An Economic Perspective on Diversity within Organizations

Working Paper #0109

Ashley Perry and Ernesto Reuben

NYU Abu Dhabi

May 2025

جامعة نيويورك أبوظبي

NYU ABU DHABI

An Economic Perspective on Diversity within Organizations

Ashley Perry
New York University Abu Dhabi, e-mail: ashley.perry@nyu.edu

Ernesto Reuben
New York University Abu Dhabi, Center for Behavioral Institutional Design, and
the Luxembourg Institute of Socio-Economic Research, e-mail: ereuben@nyu.edu

Abstract

This paper summarizes the theoretical and empirical research in economics on the impact of employee diversity on organizational performance, where diversity is predominantly viewed through the lens of gender and ethnicity/nationality. The literature has studied this topic through two types of interactions: horizontal interactions between workers who are peers and vertical interactions between managers and workers. The theory of horizontal interactions highlights the conditions under which diversity is beneficial, such as when different groups bring complementary knowledge, but also when it is costly, as in the case of intergroup communication frictions. The theory of vertical interactions focuses on discrimination against different social groups due to the hierarchical nature of these interactions. Discrimination can result from preferences that favor or disfavor certain groups, or from imperfect information and the use of beliefs about group productivity to infer individual productivity. For both horizontal and vertical interactions, the empirical findings on the impact of diversity are mixed, with varying effects on organizational performance, ranging from positive to negative to no impact. Although this may suggest the field has little to say on the subject, the impact of diversity is often in line with the theoretical predictions. This suggests that for organizations to reap the potential rewards from diversity, they must consider how their context relates to the theory and act accordingly.

Acknowledgments: This is the authors' current version of a chapter on diversity that will be part of a volume titled "Interdisciplinary Foundations for Organizational Science and Application: A Dialogue between Psychology and Economics." This volume will appear in a book series titled "The SIOP Organizational Science, Translation, and Application Series," co-sponsored by the Society of Industrial/Organizational Psychology and Oxford University Press. We gratefully acknowledge financial support from Tamkeen under the NYU Abu Dhabi Research Institute Award CG005.

Since the seminal work of Becker (1957), economists have been interested in the economic effects resulting from the interaction of different social groups. A central question in the literature has been the effect of the composition or diversity of the workforce on firm performance. This question has been studied both theoretically and empirically. Although diversity could relate to a wide variety of traits, the literature has focused predominantly on observable characteristics, such as gender and ethnicity/nationality. In this paper, we briefly describe the theoretical models used by economists to study the impact of diversity on organizational performance and empirical studies that shed light on this question. The aim of the paper is not to provide a comprehensive review of the literature but to give a general overview of the main approaches and findings.

Broadly speaking, the impact of diversity within organizations has been studied through two types of interactions: horizontal and vertical. Horizontal interactions correspond to situations where people work with each other in teams or projects at similar hierarchical levels, meaning without significant power asymmetries between employees. By contrast, vertical interactions correspond to situations where there are clear hierarchies, giving some employees (e.g., supervisors) authority over others. Although these two types of interactions do not encompass all possible ways employees interact within organizations, they represent many common scenarios, and, as we discuss later, diversity has distinct effects across these types of interactions, affecting organizational performance and overall efficiency in different ways.

1. Diversity in horizontal interactions

Economic theory identifies various mechanisms, both positive and negative, through which diversity impacts horizontal interactions within organizations. On the positive side, diversity enhances the information and skills available to a team. In particular, diversity is beneficial when the information, knowledge, or skills individuals from different social groups bring to their team are *complementary* (Lazear, 1999b; Prat, 2002). More precisely, information and skills across social groups must be both relevant and disjoint, meaning individuals from one social group possess information and skills pertinent to the task at hand that others from different groups do not. These differences can arise from homophily, where individuals from similar backgrounds are more likely to share similar experiences (Jackson, 2008). For instance, individuals from specific ethnicities are likely to have unique knowledge about the preferences

of consumers from their ethnic group because they were exposed to different institutions and social norms during their upbringing. In this context, the benefit of workplace diversity is increasing the likelihood that information and skills within teams are disjoint.

On the negative side, there are several potential costs to diversity. One of these costs is that, even if the information is complementary, effectively communicating and sharing this information between social groups can be challenging (Lazear, 1999a; Lang, 1986). A stark example is when groups do not speak the same language. However, difficulties communicating can easily arise, even in the presence of a common language, as different groups may develop different understandings of what a word or concept means. For example, the term "efficiency" has a very specific meaning for economists, which differs from its common usage or its interpretation in other fields. Similarly, the concept of "evolution" to biologists differs from its colloquial use. These communication frictions can negatively impact group work as it can lead to important information being "lost in translation" and to less productive teamwork.

