NBER WORKING PAPER SERIES

UNDERSTANDING PATENTING DISPARITIES VIA CAUSAL HUMAN+MACHINE LEARNING

Lin William Cong
Stephen Q. Yang

Working Paper 34197
http://www.nber.org/papers/w34197

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 2025

We are especially grateful to Jillian Grennan, Gerard Hoberg, Po-Hsuan Hsu, Kenneth Huang, Ke Shi (discussant), and Ting Xu for detailed feedback and suggestions. We also thank Waverly Ding, Josh Lerner, Asaf Manela, Yifei Mao, Amit Seru, Xuan Tian, Jun Yang, Liyan Yang, and conference and seminar participants at the 2025 China Financial Research Conference (Beijing), China International Conference of Finance (Shenzhen), Cornell University DEFT Lab, 2025 Forum on FinTech and Behavioral Finance at the Shanghai Jiaotong University, 2025 Global Entrepreneurship and Innovation Research Conference, the 5th Machine Lawyering Conference at CUHK, Nanyang Business School Finance Seminar, and NTU Dean's Distinguished Public Lecture for helpful comments and discussions. We acknowledge Tanvi Kiggar, Rishi Kumar, Jordan Velte, Jonathan Braun, Danish Safdariyan, Vineet Saraf, Lauren Sutioso, and Steven Yu for outstanding research assistance. Send correspondence to Cong at will.cong@cornell.edu. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Understanding Patenting Disparities via Causal Human+Machine Learning
Lin William Cong and Stephen Q. Yang
NBER Working Paper No. 34197
September 2025
JEL No. G30, J15, J16, O31

ABSTRACT

We develop an empirical approach for analyzing multi-dimensional discrimination using multimodal data, combining human perception measures with language-embedding-based, nonlinear controls for latent quality to relax restrictive assumptions in causal machine learning. Applying it to the U.S. patent examination process, we find that, ceteris paribus, applications from female inventors are 1.8 percentage points less likely to be approved, and those from Black inventors are 3 percentage points less likely—inconsistent with legally prescribed criteria. Jointly studying multiple bias dimensions and their intersections for the first time, we uncover new biases, including an affiliation bias—individual inventors are disadvantaged by 6.6 percentage points relative to employees of large, public firms, a disparity larger than any demographic gap. Moreover, innovation quality, location, and other factors can mitigate or compound discrimination, and the disparities interact: for example, racial gaps vanish among public-firm employees, masking more severe discrimination against individuals. Existing theories such as homophily cannot fully explain the results, but a simple model of correlation neglect does.

Lin William Cong
Cornell University
and NBER
will.cong@cornell.edu

Stephen Q. Yang
Stanford University
yang29@stanford.edu

1. Introduction

Racial and gendered disparities in economic outcomes are global, well-documented, yet persistent problems (e.g., Goldin, 2014; Bertrand, Kamenica, and Pan, 2015; Blau and Kahn, 2017; Chetty et al., 2014, 2020; Derenoncourt and Montialoux, 2021), given their unfairness, hindrance to growth, and distortion on talent allocation (Altonji and Blank, 1999; Bertrand, 2011; Hsieh et al., 2019). In particular, innovation and entrepreneurship ought to bridge social groups, yet many already question whether they are largely a false hope due to discrimination by human decision-makers (e.g., Chernenko and Scharfstein, 2022; Fairlie, Robb, and Robinson, 2022; Wang, Wu, and Hitt, 2024). Beyond innovation, many other decisions that allocate scarce resources, such as loans or jobs, also rely on decisions made by human reviewers reading unstructured, textual applications. Identifying robust causal evidence of such discrimination is challenging in both policy and research because "quality" is latent and text is high-dimensional. It is harder still when multiple protected traits interact. Further, in such settings, disparities across race, gender, or affiliation may reflect two distinct forces: (i) long-term structural inequities that shift average quality, and (ii) *disparate treatment*, where reviewers put weight on applicant characteristics that are outside legally prescribed criteria. Distinguishing the two situations is critical, as while both are unjust, addressing them likely requires substantially different responses.

