda}^{GN}$ is Rater 2's best response, meaning that the only possible pure strategy equilibrium in this case is $(\boldsymbol{\lambda}^{GN}, \boldsymbol{\lambda}^{GN})$.

$\boldsymbol{\lambda_1^{A_L} = 1, \lambda_1^{B_L} < 1}$**:** First, we show that we should have $\lambda_2^{A_L} = 1$. If $\lambda_2^{A_L}, \lambda_2^{B_L} < 1$, then the stand-alone value of $\boldsymbol{\lambda}_2$ is zero, which means that the rater can increase its payoff by switching to another rating technology with a positive marginal value and a positive stand-alone value (such as generalization).

If $\lambda_2^{A_L} < 1, \lambda_2^{B_L} = 1$, then the combined value is:

$$\begin{aligned} v(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) = \lambda_1^{A_H}\lambda_2^{B_H} + \lambda_1^{A_H}[\lambda_1^{B_H}(1-\lambda_2^{B_H}) + (\lambda^* - 1)(1-\lambda_1^{B_L})]^+ \\ +\lambda_2^{B_H}[\lambda_2^{A_H}(1-\lambda_1^{A_H}) + (\lambda^* - 1)(1-\lambda_2^{A_L})]^+. \end{aligned} \tag{IA.98}$$

Rater 2 can increase both the combined value and stand-alone value of its ratings by switching to $\boldsymbol{\lambda}_2' = (\lambda_2^{A_H} + \lambda_2^{A_L} - 1, 1, \lambda_2^{B_H}, 1)$. To see this, take the derivative of (IA.98) with respect to $\lambda_2^{A_H}$ and $\lambda_2^{A_L}$, and note that the latter is larger than the former since $1 - \lambda^* > 1 \geq 1 - \lambda_1^{A_H}$.

The only remaining possibility is $\lambda_2^{A_L} = 1$. It implies that a high rating in category A, either from Rater 1 or 2, is sufficient to ensure that $w^A = H$. In this case, the combined value is:

$$v(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) = (\lambda_1^{A_H} + \lambda_2^{A_H} - \lambda_1^{A_H}\lambda_2^{A_H})M^B, \tag{IA.99}$$

where

$$\begin{aligned} M^B = \max\{\lambda_2^{B_H}\lambda_1^{B_H} + (\lambda^* - 1)(1-\lambda_2^{B_L})(1-\lambda_1^{B_L}), \lambda_2^{B_H} + (\lambda^* - 1)(1-\lambda_2^{B_L}), \\ \lambda_1^{B_H} + (\lambda^* - 1)(1-\lambda_1^{B_L}), \lambda_2^{B_H} + \lambda_1^{B_H} - \lambda_2^{B_H}\lambda_1^{B_H} + (1-\lambda^*)(1-\lambda_2^{B_L}\lambda_1^{B_L})\} \end{aligned} \tag{IA.100}$$

The four terms above reflect the four possibilities in the investment rule: The first term corresponds to the case that the investment requires a high rating in category B from both raters. The second (third) term represents the possibility that only the rating of Rater 2 (1) in category B is used for the investment. The last term corresponds to the possibility that a single high rating in category B is enough for the investment, in addition to receiving a high rating in category A. If M^B is equal to the second or the last term, then Rater 2 can increase the stand-alone and the combined value, by switching to $\boldsymbol{\lambda}_2 = (\lambda_2^{A_H}, 1, \lambda_2^{B_H} + \lambda_2^{B_L} - 1, 1)$. Likewise, if M^B is equal to the third term, Rater 1 can increase its marginal value by increasing $\lambda_1^{B_L}$. Therefore, if $(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)$ forms an equilibrium and $\lambda_1^{A_L} = \lambda_2^{A_L} = 1$, then M^B should be equal to the first item.