In addition to communication costs, there are other potential costs of increased organizational diversity. One such cost is the increased uncertainty about how others will behave in strategic interactions. Diversity can increase strategic uncertainty because individuals can typically predict the behavior of members of their own social group more accurately than that of others (Kets and Sandroni, 2021). Group norms contribute to this, as different groups may have different norms for reacting to certain situations. Increased strategic uncertainty can result from a lack of knowledge about the norms existing in different groups, or even if there is common knowledge about these norms, a multiplicity of norms can create uncertainty about which norm should apply in a particular team or situation. Another cost of diversity is that it can dampen individuals' social preferences, which guide behaviors crucial to organizations, such as individuals' willingness to reciprocate and cooperate even in the absence of formal enforcement mechanisms (Fehr and Gächter, 2000). It is well-documented that individuals' social preferences vary when interacting with in-groups compared to out-groups (Chen and Chen, 2011).[1] Both of these costs, strategic uncertainty and social preferences, affect team cohesion, and in particular, the trust between team members, which is crucial for team performance.

[1]There is evidence that sharing a common identity in a group increases the level of cooperation (e.g., Eckel and Grossman, 2005), coordination (e.g., Chen and Chen, 2011), and trust (e.g., Falk and Zehnder, 2013).

2. Evidence on the effects of diversity in horizontal interactions

The evidence in the economics literature on the effect of diversity in horizontal interactions is mixed, making it challenging to definitively argue that diversity is always beneficial or detrimental to teamwork. The methods used to identify the impact range from observational studies exploiting quasi-random variation in group composition to controlled randomized field experiments. In addition to variation in methods, there is substantial variation in the contexts of these studies, with evidence coming from several settings such as professional ice hockey, online computer programmers, factory workers, and civil servants.

Several papers examine how the diversity of student teams impacts individual and team performance in university courses. Contreras et al. (2022) conducted a randomized trial with master's-level students at a university in the United Kingdom, looking at student diversity in terms of gender and whether English was their native language. They randomly assigned students to seminar classes and study groups, varying the percentage of women and those who indicated English was their native language. Their findings suggest that the composition of seminar classes or study groups, whether by gender or native speaker status, does not affect individual exam grades. This null effect could be due to the costs of diversity canceling out its benefits. Alternatively, since the output measured was an individual essay rather than joint work, individual factors like ability might be much more influential. However, they found that reducing gender diversity among non-native English speakers improved grade performance. Since this effect was not present among native English speakers, where communication costs were presumably lower, it suggests that the negative effect of diversity was due to reduced group cohesion in non-native English-speaking teams.

Hoogendoorn et al. (2013) also conducted a randomized control trial with student teams who 'run' a startup as part of an entrepreneurship course. This course is designed to replicate the experience of running a real-world company: raising funds, selling genuine products or services, and maintaining financial records. They randomly varied the gender composition of these teams. Using sales and profit data, they found a non-linear relationship between performance and the gender diversity of teams. The relationship is an inverted-U shape, where gender-diverse teams performed better than gender-homogeneous teams and teams having an equal gender split performed the best. The fact that gender diversity was beneficial suggests that each gender brought complementary knowledge to the task.

In a related study, Hoogendoorn and van Praag (2012) varied ethnic diversity in the same course and found non-linear effects between diversity and performance, though the results differed slightly. Ethnic diversity had no effect when there was a dominant ethnic group (Dutch nationals). However, when Dutch team members became a minority, ethnic diversity had a positive effect on performance. They surveyed students to try to identify the factors explaining the effects of diversity. They did not observe evidence that diversity was related to communication costs but found evidence suggesting that more diverse teams had more complementary skills. Since the benefit of diversity occurred only once nationals no longer constituted a majority of the team, it suggests that the positive effects of diversity on performance are realized only when the costs related to group cohesion are minimized.

Whereas the previous studies report some positive effects of diversity, Calder-Wang et al. (2021) do not find such positive outcomes in a similar setting. They studied MBA students in an entrepreneurship course from 2013-2016, but unlike the other studies, students only pitched an idea rather than 'ran' a company. The teams' performance was based on their progress in developing their pitch. For the 2013 cohort, students were randomly assigned to teams based on four dimensions: gender, ethnicity, school background, and industry experience. They measured diversity along these four dimensions. For each dimension, they constructed a measure of diversity as the ratio of the number of ties between team members with the same characteristic to the total number of possible ties in a team. They did not find any effect of gender, school, or industry diversity on performance, which contrasts with the findings of Hoogendoorn et al. (2013) regarding gender diversity. However, they did find that an increase in the ethnic diversity of a team reduced team performance, which is the opposite finding of Hoogendoorn and van Praag (2012).

They further analyzed their results by creating a composite ethnicity-gender diversity score (e.g., those with the same gender and ethnic group are considered to have a tie) and found that teams with the greatest diversity along both characteristics were the ones that performed the worst. The 2014-2016 cohorts were not randomly assigned to teams. Instead, students were allowed to choose their team members. Using the same measure of ethnic diversity, they still found a negative effect on team performance, but the effect was half the size of the effect in the randomized 2013 cohort. This suggests that the choice of team members can alleviate the costs of diversity, although it is unclear whether this is because teams can identify individuals who minimize communication costs or help with social cohesion. However, some caution should be

used when comparing these studies since Calder-Wang et al. (2021) assume a linear relationship between team performance and their measure of diversity, whereas Hoogendoorn and van Praag (2012) and Hoogendoorn et al. (2013) allow for non-linear specifications.