The so-called causal machine learning (Causal ML) methods show great promise for estimating contrasts of potential outcomes under different treatments or interventions, and have found fitting applications in economics (e.g., Hochberg et al., 2023). But a general adaptation of them for studying multi-dimensional discrimination is lacking. Moreover, they identify causal patterns from observational data typically only under strong "ignorability" assumptions, which preclude wider applications and robust conclusions. Specifically in our context of patenting approvals, Causal ML assumes that correlation between a discriminating dimension (e.g., gender) and innovation quality is *fully captured* by the language model, which may not hold and is often impossible to verify.

With this in mind, we develop a systematic and robust framework for identifying and quantifying unfair disparities under human review: causal human+machine learning (C-HML). We then apply C-HML to patenting and conduct the first simultaneous investigation (to our knowledge) of discrimination across multiple dimensions, assessing their relative scales and intersectionality. We clarify why standard Causal ML is insufficient in this setting, and provide novel, strong causal evidence that patenting discrimination operates through inventor names. Our analysis also uncovers a novel and economically significant affiliation bias that disadvantages individual inventors more than any single demographic factor. Finally, we propose correlation neglect as a new theoretical mechanism that parsimoniously rationalizes how these biases arise. Its additional empirical predictions are also consistent with the data.

The critical innovation of C-HML is to use decision-makers' learning (i.e., *perceptions*) of applicant characteristics, in addition to applying multimodal, nonlinear tools from Causal ML for estimating true patent quality. Importantly, perceived applicant characteristics, beyond the actual characteristics, are plausibly not correlated with merit or quality. For example, if one woman is named "Alex" and the other "Alexandra," *ceteris paribus*, they should receive the same outcomes (hiring, loans, patent acceptances). If "Alexandra" receives worse outcomes, then the human reviewer appears biased.[1] Because whether or not a name is obviously gendered or racialized is an intrinsic non-economic property of the name, C-HML significantly relaxes the otherwise strong identifying assumption in Causal ML, making it closer to what economists would view as causal without just assuming away confoundedness. C-HML also improves over state-of-the-art Causal ML techniques, allowing for the study of interactions and multi-dimensional discrimination.[2]

C-HML has the potential to serve as the next-best substitute for field experiments involving perceptions, as pioneered by Bertrand and Mullainathan (2004) and applied by recent studies such as Kline, Rose, and Walters (2022), when they are infeasible for practical reasons. For example, submitting large quantities of real PPP loan requests under fictitious identities would be illegal. We demonstrate via simulations as well as a simple theoretical model that C-HML far exceeds the estimation accuracy of ML-only or Perceptions-only approaches when critical assumptions are violated to a large degree. This superior causal identification also comes from the "complementary identification" in C-HML, i.e., it still works even when one of the Causal ML or human perception assumptions completely fails to hold.

Having developed the C-HML framework, we apply it empirically to the patent examination

[1] C-HML, with an over-saturated design, is effective except in the unlikely case where assumptions for both human perception or machine learning approaches fail by a wide margin.

[2] Our C-HML builds on the ML measure of patent application quality in Yang (2025), which shares similar identifying assumptions with approaches like CausalBERT, functioning as a data-driven nonlinear control on application quality (which is both objective and rank-preserving). It differs from CausalBERT in implementation to allow for multi-dimensional analysis. This quality measure is a powerful predictor for patent outcomes and, as argued in previous Causal ML studies, captures the sole determinants of patent acceptance.

process. We choose patenting for several reasons. First, innovation is crucial to growth and the economy (Griffith, Redding, and Reenen, 2004; Bloom, Schankerman, and Van Reenen, 2013; Drucker, 2014). Second, patenting is a critical component of innovation, especially apropos social mobility (Akcigit, Grigsby, and Nicholas, 2017; Aghion et al., 2019). Despite previous studies investigating gaps in patenting activity, causal evidence is limited, and a unifying analysis of the relative effects and interactions of different forms of patenting discrimination is lacking—which we show theoretically to be necessary. Third, patent examiners (henceforth, examiners) see only inventor names, not true demographics, which is representative of human review processes and renders our perceptions approach exploiting names of ambiguous or "opposite" demographics relevant. Finally, at least in the United States, what constitutes discrimination in patenting is legally well-defined without the need for researchers to judge any welfare consequences of the examiners' actions.