With this observation, we can write the combined value as below:

$$v(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) = (\lambda_1^{A_H} + \lambda_2^{A_H} - \lambda_1^{A_H}\lambda_2^{A_H})(\lambda_2^{B_H}\lambda_1^{B_H} + (\lambda^* - 1)(1 - \lambda_2^{B_L})(1 - \lambda_1^{B_L})). \tag{IA.101}$$

Note that Rater 1 chooses $\boldsymbol{\lambda}_1$ to maximize the combined value. The first-order conditions with respect to $\lambda_1^{B_H}$ and $\lambda_1^{B_L}$ reveal that we should have $\lambda_1^{B_H} = 1$. This is because $\boldsymbol{\lambda}_2$ has a positive stand-alone value, which implies that $\lambda_2^{B_H} > (1 - \lambda^*)(1 - \lambda_2^{B_L})$. Therefore, $\boldsymbol{\lambda}_1 = (a, 1, 1, 1 - a)$, for some $a \in [0, 1]$.

Furthermore, $a \leq \frac{1}{1-\lambda^*}$. This is because if $a > \frac{1}{1-\lambda^*}$, Rater 2 can increase the combined value and its stand-alone value by increasing $\lambda_2^{B_L}$ to one. This implies that Rater 1's best response is to specialize in category A, i.e., $a = 1$,[20] which is ruled out earlier.

Now, we consider two possibilities for $(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)$. If $(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)$ are substitutes

[20]Note that both rating technologies $(1, 1, 1, 0)$ and $(1, 1, 0, 1)$ represent specialization in category A.

and form an equilibrium, then by examining the first-order conditions with respect to $\lambda_2^{B_H}$ and $\lambda_2^{B_L}$, we find that we should have $\lambda_2^{B_H} = 1$ since we just showed that $1 > (1-\lambda^*)a$, implying that $\boldsymbol{\lambda}_2 = (b, 1, 1, 1-b)$ for some $b \in [0, 1]$. However, in the inequalities below, we show that rating technologies of this form are all complements:

$$\begin{aligned} v((a,1,1,1-a),(b,1,1,1-b)) - v((a,1,1,1-a),\boldsymbol{O}) - v(\boldsymbol{O},(b,1,1,1-b)) \\ = (a+b-ab)(1+(\lambda^*-1)ab) - a(1+(\lambda^*-1)a) - b(1+(\lambda^*-1)b) \\ = -ab + (1-\lambda^*)(a^2+b^2-ab(a+b-ab)) \\ \geq a^2+b^2-ab(1+a+b-ab) \geq (a-b)^2 \geq 0. \end{aligned} \tag{IA.102}$$

Therefore, $\boldsymbol{\lambda}_1 = (a, 1, 1, 1-a)$ and $\boldsymbol{\lambda}_2 = (\lambda_2^{A_H}, 1, \lambda_2^{B_H}, \lambda_2^{B_L})$ should be complements. The combined value is:

$$v(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) = (a + \lambda_2^{A_H} - a\lambda_2^{A_H})(\lambda_2^{B_H} + (\lambda^* - 1)a(1 - \lambda_2^{B_L})). \tag{IA.103}$$

Furthermore, we should have $\lambda_2^{B_L} < 1$, as otherwise, $a = 1$, namely specialization in category A would be optimal, to which generalization is Rater 2's best response. However, one can show that Rater 2 can increase the stand-alone value of its ratings by changing $\lambda_2^{B_H}$ and $\lambda_2^{B_L}$ to $\lambda_2^{B_H} - \varepsilon$ and $\lambda_2^{B_L} + \varepsilon$, for some sufficiently small $\varepsilon > 0$.[21] This perturbation is not feasible only

[21] Note that $\lambda_2^{B_H} > 0$, as otherwise, the stand-alone value of $\boldsymbol{\lambda}_2$ would be zero.

when the sum of the stand-alone values is equal to the combined value:

$$v(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) = v(\boldsymbol{\lambda}_1, \boldsymbol{O}) + v(\boldsymbol{O}, \boldsymbol{\lambda}_2) \Rightarrow (a + \lambda_2^{A_H} - a\lambda_2^{A_H})(\lambda_2^{B_H} + (\lambda^* - 1)a(1 - \lambda_2^{B_L})) \\ = a(1 + (\lambda^* - 1)a) + \lambda_2^{A_H}(\lambda_2^{B_H} + (\lambda^* - 1)(1 - \lambda_2^{B_L})). \tag{IA.104}$$