Outside of classroom settings, several observational studies use non-random variation in group composition to try to identify the effect of diversity on team performance. Kahane et al. (2013) examined NHL team performance in the 2000s, exploiting the natural variation in the share of European players within teams, which varied from very few to almost a third of the squad. They found that an increase in the share of Europeans in a team was associated with better league performance. Consistent with the theory on the benefits of diversity, North American and European players have complementary skill sets due to how ice hockey is taught at the youth level. In North America, the focus is on learning by playing, whereas in Europe, the focus is on skill development. Hence, North Americans tend to be more physical, which complements Europeans being more skillful. Tellingly, Kahane et al. (2013) also show that the fewer different countries the European players come from, the better a team performs. Specifically, while North American players do not benefit from the increased presence of Europeans in the team, European players perform better with more players from their home country, which is consistent with the existence of communication costs between players of different nations.

Using a similar empirical strategy in a very different context, Rasul and Rogger (2015) show that diversity improves performance in the Nigerian civil service, a country with many ethnolinguistic groups. They used a large administrative dataset on the ethnic makeup of departments working on various projects and the corresponding project completion rates. Exploiting the natural variation in the concentration of ethnic groups across departments, they found a strong positive correlation between increased department diversity and project completion rates. They also find significant heterogeneity by ethnicity in workplace experience and beliefs about work, which they suggest indicates that the benefit of diversity is realized in this context through the complementary skills different groups bring to a project.

As in the public sector, there is evidence that diversity can be beneficial in the private sector as well. Calder-Wang and Gompers (2021) use the variation in the percentage of women hired by venture capital funds to determine how gender diversity impacts fund performance as measured by startups that successfully IPO or are bought out at a value above the initial investment. They use the number of female children of incumbent partners in a fund as an instrument to predict the percentage of females hired. They find a positive association between

the percentage of females hired and fund performance. While they cannot identify the exact mechanism, they suggest that greater gender diversity in a male-dominated industry could either reduce the likelihood of correlated errors in judgment or improve the quality of deals the fund can attract. This is because female employees tend to have different backgrounds and skills from their male counterparts, which are useful to the objectives of venture capital funds. Under these conditions, as the theory predicts, diversity is beneficial. However, caution is warranted when interpreting these results, as the authors cannot entirely rule out the possibility that the number of female children of incumbent partners impacts fund performance through channels other than the hiring of female employees.

In a field experiment designed to quantify the price of discrimination, Hedegaard and Tyran (2018) recruited participants to work stuffing envelopes initially by themselves and subsequently with another participant of their choosing. They recruited school students with typically Danish-sounding and Muslim-sounding names. They find no difference in the number of envelopes stuffed by homogeneous teams compared to diverse teams. However, potential selection effects may confound this result since participants could choose their partners, presumably based on how well they think they work in homogeneous and diverse teams. Additionally, the envelope-stuffing task requires little teamwork. Hence, finding a null effect is unsurprising, given that there is no clear mechanism through which diversity could be beneficial.

Lyons (2017) use an online field experiment to explicitly test the performance of diverse and homogeneous teams. The experiment recruited contractors to complete a web programming task from the online platform oDesk, which companies use to hire contractors for computer development tasks. These contractors were randomly assigned to teams of two to complete the task, with team compositions being either homogeneous (of the same nationality) or diverse (of different nationalities). Teams were given three features to add to a website: one with JavaScript code, one with PHP code, and one with both languages. Relative to the number of features successfully added by individuals, homogeneous teams showed increased performance, whereas diverse teams showed decreased performance. It looks like a lack of disjoint skills across social groups and communication costs in diverse teams drove these outcomes. First, the likelihood that a pair of workers had disjoint programming knowledge was the same for homogeneous and diverse teams, indicating that the complementarity of information was not a function of diversity. Second, communication was less clear in diverse teams compared to

homogeneous teams, as evidenced by individuals in diverse teams misreporting what features had been added and sending more messages in the chat logs.

In different context Marx et al. (2021) conducted a field experiment with canvassers for voter registration over two weeks in Kenya before the 2013 general election. Canvassers were randomly paired and allocated households to visit and provide information on voter registration in Kibera, a large slum in Nairobi where voter registration is low. The pairs were considered diverse if they did not belong to the same ethnic group, which was the case for roughly 77% of pairs. They found that homogeneous teams performed better than diverse ones, visiting more households and spending longer during each visit. It appears that co-ethnic team members trusted each other not to slack off and, therefore, were more likely to split up and visit households separately. Furthermore, the performance benefit to homogeneous teams increased with the number of days pairs worked together over the two weeks. Language did not appear to be an impediment in this setting, as most spoke English or Swahili, and there was no difference in who got to speak their preferred language between homogeneous or diverse pairs. This suggests that diversity between pairs was costly not because it increased communication costs but because it reduced team cohesion through lower levels of trust.