We confirm empirically that minorities are indeed at a systematic disadvantage. Specifically, while quality differences add significantly to the aggregate under-acceptance of minority patents, the gap seems to be primarily a result of discriminatory treatment. Even when controlling for application quality including its nonlinear effects, Black and Hispanic inventors' applications are much less likely to be accepted than White and Asian & Pacific Islander (henceforth, API) inventors', with Black inventors experiencing the worst discrimination at 3 percentage points less relative to Whites. Meanwhile, male inventors experience an advantage of 1.8 percentage points, controlling for quality (including nonlinear effects).

Bias could also manifest in non-demographic characteristics such as firm affiliation, which is new to the literature—examiners may unjustifiably favor applications filed by inventors at, e.g., Microsoft, over individuals. We find a substantial gap between public firms and individuals, with the acceptance rate of public firm applications over 15 percentage points higher. This difference is not driven by skill only; we find significant causal discrimination and preferential bias for public firms reaching 6.6 percentage points, more than twice as large as any (one-dimensional) demographic gap. Patents are supposedly judged solely on their content, so there is no legal justification for why the same application filed by Apple instead of an individual ought to be accepted instead of rejected.[3] Moreover, discriminatory treatment against individual inventors could discourage

[3]This bias may be economically efficient from the perspective of total welfare, since the patent would generate more economic value in the hands of Apple. However, this is irrelevant for the purposes of patent approval. Legally, patentability depends on subject-matter eligibility and utility (35 U.S.C. §§ 101), novelty (102), and non-obviousness (103). Applicant identity is not a statutory criterion.

encourage innovation by broader subsets of the population.

We also document several other hitherto unstudied dimensions of discrimination. In particular, we find bias based on name frequency; inventors with rare names are at a substantial disadvantage. Moreover, inventors located in big cities experience a modest but significant advantage. While the average quality is indeed higher in big cities than in rural areas, examiners overweigh the importance of this distinction, giving an unfair advantage to better-located inventors.[4] Moreover, using examiners' characteristics, we document a homophily effect in race and last names. However, they do not fully explain our empirical results, given that female/Black examiners still exhibit significant bias against inventors of their own gender/race.

Returning to the outsized importance of affiliation bias, we next study cross-sectional variation in demographic discrimination across types of applicants. This is a unique benefit of our implementation of Causal ML relative to previous studies. The hypothesis is that if a discriminating examiner reviews a patent filed by a minority inventor working for Ford or Pfizer, they may be more prone to overlook their racial bias in favor of bias toward the company name. If examiners were intentionally racist, the effect would be homogeneous.

We find three levels of discrimination. In the case of public firms, racial treatment is much more equitable, with the exception that White inventors fare slightly better than all other races. However, for smaller private firms, the discrimination effect reappears, albeit at a less significant magnitude. Finally, for individual inventors, Black and Hispanic inventors are at even more of an unfair disadvantage than indicated by the entire sample. From a social mobility perspective, this means that prior studies systematically under-estimate racial disparities for individual inventors.

Unlike race, the gendered gap does not disappear for inventors associated with public firms. These surprising observations prompt us to conduct general tests for interaction effects of race and gender, which is also novel to the literature. The findings are nuanced: for example, among females, Black inventors experience somewhat less discrimination, while API inventors experience more of an advantage. Among White inventors, males experience a much larger advantage than in Black or API inventors. Meanwhile, in API inventors, gendered discrimination is substantially reduced: because these names are less obviously gendered to native English speakers, less discrimination occurs.