Note that Rater 1 chooses $\boldsymbol{\lambda}_1$ to maximize the combined value given $\boldsymbol{\lambda}_2$. Therefore, the combined value should not increase by switching to $\boldsymbol{\lambda}^{SP_B}$:

$$V(\boldsymbol{\lambda}^{SP_B}, \boldsymbol{\lambda}_2) \leq V(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) \Rightarrow \lambda_2^{A_H} \leq a(1 + (\lambda^* - 1)a) + \lambda_2^{A_H}(\lambda_2^{B_H} + (\lambda^* - 1)(1 - \lambda_2^{B_L})) \\ \Rightarrow \lambda_2^{A_H}(1 - \lambda_2^{B_H} + (1 - \lambda^*)(1 - \lambda_2^{B_L})) \leq a(1 + (\lambda^* - 1)a). \tag{IA.105}$$

Note that Constraint (IA.21), along with $\lambda_2^{A_L} = 1$, implies that:

$$\lambda_2^{A_H} \leq 2 - \lambda_2^{B_H} - \lambda_2^{B_L} \leq 1 - \lambda_2^{B_H} + (1 - \lambda^*)(1 - \lambda_2^{B_L}). \tag{IA.106}$$

Moreover, note that

$$a(1 + (\lambda^* - 1)a) \leq \frac{1}{4(1 - \lambda^*)}, \tag{IA.107}$$

where the upper bound is attained at $a = \frac{1}{2(1-\lambda^*)}$. Therefore, by combining inequalities IA.105-IA.107, we find:

$$\lambda_2^{A_H} \leq \sqrt{\frac{1}{4(1 - \lambda^*)}} < \frac{1}{2}. \tag{IA.108}$$

Likewise, the combined value should not increase if Rater 1 switches to

$\boldsymbol{\lambda}^{SP_A}$:

$$V(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{\lambda}_2) \leq V(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)$$

$$\Rightarrow \lambda_2^{B_H} + (\lambda^* - 1)(1 - \lambda_2^{B_L}) \leq a(1 + (\lambda^* - 1)a) + \lambda_2^{A_H}(\lambda_2^{B_H} + (\lambda^* - 1)(1 - \lambda_2^{B_L}))$$

$$\Rightarrow \lambda_2^{B_H} + (\lambda^* - 1)(1 - \lambda_2^{B_L}) \leq \frac{1}{1 - \lambda_2^{A_H}} a(1 + (\lambda^* - 1)a) \leq \frac{1}{2(1 - \lambda^*)}. \tag{IA.109}$$

Therefore, by combining the last two inequalities, we find:

$$\phi_2 \leq v(\boldsymbol{O}, \boldsymbol{\lambda}_2) = \lambda_2^{A_H}(\lambda_2^{B_H} + (\lambda^* - 1)(1 - \lambda_2^{B_L})) < \frac{1}{4(1 - \lambda^*)}. \tag{IA.110}$$

However, it is a contradiction since Rater 2 can increase its payoff by switching to $\boldsymbol{\lambda}_2' = (\frac{1}{2(1-\lambda^*)}, 1, 1, 1 - \frac{1}{2(1-\lambda^*)})$. Because, it achieves a stand-alone value of $\frac{1}{4(1-\lambda^*)}$, and it is a complement for $\boldsymbol{\lambda}_1 = (a, 1, 1, 1 - a)$, as demonstrated in (IA.102).

As such, there is no equilibrium in which $\lambda_1^{A_L} = 1, \lambda_1^{B_L} < 1$. Likewise, we can rule out the possibility that $\lambda_1^{A_L} < 1, \lambda_1^{B_L} = 1$. It completes the proof of the second statement.