In summary, the empirical evidence from economics on the impact of workplace diversity on teamwork is mixed, with some studies finding positive effects of diversity through skill complementarity, while others report negative effects due to costs and reduced team cohesion. These findings indicate that the impact of diversity on team performance is contingent on several factors, such as the nature of the task, the degree of interdependence among team members, and the existing social and communication dynamics within teams. Therefore, while diversity has the potential to enhance organizational outcomes, its successful integration requires careful consideration of these contextual elements to minimize potential costs and maximize its benefits. In this respect, it is interesting that there appears to be an association between the benefits of diversity and financial incentives. Studies that found a negative effect of diversity typically had fixed-pay structures, whereas studies reporting positive effects were more likely to involve performance pay (e.g., ice hockey players receiving bonuses for winning). This suggests that some of the frictions present in diverse teams might be overcome if the financial incentives are appropriately designed.

3. Diversity in vertical interactions

Vertical interactions may include elements of teamwork, meaning that the mechanisms described above can also be at play. However, the hierarchical nature of the relationships and responsibilities sets vertical interactions apart from horizontal ones. Those higher up in the hierarchy, such as supervisors, have the explicit ability to influence their subordinates through task assignments, evaluations, or promotion and firing decisions. This asymmetry introduces additional ways for diversity to impact organizational performance. Here, we concentrate on discrimination in vertical interactions within organizations, defined as the differential treatment in the workplace of employees from different social groups.

One of the canonical economic models of discrimination was proposed by Becker (1957), where a *preference* favoring or disfavoring particular social groups is the reason behind someone's differential treatment. This approach is known as taste-based discrimination. For example, a CEO who dislikes interacting with individuals from ethnic minorities might prefer to promote to the C-suite an employee from the ethnic majority, even if more productive employees from ethnic minorities are available. Taste-based discrimination can be costly to organizations when discriminatory preferences lead to the favoring of less productive employees.

The second canonical approach to discrimination, introduced by Phelps (1972) and Arrow (1973), explains differential treatment through imperfect information about individuals' productivity. When workers' productivity is unknown or observed imperfectly, supervisors may infer an individual's productivity based on the productivity distribution of their social group. This model is known as statistical discrimination since differential treatment is the consequence of different but nonetheless accurate beliefs about the productivity of different social groups. For instance, a manager promoting someone for a position that requires constant availability might discriminate against female employees, assuming they will be less available than men because of the general empirical observation that women are often the main caregivers within households. Statistical discrimination can also be costly to organizations due to the misallocation of skills or abilities. In the example above, the manager might promote a male employee with family responsibilities over a female employee with more availability.

More recent models have questioned the conventional assumption in statistical discrimination models that individuals hold accurate beliefs about the productivity distribution of different social groups (Bohren et al., 2023; Campos-Mercade and Mengel, 2024). Instead, these

newer models consider the impact of inaccurate beliefs, which deviate from the true productivity distribution due to stereotypes (Reuben et al., 2014; Bordalo et al., 2019). These inaccurate beliefs have new policy implications and can lead to unexpected outcomes in dynamic settings, such as discrimination *favoring* a social group as a direct result of historical discrimination *against* it.

Finally, in the models of discrimination discussed so far, discrimination directly results from decisions made by individuals at higher levels in a hierarchy. Other models have explored the role of indirect or systemic factors beyond an individual's control (Bohren et al., 2022; Campos-Mercade and Mengel, 2024). In these models, the institutions and strategic environment in which individuals are embedded determine the presence and persistence of discriminatory behavior. For instance, negatively biased beliefs about a minority group can lower *both* supervisors' expectations of the performance of minority workers and the effort those workers are willing to exert due to the supervisors' expectations. This self-fulfilling cycle can lead to discrimination as the outcome of an equilibrium that individuals alone cannot break. To disrupt this cycle, centralized policies are necessary to change the "rules of the game" or enable coordinated actions.

In practice, it is challenging to identify the exact reasons behind discriminatory behavior. Take, for example, correspondence studies, which examine differences in callback rates for resumes that are identical except for the social group of the job candidates (for a review of these studies, see Bertrand and Duflo, 2017). When these studies find discrimination, evidenced by differential callback rates, it is often unclear whether it stems from employers' tastes or their beliefs about productivity. Hence, in some of the research we discuss next, the best we can do is highlight where evidence of differential treatment occurs.