A natural interpretation of the findings is intentional discrimination; but one should not assert

[4] Differences in quality may be a result of either coagglomeration and concentration of firms (e.g., Ellison, Glaeser, and Kerr (2010)) or social connectedness (e.g., Bailey et al. (2018)).

this without clear scientific evidence. Instead, we demonstrate in a simple model that discrimination and biases may arise unintentionally as a result of *correlation neglect*. Because race, gender, and other factors are fundamentally correlated with quality, examiners mistakenly interpret these characteristics as an orthogonal source of information when the correlation is, in fact, already reflected in the patent's content. Thus, an examiner without intentional biases may still make discriminatory decisions. Our model provides a unifying framework better suited to settings involving examiner reviews like patenting. Correlation neglect predicts outsized discrimination for characteristics with higher explanatory power and lower perception noise, as well as for lower-quality and borderline applications, both of which we confirm empirically to be true.

Our study contributes to several strands of research. The first concerns economic discrimination against minority groups. Human reviewers are often known to be biased when conducting non-blind review (e.g., Goldin and Rouse (2000)). Bertrand and Mullainathan (2004), Chetty et al. (2014), Chetty et al. (2020), and Derenoncourt and Montialoux (2021) consider racial disparities in social mobility. Bertrand, Kamenica, and Pan (2015) and Blau and Kahn (2017) examine gender gaps in economic outcomes, mostly due to occupational and industry effects. Hsieh et al. (2019) demonstrates the economic inefficiency arising from the misallocation of talent as a result of discrimination. Concerning racial disparities in entrepreneurship and business, Aghion et al. (2017) and Bell et al. (2019) study how factors such as neighborhoods "nurturing" could explain the gap in patenting quality. Discrimination against minorities has also been prominent in early-stage start-ups, venture capital, and loans (Chernenko and Scharfstein, 2022; Fairlie, Robb, and Robinson, 2022; Fairlie and Robinson, 2023; Wang, Wu, and Hitt, 2024). We show yet another avenue of discrimination.

More specifically for disparities in patenting, Tabakovic and Wollmann (2018) and Coluccia, Dossi, and Ottinger (2023) both show potential racial discrimination. In gendered discrimination, Hunt et al. (2012) attempts to explain why the majority of inventors are male, while Schuster, Marcowitz-Bitton, and Gerhardt (2022) observes a lack of female participation in transfers from academia to industry and Jensen, Kovács, and Sorenson (2018) generally confirms worse proportional outcomes. Few papers unify these two directions. We seek to build upon Cook and Kongcharoen (2010) and Cook, Gerson, and Kuan (2022), who broadly investigate the "idea gap in pink and black," and provide a rigorous, extensive comparison and investigation of interactive effects.[5] Moreover, Chien and

[5] Lerner (2002) observes that some countries also historically prefer domestic over foreign applicants.

Grennan (2024) and Chien, Grennan, and Sandvik (2025) investigate voluntary participation in patenting, i.e., the "innovator-inventor gap," and related solutions. Brouillette (2025) models labor market discrimination, child penalties, and limited exposure to opportunities as barriers to female innovation. We directly study patent review, another potentially large barrier that compounds with bias at other stages of the innovation chain, given that women's accepted patents also receive fewer citations (Li and Liu, 2023), and their research is systematically under-referenced (Koffi, 2025).

Our paper also complements the literature studying examiners' limitations by demonstrating that examiners discriminate based on inventor names. For example, Frakes and Wasserman (2017) shows the variance in standards driven by time pressure. Many studies apply examiner assignment as an instrument, such as Farre-Mensa, Hegde, and Ljungqvist (2020) and Sampat and Williams (2019). Yang (2025) argues that examiners may prioritize applications with strong writing despite weak underlying technology.