$\lambda^* \in [0, 1)$

This case corresponds to parts (b.2) and (b.3) of the proposition. We examine the robust equilibrium outcomes by dividing the cases based on $\lambda_2^{A_L}, \lambda_2^{B_L}, \lambda_1^{A_L}, \lambda_1^{B_L}$. Specifically, we examine which pairs of $(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)$ form a robust equilibrium under Definition 1, namely the pair is an equilibrium in a neighborhood of λ^*.

Note that the stand-alone value of $\boldsymbol{\lambda}_2$ should be positive so Rater 2 obtains a positive fee. Therefore, $\lambda_2^{A_L} \geq 1$ or $\lambda_2^{B_L} \geq 1$.

$\boldsymbol{\lambda_2^{A_L} = \lambda_2^{B_L} = 1}$**:** In this case, $\boldsymbol{\lambda}_2$ generates no false-positive error, and belongs to the baseline set of feasible rating technologies. However, $\boldsymbol{\lambda}_1$ can be chosen from the more general set specified by (IA.21). By analyzing the first-order conditions, one can show that if $\lambda_2^{A_H} > \lambda^*$, then $\lambda_1^{A_L} = 1$. Moreover, if $\lambda_2^{A_H} < \lambda^*$, then $\lambda_1^{A_H} = 1$.[22] Likewise, either $\lambda_1^{B_H}$ or $\lambda_1^{B_L}$ is equal to one depending on whether $\lambda_2^{B_H}$ is smaller or bigger than λ^*.

Therefore, $\boldsymbol{\lambda}_1$ has to take one of the following three forms:

- $\boldsymbol{\lambda}_1 = (1, x, 1, 1-x)$ for some $x \in (0,1)$: The combined value in this case is

$$v(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) = \lambda_2^{A_H}\lambda_2^{B_H} + \lambda_2^{A_H}[1-\lambda_2^{B_H}+(\lambda^*-1)(1-x)]^+ + \lambda_2^{B_H}[1-\lambda_2^{A_H}+(\lambda^*-1)x]^+. \tag{IA.111}$$

  If $1 - \lambda_2^{B_H} + (\lambda^* - 1)(1-x) \leq 0$, then the realization of s_1^B has no impact on the investor's decision. As a result, Rater 1 could increase its payoff by specializing in category A, which is against $\boldsymbol{\lambda}_1$ having no false-positive error. Therefore, $1 - \lambda_2^{B_H} + (\lambda^* - 1)(1-x) > 0$. With the same logic, the last term in equation IA.111 should also be positive.

  Since x should maximize the combined value, the only possibility for an interior solution is that $\lambda_2^{A_H} = \lambda_2^{B_H} = 0.5$. Therefore, $v(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}^{GN}) = 0.25 + 0.5\lambda^*$. However, it is less than $v(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{\lambda}^{GN})$:

$$v(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{\lambda}^{GN}) = \max\{0.5, \lambda^*\} \geq 0.5(0.5 + \lambda^*) = v(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}^{GN}). \tag{IA.112}$$

[22] We do not analyze the knife-edge case that $\lambda_2^{A_H}$ or $\lambda_2^{B_H}$ is λ^* since the outcome cannot be a robust equilibrium.

Equality is achieved only for $\lambda^* = 0.5$. Therefore, there cannot be a robust equilibrium in this scenario.

- $\boldsymbol{\lambda}_1 = (x, 1, 1, 1-x)$ for some $x \in (0,1)$:[23] The combined value is

$$v(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) = (x + \lambda_2^{A_H} - x\lambda_2^{A_H})(\lambda_2^{B_H} + [1 - \lambda_2^{B_H} + (\lambda^* - 1)x]^+). \quad \text{(IA.113)}$$

  If $1 - \lambda_2^{B_H} + (\lambda^* - 1)x \leq 0$, then s_1^B has no impact on the investor's decision since the investor would invest only when $s_2^B = h$ and either s_1^A or s_2^A are also h. This implies that $\boldsymbol{\lambda}_1$ is a suboptimal choice. If $1 - \lambda_2^{B_H} + (\lambda^* - 1)x > 0$, then s_2^B is redundant since $M^B = 1 + (\lambda^* - 1)x$, meaning that the investor invests when $s_1^B = h$ and either s_1^A or s_2^A is h.