4. Evidence on the effect of diversity in vertical interactions

As with horizontal interactions, diversity in vertical interactions has been studied in classroom settings. For example, Calder-Wang et al. (2021) study the same MBA teams tasked with preparing a pitch for a startup, but focus on the impact of vertical interactions. They leveraged the exogenous allocation of teams to section leads−—professors responsible for guiding students through the course and assessing their performance. The study found that the gender composition of teams did not affect performance when advised by male section leads. However,

increasing the number of women positively impacted team performance when advised by female section leads. While the motives of section leads are unknown, one possible explanation is that female section leads, being aware of negative gender stereotypes in business, provided more support to teams with more females. Alternatively, female section leads may prefer to work with majority-female teams, whereas male section leads do not have such preferences.

Outside classroom settings, Hjort (2014) examine the effect of assigning workers of different ethnicities to teams in a vertical production process. In a Kenyan plant, a single upstream worker passes flowers to two downstream workers who create bunches, an arrangement mirroring a manager assigning work to subordinates. To study the impact of diversity, they exploit the quasi-random assignment of workers to homogeneous teams, where all workers are co-ethnic, or diverse teams, where downstream workers do not share co-ethnic ties with upstream workers. They found that the output in homogeneous teams was higher than in diverse teams. The negative effect of diversity appeared to be driven by a preference for in-groups (i.e., taste-based discrimination), as performance differences between homogeneous and diverse teams increased after the 2007 contentious election, which led to widespread ethnic conflict. Interestingly, the negative impact of diversity was ameliorated by changes to the financial incentives for workers. Specifically, introducing team pay instead of piece-rate pay for downstream workers improved the performance of diverse teams.

In the canvassing study of Marx et al. (2021), in addition to studying horizontal interactions, they examine the impact of diversity on vertical interactions. They randomly assigned a supervisor to canvassing teams. As Hjort (2014), the diversity of supervisor-team pairs is defined by the number of co-ethnic links between supervisors and canvassers.[2] They found that ethnically diverse teams performed slightly better than more homogeneous teams, spending more time at each household despite visiting a similar number of households. That a positive effect is found in this setting but not in Hjort (2014) suggests that context matters. In this setting, supervisors were responsible for monitoring rather than assigning tasks, resulting in a different manifestation of in-group preferences. It appears that supervisors monitored more homogeneous teams less stringently than diverse ones, resulting in homogeneous teams spending less time in the field.

[2]Due to the small sample of teams and the large number of Kenyan ethnic groups, there were no instances where the supervisor and both canvassers were all co-ethnic. Hence, comparisons are between diverse teams without any co-ethnic links and more homogeneous teams where the supervisor was co-ethnic with one canvasser.

Bandiera et al. (2009) conduct a field study that, while not explicitly about the effect of diversity, provides valuable insights into the topic. They study fruit pickers in the UK, investigating how the social connectedness between managers and workers impacts productivity. Managers were responsible for assigning workers to specific locations in the field, and hence, they could send workers to parts of the field where fruit was more plentiful. A manager and worker were considered socially connected if they shared the same nationality, joined the fruit-picking company in the same cohort, or lived in close proximity at the on-site accommodation. Since nationality was the primary source of connection, their findings shed light on the effects of diversity in vertical interactions on performance. In their study, managers were randomly assigned to workers and to either a fixed wage or performance-based remuneration scheme. Under the fixed pay regime, managers tended to assign the most productive fields to workers with whom they were connected, resulting in these workers picking more fruit. However, since field assignment depended on a factor unrelated to workers' productivity, the overall amount of fruit picked decreased with connectedness. These findings are consistent with those of Marx et al. (2021), suggesting that diversity in vertical interactions can increase organizational performance by reducing taste-based discrimination that favors in-group workers at the expense of overall productivity. Consistent with this interpretation, the effect of diversity was eliminated when managers' incentives were aligned with those of the firm through performance-based pay.

Other work has focused on settings where differential treatment is attributed to statistical discrimination. For instance, Sarsons et al. (2021) investigate promotion decisions by analyzing tenure outcomes in the top 30 US economics departments. As expected, the probability of getting tenure increases with the number of publications. However, this probability varies between men and women depending on whether the publications were written alone or with co-authors. Specifically, co-authored publications increased the probability of tenure for men about twice as much as for women. Although women were not penalized for solo-authored papers, their overall disadvantage was substantial since co-authored papers were twice as common. The fact that there was discrimination against women for co-authored but not for solo-authored work makes taste-based discrimination unlikely. Instead, the authors propose that tenure committees held lower prior beliefs about women's productivity and, therefore, gave them less credit for co-authored publications, where individual contributions are uncer-

tain.[3] Bohren et al. (2019) also study performance evaluation by examining an online forum where users post math questions, which are then evaluated by others in the forum. The authors posted questions using usernames that randomly varied by gender and reputation. They found that posts by male usernames receive better evaluations than posts by female usernames if the usernames have no prior reputation. However, the reverse was true for usernames with positive reputations, where female usernames received more favorable evaluations. Once again, since the gender difference in evaluations changes depending on the username's reputation, it is unlikely that it is driven by tastes. A more reasonable explanation is that reputation influences evaluators' beliefs. We should note that not all papers show discrimination against women. For example, when Card et al. (2022) examined whether there is gender bias in electing fellows to the Econometric Society (one of the top honors for academic economists), they found that, for the last couple of decades, there has been discrimination in favor of women. Unfortunately, Sarsons et al. (2021) and Bohren et al. (2019) do not observe the gender composition of the evaluators. However, Sarsons et al. (2021) also present evidence from online experiments showing that gender difference in evaluations was driven primarily by male evaluators.[4] These papers reveal that statistical discrimination can occur not only during hiring, where employers have limited information about candidates but also in the evaluation of performance within organizations, leading to disparities in promotion decisions.