The last point implies that our paper adds to the study of how names relate to economic outcomes. Fryer and Levitt (2004) show a potential alternative, non-causal explanation for the relation between names and socioeconomic outcomes. In the specific case of patenting, we find that distinctively racial or gendered names do lead to discrimination by examiners. This is consistent with a wealth of literature that shows the importance of race when an application undergoes review in the job market (Bertrand and Mullainathan, 2004; Bertrand and Duflo, 2017).

Importantly, we highlight, for the first time, the importance of applicant credibility (via firm affiliation) in patenting, in line with similar results in venture capital. Due to the high risk of entrepreneurship and innovation, decision-makers (venture capital investors) often pay close attention to the identity and credibility of the applicant (Gompers et al., 2020; Bernstein, Korteweg, and Laws, 2017). Particularly, Gompers, Lerner, and Scharfstein (2005) and Hacamo and Kleiner (2022) show that association with public firms can often serve as a strong signal for performance and predict funding outcomes. In venture capital, such decisions are likely justified. In contrast, the legal standards for patenting do not permit consideration of individual characteristics, but examiners nevertheless exhibit bias contributing to large disparities.

More broadly in terms of methodological innovation, this paper contributes to the application of machine learning to economics (Dell, 2025) and, specifically, bias and causal effects (e.g., Athey and Imbens, 2015; Chernozhukov, Newey, and Singh, 2022), especially involving texts, which is rare, as revealed in recent comprehensive surveys on the topic (e.g., Hoberg and Manela, 2025). Both Veitch, Sridhar, and Blei (2020) and Khetan et al. (2022) study embedding based techniques for causal inference. Hochberg et al. (2023) use similar techniques to demonstrate that female inventors receive less citations. We combine powerful LLM embeddings with human perception tests to not only extend causal text models in computer science—often just a nonlinear control assuming conditional independence—to a causal econometric framework, but also offer a timely and necessary study on the multiple dimensions and their interactions of patenting disparities for the first time.

The remainder of the paper is organized as follows: Section 2 describes our framework for modeling and causally identifying discrimination. Section 3 outlines the data sources and variables used in the empirical tests. Section 4 reports evidence of racial and gendered disparities in patenting, while Sections 5 and 6 uncover novel results on discriminating factors, their interactions, and heterogeneity. Section 7 concludes. The appendix contains the proofs of all propositions.

2. A Framework for Studying Discrimination in Human Reviews

2.1. Identifying (High-Dimensional) Causal Discrimination

Regulators' decision-making and discrimination. Many important economic outcomes depend on decision-makers' evaluation of an application submitted by an individual, ranging from online lending to mortgage applications to patenting. In all of these cases, the quality of the application (perhaps including personal characteristics such as credit score) may be the only legal determinant of outcomes. Quality may correlate with race or other demographic characteristics, and often does. Long-term structural inequities and other factors lead to, for example, aggregate differences in credit score. Nevertheless, a decision-maker should never reject an application that would have been accepted in the counterfactual with a different race or gender, *ceteris paribus*.

Clearly, it is difficult but imperative to extract the direct causal effect (if any) of discriminatory factors on outcomes through unfair regulatory/approval treatment. The econometrician's objective then lies in determining the influence (if any) of said factors on regulators' decisions *beyond* the extent to which they drive actual economic variables. Otherwise, it is impossible to distinguish *bias* from *structural inequity*.

For the ease of exposition, we now describe all variables within the context of patenting, but the findings remain relevant for general applications potentially subject to discrimination. Suppose that the merit, or quality, q of a patent application, i.e., whether it fairly deserves to be approved, is driven by two independent groups of factors such that $q = q_i + q_c$. q_i is determined by the idiosyncratic quality of the application, i.e., the ingenuity and quality of the idea itself. Meanwhile, q_c is driven by the individual characteristics such as race, gender, or socioeconomic background. Akcigit, Grigsby, and Nicholas (2017), Bell et al. (2019), and others document the importance of environmental factors in nurturing innovative skills. Also potentially included in q_c is current employment status: an inventor affiliated with Apple and with access to a successful patent lawyer, on average, produce higher-quality inventions. Notation is summarized in Table A.1 in the Appendix.