  If $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2$ are substitutes or the complementarity condition in Definition A1 holds with equality, then Rater 2 can increase its payoff by slightly increasing $\lambda_2^{A_H}$ and reducing $\lambda_2^{B_H}$. If $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2$ are complements and the complementarity condition does not hold with equality, then Rater 2 can increase payoff by switching to $(\lambda_2^{A_H}, 1, 1, \lambda_2^{B_H})$, which increases its stand-alone value without reducing its marginal value. As such, there is no equilibrium in this scenario.

- $\boldsymbol{\lambda}_1 = (x, 1, 1-x, 1)$ for some $x \in [0,1]$: In this case, $\boldsymbol{\lambda}_1$ also belongs to the set of rating technologies in the baseline model. In Proposition 1, we show that the only possible equilibria in pure strategies are

[23] The extreme cases (i.e., $x = 0, 1$) correspond to specialization in one of the categories, which are analyzed in the next case.

generalization by both raters and specialization in different categories. Therefore, we only need to examine if those outcomes remain in equilibrium when we expand the set of rating technologies available to the raters.

– The generalization outcome $(\boldsymbol{\lambda}^{GN}, \boldsymbol{\lambda}^{GN})$: We show that the generalization outcome is not an equilibrium for any value of $\lambda^* \in (0,1)$. To demonstrate this point, we show that **(1)** $\boldsymbol{\lambda}^{GN}$ is not Rater 1's best response to $\boldsymbol{\lambda}^{GN}$ when $\lambda^* > 0.5$. And, **(2)** $\boldsymbol{\lambda}^{GN}$ is not Rater 2's best response to $\boldsymbol{\lambda}^{GN}$ when $\lambda^* \leq 0.5$.

  Define $\boldsymbol{\lambda}^{(x)} = (x, 1, 1, 1-x)$. We show that when $\lambda^* \in (0.5, 1)$, $v(\boldsymbol{\lambda}^{(x)}, \boldsymbol{\lambda}^{GN}) > v(\boldsymbol{\lambda}^{GN}, \boldsymbol{\lambda}^{GN})$ for some $x \in (0,1)$, which proves (1) since Rater 1 effectively maximizes the combined value.

$$v(\boldsymbol{\lambda}^{(x)}, \boldsymbol{\lambda}^{GN}) = 0.5(1+x)(\frac{1}{2} + [\frac{1}{2} + (\lambda^* - 1)x]^+). \qquad \text{(IA.114)}$$

  For $\hat{x} = \frac{\lambda^*}{2(1-\lambda^*)}$, we have:

$$v(\boldsymbol{\lambda}^{(\hat{x})}, \boldsymbol{\lambda}^{\boldsymbol{GN}}) = \frac{(2-\lambda^*)^2}{8(1-\lambda^*)} > \frac{9}{16} = v(\boldsymbol{\lambda}^{\boldsymbol{GN}}, \boldsymbol{\lambda}^{\boldsymbol{GN}}) \qquad \lambda^* > 0.5. \qquad \text{(IA.115)}$$

  It proves (1).

  To prove (2), we note that $\boldsymbol{\lambda}^{GN}$ is a complement for itself when $\lambda^* \leq \frac{17}{32}$. In fact, any two rating technologies are complements in a neighborhood of $(\boldsymbol{\lambda}^{\boldsymbol{GN}}, \boldsymbol{\lambda}^{\boldsymbol{GN}})$ when $\lambda^* \leq 0.5 < \frac{17}{32}$. Therefore, $\boldsymbol{\lambda}^{GN}$ is a complement for $\boldsymbol{\lambda}'_2 = (\frac{1}{2}, 1, \frac{1}{2} + \varepsilon, 1 - \varepsilon)$ for sufficiently small values of $\varepsilon > 0$. $\boldsymbol{\lambda}_2'$ generates a larger stand-alone value:

$$v(\boldsymbol{O}, \boldsymbol{\lambda}_2') = \frac{1}{2}(\frac{1}{2}+\varepsilon+(\lambda^*-1)\varepsilon) = 0.25+0.5\lambda^*\varepsilon > 0.25 = v(\boldsymbol{O}, \boldsymbol{\lambda}^{GN}). \tag{IA.116}$$

As such, the generalization outcome is not an equilibrium when $\lambda^* \in (0, 1)$, whereas it was an equilibrium in the baseline case for some values of λ^* in this range.