Recent evidence, such as the one reported by Sarsons et al. (2021) and Bohren et al. (2019), has called into question the assumption of accurate beliefs in statistical discrimination and has given rise to a growing literature on inaccurate statistical discrimination (Bohren et al., 2023; Campos-Mercade and Mengel, 2024). The deviation of beliefs from the true distribution of performance can arise due to over-reliance on general stereotypes (Reuben et al., 2014; Bordalo et al., 2019) and lead to counter-intuitive reversals in discriminatory behavior. For example, Bohren et al. (2019) developed a model showing how the dynamics of discrimination

[3]Suggestively, Sarsons et al. (2021) show that in sociology, where the contribution of co-authors is discernible by the order in which they are listed, there was no gender difference in the impact of a co-authored publication on the probability of tenure.

[4]A similar finding is reported by Moisan et al. (2024), who examine university students' beliefs about their classmates' mathematical and verbal abilities. They found that the students' beliefs were positively correlated with their classmates' performance in both subjects. However, for math, this correlation was much stronger for male classmates, indicating that high-performing women in math were more likely to be overlooked. The gender difference in beliefs about math ability was driven by male students.

depend on its underlying source. To illustrate this, imagine a faculty committee evaluating the abilities of a male and a female researcher based on letters of recommendation. If the committee's prior belief is that, on average, women are less capable than men and these beliefs are accurate and shared by the letter writers, then receiving equally positive letters for both researchers would narrow the committee's belief gap but not eliminate it. By contrast, if the committee's prior is that women are equally capable as men but knows that some letter writers inaccurately think men are better, then receiving equally positive letters would lead the committee to conclude that the female researcher is actually more capable. These predictions are not only consistent with the empirical results of Bohren et al. (2019) but also the findings of Sarsons et al. (2021), who show that female economists benefit more than male economists from solo-authored publications in the same journal. These findings illustrate how statistical discrimination at one stage can influence, or be influenced by, discrimination at different stages, offering a different interpretation of conflicting findings in the literature. Namely, discrimination favoring women at the top of the economics profession in the last two decades (as reported by Card et al., 2022) may result from increasing awareness of discrimination against women at earlier stages of their academic careers (as found by Sarsons et al., 2021).

In addition to recognizing the role of inaccurate beliefs, economists increasingly recognize that discrimination can be systemic, arising from the strategic environment and institutions individuals are part of. Glover et al. (2017) provide suggestive evidence of systemic discrimination in a French grocery store. In their study, the vast majority of managers (94%) belong to the majority group, while a substantial number of cashiers (28%) are from minority groups (as categorized by their names). They used the implicit association test to determine whether individual managers were biased toward minorities. Using scanning and timesheet data to measure performance and the quasi-random assignment of shifts, Glover et al. (2017) found that minority workers performed worse and had higher absenteeism when assigned to shifts with biased managers. A compelling explanation is that biased managers have low expectations and fail to recognize the good performance of minority cashiers. Anticipating discrimination, minority cashiers put in less effort when a biased manager is in charge, thereby confirming the manager's beliefs. Reuben et al. (2025) use a lab experiment to test this type of self-fulfilling system of statistical discrimination.[5] In the experiment, workers chose whether to improve

[5]De Haan et al. (2017) and Campos-Mercade and Mengel (2024) also found self-fulfilling statistical discrimination in experiments where groups were randomly assigned using the minimal group paradigm.

their performance in one of two tasks through training, knowing that managers in charge of task assignment observe only the workers' training decisions and gender. Gender stereotypes were introduced by providing information about past performance, with some groups seeing higher performance for men and others for women. Without stereotypes, workers trained in the task they were better at, and managers assigned tasks based on training, ignoring gender. With stereotypes, managers discriminated according to the stereotype when they assigned tasks between equally trained workers. Anticipating discrimination, workers trained in tasks stereotyped for their gender, making the stereotypes self-fulfilling. This led to inefficient outcomes as decisions were based on gender rather than ability. Illustrating the pervasive nature of systemic discrimination, Reuben et al. (2025) further show that discrimination persists even after it is known that there are no gender differences in ability. This persistence is because unilateral change is ineffective. Workers cannot ignore gender in their training decisions if managers continue to discriminate, and managers who do not discriminate will make poor assignments if workers are still basing their training decisions on gender. In addition to arising from strategic considerations, systemic discrimination can also occur indirectly. For example, requiring a certificate for a position can result in indirect discrimination if, due to historical reasons, access to certification has been unequal among different social groups (Bohren et al., 2022).