The examiner's legal obligation is to make decisions based on their judgment of q; essentially, they must attempt to estimate q by reading the application. In reading the application, the examiner obtains a noisy signal $s_1 = q+\epsilon_1$, with ϵ_1 independent from q, q_i, and q_c. Note that because examiners are time-constrained and must make subjective judgements, ϵ_1 can potentially be significant. Legally, only content-based assessment may be used, s_1. In addition, examiners can observe, through the application, a legally inadmissible signal $s_N = q_c + \epsilon_N$. If q_c depends on gender, for example, examiners can infer gender from the listed name of the inventor.

Legally, examiners must make a decision based only on s_1, i.e., their reading of the application. Modeling examiner decisions as $y = \beta_1 s_1 + \beta_2 s_N$, we can draw the following correspondences:

$\beta_2 = 0$ **(no discrimination):** The examiner indeed follows legal requirements and makes the same decision as if they had not observed the individuals gender, race, and so on.

$\beta_2 \neq 0$ **(discrimination):** A biased examiner discriminates, incorporating the characteristics of the individual into the decision.

Now, consider the role of the empiricist studying the examination process, who seeks to determine whether unfair discrimination is taking place. Optimally, one could conduct a large-scale randomized controlled trial, holding s_1 constant and varying s_N to observe outcomes for identical applications filed by, for example, inventors of differing genders. Nevertheless, most settings lack appropriate conditions for such experimental techniques. Thus, the econometrician must investigate the regulator's (lack of) biases using only ex-post observation of the examiner's ultimate decisions y. In other words, the econometrician must estimate β_2 without knowledge of s_1 and s_N, which are subjective signals the examiner determines based on priors, experience, and so on.

Causal ML for estimating β_2? An econometrician often has access to data regarding q_c. For concreteness, suppose q_c is gender. An econometrician may, in attempting to determine the presence of gendered discrimination, run the following regression:

$$y = \hat{\beta}_2 q_c + \epsilon. \tag{1}$$

Nevertheless, estimating such a model leads to biased results.

Proposition 1. *Even if $\beta_2 = 0$, $corr(y, q_c) \neq 0$ and $\hat{\beta}_2 \neq 0$.*

Essentially, regardless of true discrimination, the econometrician always observes an association between outcomes and gender, because fair quality q is related to q_c. Clearly, controlling for q would allow the econometrician to obtain a better estimate $\hat{\beta}_2$, but q is not directly observable.

Machine learning techniques present a potential solution to this identification issue. In patenting (as with most settings), the entire text of the application is available to the econometrician. By using a textual machine learning model to analyze the application, the econometrician can generate a noisy signal $q_{ML} = q + \epsilon_{ML}$, where the variance of ϵ_{ML} decreases with the quality of the model. If ϵ_{ML} is identically zero, then the machine learning model is perfect. Otherwise, let ϵ_{ML} be distributed with zero mean and variance σ^2_{ML}. With this signal, the econometrician can run the following model:

$$y = \hat{\beta}_1 q_{ML} + \hat{\beta}_2 q_c + \epsilon. \tag{2}$$

Proposition 2. *If $corr(\epsilon_{ML}, q_c) = 0$, the estimation error is given by the following expression:*

$$\hat{\beta}_2 - \beta_2 = \beta_1 \frac{\sigma^2_{ML}}{\sigma^2_{q_i} + \sigma^2_{ML}}.$$

If the machine learning model is highly accurate and unbiased (i.e., $\sigma^2_{ML} \approx 0$), then it essentially "controls" for quality, allowing the econometrician to identify any potentially illegal consideration of q_c and extract the value of β_2. However, σ^2_{ML} may not always be sufficiently small for accurate estimation of β_2. This can lead to an over-estimation of the effect and could even cause the identification of false positives even if $\beta_2 = 0$. In addition, should $corr(\epsilon_{ML}, q_c) = 0$ fail to hold, it is possible to not only over-estimate β_2, but also flip the sign, making it difficult to draw conclusive results.