– Specialization in different categories $(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{\lambda}^{SP_B}), (\boldsymbol{\lambda}^{SP_B}, \boldsymbol{\lambda}^{SP_A})$: We show that **(1)** this outcome is not an equilibrium when $\lambda^* \in [0, 0.5)$, and **(2)** this outcome remains an equilibrium when $\lambda^* \in [0.5, 1]$.

  To prove (1), note that $\boldsymbol{\lambda}^{SP_A}$ and $\boldsymbol{\lambda}^{SP_B}$ are complements when $\lambda^* \in [0, 0.5)$:

$$v(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{O}) + v(\boldsymbol{O}, \boldsymbol{\lambda}^{SP_B}) = 2\lambda^* < 1 = v(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{\lambda}^{SP_B}). \tag{IA.117}$$

  Therefore, rating technologies are complements in a neighborhood of $(\boldsymbol{\lambda}^{SP_A}, \boldsymbol{\lambda}^{SP_B})$. Let $\varepsilon > 0$ be sufficiently small such that $\boldsymbol{\lambda}^{(1-\varepsilon)} = (1-\varepsilon, 1, 1, \varepsilon)$ and $\boldsymbol{\lambda}^{SP_B}$ are complements. The inequality below demonstrates that the stand-alone value of $\boldsymbol{\lambda}^{(1-\varepsilon)}$ exceeds that of specialization:

$$\begin{aligned} v(\boldsymbol{O}, \boldsymbol{\lambda}^{(1-\varepsilon)}) &= (1-\varepsilon)(1+(\lambda^*-1)(1-\varepsilon)) \\ &= \lambda^* + (1-2\lambda^*)\varepsilon(1-\frac{1-\lambda^*}{1-2\lambda^*}\varepsilon) > \lambda^* = v(\boldsymbol{O}, \boldsymbol{\lambda}^{SP_A}). \end{aligned} \tag{IA.118}$$

  Therefore, $\boldsymbol{\lambda}^{SP_A}$ is not Rater 2's best response to $\boldsymbol{\lambda}^{SP_B}$.

To prove (2), we show that specialization in a category maximizes the stand-alone value. From Proposition 2, we know that for any value of λ^*, specializing in category A (B) obtains the highest combined value when the other rater specializes in category B (A). Therefore, according to lemma A2. Each rater's best response to specialization is to specialize in the other category.

To prove that specialization obtains the largest stand-alone value, we need to show this choice maximizes the following objective function (assuming $\lambda_2^{A_L} = 1$ to ensure a positive stand-alone value):

$$\max_{\lambda_2^{A_H}+\lambda_2^{B_H}+\lambda_2^{B_L}\leq 2} v(\boldsymbol{O}, \boldsymbol{\lambda}_2) = \lambda_2^{A_H} \max\{\lambda^*, \lambda_2^{B_H} + (\lambda^* - 1)(1 - \lambda_2^{B_L})\}. \tag{IA.119}$$

Since $\lambda^* \in (0, 1]$, $\lambda_2^{B_H} = 1$, as otherwise, the stand-alone value could be increased by slightly increasing $\lambda_2^{B_H}$ and decreasing $\lambda_2^{B_L}$ by the same amount. Therefore, the rating technology with the largest stand-alone value is $\boldsymbol{\lambda}^{(x)} = (x, 1, 1, 1-x)$ for some $x \in [0, 1]$. In fact, x should maximize $x(1+(\lambda^*-1)x)$. The maximizing value of x is one for $\lambda^* \in [0, 1]$ and it is $x^* = \frac{1}{2(1-\lambda^*)}$ for $\lambda^* \in (0, 0.5)$. Since $x = 1$ corresponds to $\boldsymbol{\lambda}^{SP_A}$, we see that specialization obtains the highest stand-alone value. Hence, specialization in different categories is an equilibrium when $\lambda^* \in [0.5, 1]$.