In summary, the evidence on the effects of diversity in vertical interactions highlights additional challenges and potential benefits for diverse organizations. Taste-based discrimination, where supervisors favor their in-group subordinates, can have different implications depending on the incentives and tasks available to supervisors. A diverse workforce can result in supervisors allocating tasks based on group affiliation rather than productivity, especially if their pay is not tied to subordinates' performance. Conversely, having subordinates from different social groups may reduce supervisors' leniency in monitoring. The literature on statistical discrimination highlights other challenges in a diverse workforce. Inaccurate stereotypes about the performance of certain social groups can lead to biased evaluations and promotion decisions. Moreover, when these groups anticipate discrimination, it can result in underinvestment and misallocation of talent. While the evidence is not extensive, various studies suggest that diversity among individuals higher up the hierarchy can mitigate these biases, as minorities are generally less likely to hold biased beliefs about the performance of subordinates within their social group.

5. Conclusion

In economics, the impact of diversity within organizations has been studied through two types of interactions: horizontal and vertical. Horizontal interactions occur between workers at similar hierarchical levels, while vertical interactions occur between workers at different hierarchical levels.

For horizontal interactions, economic theory suggests that diversity can be beneficial when the information or skills possessed by individuals from different groups are complementary (Lazear, 1999a; Prat, 2002). However, increased diversity also presents challenges. Diverse teams may face higher communication frictions, leading to information being "lost in translation" and impeding team effectiveness (Lazear, 1999b; Lang, 1986). Additionally, group cohesion may be reduced in diverse teams due to higher strategic uncertainty or weakened social preferences. (Kets and Sandroni, 2021; Chen and Chen, 2011)

Vertical interactions may enjoy similar benefits and face similar costs from diversity as horizontal interactions, but the hierarchical nature of these interactions introduces additional considerations. In particular, the potential for differential treatment or discrimination based on individuals' social groups. Discrimination can occur due to tastes: preferences in favor of or against a particular social group (Becker, 1957). The effects of taste-based discrimination manifest depending on how supervisors favor workers from their preferred group. For instance, in some cases, favoritism may be expressed as laxer monitoring, while in others as preferential task assignment. Discrimination can also occur due to imperfect information about individuals' productivity, known as statistical discrimination (Phelps, 1972; Arrow, 1973). Statistical discrimination can occur when those high up in the hierarchy believe that workers' productivity varies across social groups, regardless of the accuracy of these beliefs. Finally, discrimination can be systemic, emerging from the institutions or strategic environment in which individuals are embedded. Generally, when diversity in vertical interactions leads to discriminatory behaviors, it negatively impacts organizational performance through the misallocation of resources and talent.

Overall, the empirical findings on the effects of diversity in horizontal and vertical interactions are mixed, with evidence of positive, negative, and sometimes no discernible impact on productivity. These mixed findings could suggest that the effect of diversity is idiosyncratic. However, the mixed evidence often aligns with theoretical predictions. For instance, Lyons

(2017) finds negative effects of diversity on a programming team task. The lack of benefits from diversity in this setting is consistent with the observation that social groups do not have differential knowledge of programming languages. For organizations to reap the benefits of diversity, they must carefully consider the context in which they operate to maximize their benefits and minimize their costs.

References

Arrow, K. J. (1973). The Theory of Discrimination. In Ashenfelter, O. and Rees, A., editors, *Discrimination in Labor Markets*, pages 3–33. Princeton University Press, Princeton, NJ.

Bandiera, O., Barankay, I., and Rasul, I. (2009). Social Connections and Incentives in the Workplace: Evidence From Personnel Data. *Econometrica*, 77(4):1047–1094.

Becker, G. S. (1957). *The Economics of Discrimination*. University of Chicago Press, Chicago, 1st ed edition.

Bertrand, M. and Duflo, E. (2017). Field Experiments on Discrimination. In Banerjee, A. V. and Duflo, E., editors, *Handbook of Economic Field Experiments*, pages 309–393. Elsevier B.V., Amsterdam, NL.

Bohren, J. A., Haggag, K., Imas, A., and Pope, D. G. (2023). Inaccurate Statistical Discrimination: An Identification Problem. *Review of Economics and Statistics*, forthcoming.

Bohren, J. A., Hull, P., and Imas, A. (2022). Systemic Discrimination: Theory and Measurement. Working paper 29820, NBER.

Bohren, J. A., Imas, A., and Rosenberg, M. (2019). The Dynamics of Discrimination: Theory and Evidence. *American Economic Review*, 109(10):3395–3436.

Bordalo, P., Coffman, K., Gennaioli, N., and Shleifer, A. (2019). Beliefs about Gender. *American Economic Review*, 109(3):739–773.