The C-HML solution to estimating β_2. Under the above formulation of the econometrician's problem, it is indeed very difficult (absent perfect machine learning) to always extract an unbiased $\hat{\beta}_2$. However, let us return to the assumption above that the econometrician observes only q_c and not s_N. To be clearer in the case of inventor gender in patenting, q_c depends on the true gender of the applicant, while s_N is the examiner's opinion of the inventor's gender (since true demographics are not available). It seems that the subjective s_N should be unobservable, but, in fact, it is possible for the econometrician to approximate this belief ex-post.

Recall that the examiner uses s_N because q_c is not available, and the examiner is inferring characteristics from the application. Specifically, if examiners consider gender in their decisions, they must do so based on inventor names. Thus, to be more precise, s_N is actually $s_N = q_c + \epsilon_N$. By observing the names (which are also available ex-post), the econometrician can individually and carefully judge the impressions the names make on a reader, i.e., s_N. For gendered discrimination, this means the perceived gender.

Now, a critical point to notice is that ϵ_N, the difference in perceived from actual gender, ought to be non-economic in nature, and thus independent of other variables (specifically, q_i and ϵ_1). While many discriminatory factors such as race or gender are economic, the appearance of a name is in all likelihood not. That is, whether one has a gender-neutral name should not correspond to any differences in quality across applications, *holding gender constant.* With this in mind, the econometrician can incorporate s_N, the *human perception* variable into the regression model:

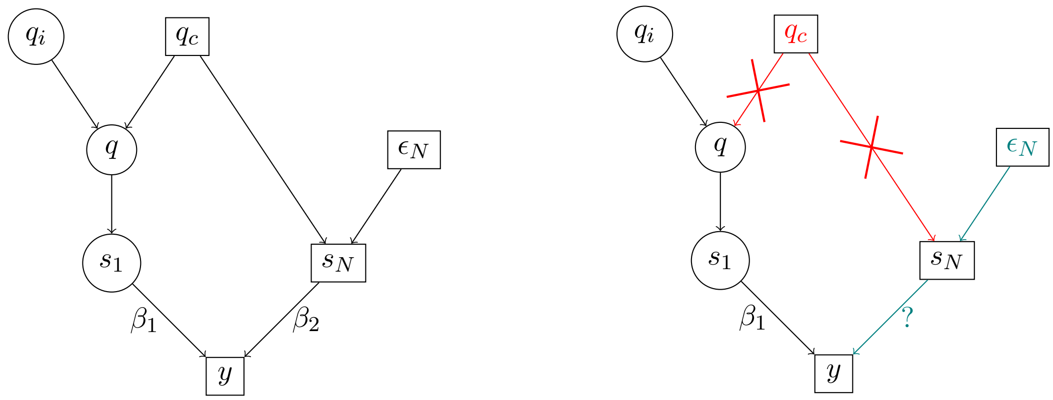
$$y = \hat{\beta}_1 q_{ML} + \hat{\beta}_2 s_N + \epsilon, \text{holding } q_c \text{ constant.} \tag{3}$$

Proposition 3. *If $corr(\epsilon_N, q_{ML}) = 0$ and $corr(\epsilon_N, q_i + \epsilon_1) = 0$, then $\hat{\beta}_2 = \beta_2$.*

Essentially, if the examiner is not discriminating and the correlation between y and q_c occurs only because of q_c's relation to q, then ϵ_N should have no effect on y. The key is that examiners only observe characteristics indirectly through the non-economic variable of names, which allows for a clean identification leveraging only the variation in ϵ_N.

Figure 1 visualizes these relations between variables in the model. Squares represent observable features (for the econometrician, not the examiner), while circles are latent. Again, as shown in Panel B, the econometrician eliminates any residual relation between s_1 and s_N by holding q_c constant.