$\boldsymbol{\lambda_2^{A_L} = 1}$ and $\boldsymbol{\lambda_2^{B_L} < 1}$:[24] First, we show that it is not possible to have

[24]The argument is similar for $\lambda_2^{A_L} < 1$ and $\lambda_2^{B_L} = 1$

$\lambda_1^{A_L} < 1$ in equilibrium. If $\lambda_1^{B_L} < 1$ as well, then s_1^A has no impact on the investor's decision, implying that Rater 1 would be able to increase its payoff by switching to $\boldsymbol{\lambda}^{SP_B}$. If $\lambda_1^{B_L} = 1$, the combined value is:

$$\begin{aligned} v(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) = \lambda_2^{A_H}\lambda_1^{B_H} &+ \lambda_2^{A_H}[\lambda_2^{B_H}(1-\lambda_1^{B_H}) + (\lambda^* - 1)(1-\lambda_2^{B_L})]^+ \\ &+\lambda_2^{A_H}[(1-\lambda_2^{B_H})(1-\lambda_1^{B_H}) + (\lambda^* - 1)\lambda_2^{B_L}]^+ \\ &+\lambda_1^{B_H}[\lambda_1^{A_H}(1-\lambda_2^{A_H}) + (\lambda^* - 1)(1-\lambda_1^{A_L})]^+ \\ &+\lambda_1^{B_H}[(1-\lambda_1^{A_H})(1-\lambda_2^{A_H}) + (\lambda^* - 1)\lambda_1^{A_L}]^+. \end{aligned} \quad \text{(IA.120)}$$

In equation IA.120, the terms correspond to the investor's value from signal realizations $(s_2^A = h, s_1^B = h)$,$(s_2^A = h, s_1^B = l, s_2^B = h)$,$(s_2^A = h, s_1^B = l, s_2^B = l)$,$(s_2^A = l, s_1^A = h, s_1^B = h)$, and $(s_2^A = l, s_1^A = l, s_1^B = h)$, respectively. If the third or fifth term is positive, then signal realizations in categories B and A, respectively, have no impact on the investor's decision, which cannot happen in equilibrium since Rater 1 could increase the combined value by specializing in one of the categories. Moreover, the fourth term should be positive, as otherwise, the realization of s_1^A has no impact on the investor's decision. Moreover, if $\lambda_2^{A_H} \geq \lambda^*$, Rater 1 can increase its payoff by switching to $\boldsymbol{\lambda_1}' = (\lambda_1^{A_H} + \lambda_1^{A_L} - 1, 1, \lambda_1^{B_H}, 1)$, which would violate the assumption that $\lambda_1^{A_L} < 1$. Therefore, we should have $\lambda_2^{A_H} < \lambda^*$, which implies that $\lambda_1^{A_H} = 1$. That is, $\boldsymbol{\lambda}_1 = (1, 1-y, y, 1)$ for some $y \in [0,1]$. Thus, the combined value can be rewritten as below:

$$v(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) = y[1+(\lambda^*-1)y] + \lambda_2^{A_H}[\lambda_2^{B_H}(1-y) + (\lambda^*-1)(1-\lambda_2^{B_L})]^+. \quad \text{(IA.121)}$$