Calder-Wang, S. and Gompers, P. A. (2021). And the children shall lead: Gender diversity and performance in venture capital. *Journal of Financial Economics*, 142(1):1–22.

Calder-Wang, S., Gompers, P. A., and Huang, K. (2021). Diversity and Performance in Entrepreneurial Teams. Working paper 28684, NBER.

Campos-Mercade, P. and Mengel, F. (2024). Non-Bayesian Statistical Discrimination. *Management Science*, 70(4):2549–2567.

Card, D., DellaVigna, S., Funk, P., and Iriberri, N. (2022). Gender Differences in Peer Recognition by Economists. *Econometrica*, 90(5):1937–1971.

Chen, R. and Chen, Y. (2011). The Potential of Social Identity for Equilibrium Selection. *American Economic Review*, 101(6):2562–2589.

Contreras, V., Orsini, C., and Ozcan, B. (2022). Effects of team diversity on performance, perceptions,

and predictions: Experimental evidence of gender composition and language. Social policy working paper 06-22, LSE Department of Social Policy.

De Haan, T., Offerman, T., and Sloof, R. (2017). Discrimination In the Labour Market: The Curse of Competition Between Workers. *The Economic Journal*, 127(603):1433–1466.

Eckel, C. C. and Grossman, P. J. (2005). Managing diversity by creating team identity. *Journal of Economic Behavior & Organization*, 58(3):371–392.

Falk, A. and Zehnder, C. (2013). A city-wide experiment on trust discrimination. *Journal of Public Economics*, 100:15–27.

Fehr, E. and Gächter, S. (2000). Fairness and Retaliation: The Economics of Reciprocity. *Journal of Economic Perspectives*, 14(3):159–182.

Glover, D., Pallais, A., and Pariente, W. (2017). Discrimination as a Self-Fulfilling Prophecy: Evidence from French Grocery Stores. *The Quarterly Journal of Economics*, 132(3):1219–1260.

Hedegaard, M. S. and Tyran, J.-R. (2018). The Price of Prejudice. *American Economic Journal: Applied Economics*, 10(1):40–63.

Hjort, J. (2014). Ethnic Divisions and Production in Firms. *The Quarterly Journal of Economics*, 129(4):1899–1946.

Hoogendoorn, S., Oosterbeek, H., and Van Praag, M. (2013). The Impact of Gender Diversity on the Performance of Business Teams: Evidence from a Field Experiment. *Management Science*, 59(7):1514–1528.

Hoogendoorn, S. and van Praag, M. (2012). Ethnic Diversity and Team Performance: A Field Experiment. Working paper 12-068/3, Tinbergen Institute.

Jackson, M. O. (2008). *Social and Economic Networks*. Princeton University Press, Princeton, NJ.

Kahane, L., Longley, N., and Simmons, R. (2013). The Effects of Coworker Heterogeneity on Firm-Level Output: Assessing the Impacts of Cultural and Language Diversity in the National Hockey League. *Review of Economics and Statistics*, 95(1):302–314.

Kets, W. and Sandroni, A. (2021). A Theory of Strategic Uncertainty and Cultural Diversity. *The Review of Economic Studies*, 88(1):287–333.

Lang, K. (1986). A Language Theory of Discrimination. *The Quarterly Journal of Economics*, 101(2):363.

Lazear, E. (1999a). Culture and Language. *Journal of Political Economy*, 107(S6):S95–S126.

Lazear, E. (1999b). Globalisation and the Market for Team-Mates. *The Economic Journal*, 109(454):15–40.

Lyons, E. (2017). Team Production in International Labor Markets: Experimental Evidence from the Field. *American Economic Journal: Applied Economics*, 9(3):70–104.

Marx, B., Pons, V., and Suri, T. (2021). Diversity and team performance in a Kenyan organization. *Journal of Public Economics*, 197:104332.

Moisan, F., Muñoz-Herrera, M., and Reuben, E. (2024). Gender biases in job referrals. Working paper,

NYU Abu Dhabi.

Phelps, E. S. (1972). The Statistical Theory of Racism and Sexism. *American Economic Review*, 62(4):659–661.

Prat, A. (2002). Should a team be homogeneous? *European Economic Review*, 46(7):1187–1207.

Rasul, I. and Rogger, D. (2015). The Impact of Ethnic Diversity in Bureaucracies: Evidence from the Nigerian Civil Service. *American Economic Review*, 105(5):457–461.

Reuben, E., Rott, C., Hopfensitz, A., and Perry, A. (2025). Arbitrary gender stereotypes cause segregation in labor markets. Working paper, NYU Abu Dhabi.

Reuben, E., Sapienza, P., and Zingales, L. (2014). How stereotypes impair women's careers in science. *Proceedings of the National Academy of Sciences*, 111(12):4403–4408.

Sarsons, H., Gërxhani, K., Reuben, E., and Schram, A. (2021). Gender Differences in Recognition for Group Work. *Journal of Political Economy*, 129(1):101–147.