The first term in equation IA.121 represents the stand-alone value of $\boldsymbol{\lambda}_1$, and the second-term is less than the stand-alone value of $\boldsymbol{\lambda}_2$. Therefore, two rating technologies are substitutes. As a result, $\boldsymbol{\lambda}_2$ should maximize the combined value given $\boldsymbol{\lambda}_1$. The first-order conditions imply that if $y > \lambda^*$, then $\lambda_2^{B_L} = 1$, which contradicts our original assumption. Therefore, $y < \lambda^*$, which further implies $\lambda_2^{B_H} = 1$. As a result, the only possibility is that $\boldsymbol{\lambda}_2 = (x, 1, 1, 1 - x)$ for some $x \in [0, 1]$. The combined value is:

$$x + y - xy + (\lambda^* - 1)(x^2 + y^2). \tag{IA.122}$$

The first-order conditions with respect to x and y imply that $x = y = \frac{1}{1+2(1-\lambda^*)}$. This implies a combined value of $v(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) = \frac{1}{1+2(1-\lambda^*)} = y$, which is less than λ^*, as shown earlier. Therefore, Rater 1 can increase the combined value by specializing in a category, meaning that $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2$ cannot form an equilibrium. Therefore, there is no equilibrium in which $\lambda_2^{A_L} = \lambda_1^{B_L} = 1$ and $\lambda_2^{B_L}, \lambda_1^{A_L} < 1$.

Lastly, we analyze the possibility that $\lambda_1^{A_L} = 1$ and $\lambda_1^{B_L} \leq 1$. The combined value is:

$$v(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) = \lambda^A M^B, \tag{IA.123}$$

where $\lambda^A = \lambda_1^{A_H} + \lambda_2^{A_H} - \lambda_1^{A_H}\lambda_1^{B_H}$, and M^B is defined in equation IA.93. By examining the first-order conditions, we find that we should have either $\lambda_1^{B_H} = 1$ or $\lambda_1^{B_L} = 1$ for a robust equilibrium. Moreover, since a robust equilibrium cannot be on the borderline of the set of complement rating technology pairs as it varies with λ^*, $\boldsymbol{\lambda}_2$ should maximize the combined value or stand-alone value in a neighborhood of $(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)$, depending on whether $\boldsymbol{\lambda}_1$

and $\boldsymbol{\lambda}_2$ are substitutes or complements.

First, we rule out the possibility that the rating technologies are complements: For $\lambda^* \in [0.5, 1]$, $\boldsymbol{\lambda}^{SP_A}$ and $\boldsymbol{\lambda}^{SP_B}$ maximize the stand-alone value. For $\lambda^* < 0.5$, the rating technology that maximizes the stand-alone value depends on λ^*, so it cannot be part of a robust equilibrium. Therefore, $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2$ should be substitutes.

By analyzing the first-order conditions, we can show that there are two possibilities: either $\lambda_2^{B_H} = 1$ or $\lambda_2^{B_L} = 1$, where the latter is ruled out by the case assumption, as analyzed earlier. Similarly, one can show either $\lambda_1^{B_H} = 1$ or $\lambda_1^{B_L} = 1$. Therefore, we only need to examine the following possibilities:

- $\boldsymbol{\lambda}_1 = (y, 1, 1-y, 1)$, $\boldsymbol{\lambda}_2 = (x, 1, 1, 1-x)$, and $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2$ are substitutes: x and y should maximize the combined value, which is:

$$v(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) = (x + y - xy)(1 - y + [y + (\lambda^* - 1)x]^+). \tag{IA.124}$$

  If the term inside the bracket is positive, we should have $y = 1$, and consequently $\boldsymbol{\lambda}_1 = \boldsymbol{\lambda}^{SP_A}$, which is analyzed earlier. Otherwise, $x = 1$.

- $\boldsymbol{\lambda}_1 = (y, 1, 1, 1-y)$, $\boldsymbol{\lambda}_2 = (x, 1, 1, 1-x)$, and $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2$ are substitutes: the combined value is

$$v(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) = (x + y - xy)(1 + (\lambda^* - 1)xy). \tag{IA.125}$$

  By examining the first-order conditions, we find that $x = y$ and they depend on λ^*. Thus, there is no robust equilibrium under this possibility.

REFERENCES

Adams, William James, and Janet L Yellen, 1976, Commodity bundling and the burden of monopoly, *The quarterly journal of economics* 90, 475–498.

Zhou, Jidong, 2017, Competitive bundling, *Econometrica* 85, 145–172.