Figure 1. Model Illustration



Panel A: Variable Relations

Panel B: ϵ_N Conditional on q_c

High-dimensional discrimination and mis-measurements. In reality, more than one discriminatory factor may constitute q_c. For example, race and gender could both simultaneously play into examiner decisions. Thus, we extend the toy model to enrich q_c. For illustration, let

$$q_c = \beta_G\, G + \beta_R\, R + \beta_{GR}\, (G \times R), \tag{4}$$

where G is a binary indicator for gender (e.g., 1 if female, 0 if male) and R is a binary indicator for race (e.g., 1 if non-White, 0 if White). Such a setup captures the separate effects of race and gender as well as intersectionality, which can potentially be of outsized importance.

Then, y may be written as

$$y = \beta_1 s_1 + \beta_2 s_N = \beta_1 q_i + (\beta_1 + \beta_2)\,(\beta_G\, G + \beta_R\, R + \beta_{GR}\, (G \times R)) + \beta_1 \epsilon_1 + \beta_2 \epsilon_N \tag{5}$$

$$= \beta_1 q_i + \gamma_G G + \gamma_R R + \gamma_{GR} (G \times R) + \beta_1 \epsilon_1 + \beta_2 \epsilon_N. \tag{6}$$

Now, a significant risk arises. Assume that a hypothetical econometrician has a completely robust identification strategy in access to a flawless machine learning model such that $q = q_{ML}$. Even in this perfect situation, the econometrician may still misidentify the presence or magnitude of bias when focusing only on one dimension, say gender, of discrimination and estimates:

$$y = \hat{\beta}_1 q_{ML} + \hat{\gamma}_G\, G + \eta, \tag{7}$$

If q_c depended only on gender (G), then the econometrician would extract the correct value of γ_G without issue as per Proposition 2. Nevertheless, the econometrician's approach is flawed because it ignores R and the interaction $G \times R$.

Proposition 4. *If the econometrician estimates a linear regression model considering only G, the estimated coefficient $\hat{\gamma}_G$ will be subject to bias such that:*

$$\hat{\gamma}_G - \gamma_G = (\gamma_R + \gamma_{GR}) \frac{\text{Cov}(G, R)}{\text{Var}(G)}. \tag{8}$$

If $\text{Cov}(G, R) \neq 0$, the estimated effect may be significantly biased and mis-estimated due to omitted variables.[6] The magnitude of the mis-estimation depends on the other discriminatory effects and the extent of the correlation, and thus may easily be large enough to *flip the sign*. Critically, without considering the other dimensions, there is no way for the econometrician to identify this bias or estimate its potential effects, *even given an otherwise perfect identification strategy*. The econometrician cannot accurately draw any conclusion whatsoever regarding the potential presence of gendered discrimination. In sum, if applications are affected by a combination of discriminatory factors, researchers ought to consider the high-dimensionality of discrimination, accounting for various contributing factors and intersectionality. Our C-HML approach enables exactly that.

2.2. The Causal Human+Machine Learning Approach

We now describe the details of C-HML, starting with its machine learning part. We also discuss how C-HML relates to and complements other methodologies.

Language-embedding-based quality measure. Recall that the ultimate goal empirically lies in distinguishing racial disparities driven by structural inequities from unfair discrimination on the part of the regulator. This is because while both present significant social challenges, they necessitate significantly different solutions. Isolating these two effects can be very challenging since it requires the determination of the merit of an application.

For patent applications, all legally relevant information to determine worthiness lies in the text of the application. Thus, we seek to construct an objective measure of patent quality based on the

[6]Why should gender and race be correlated? Obviously, in the population, this is not necessarily the case. But inventors are not a random sample of the population and equilibrium entry is affected by a variety of factors. Both gender and race play a role in this, e.g., White men may have the easiest path to innovation in the United States.

contents of the application. With such a measure as a control, potentially entering in nonlinear ways, we may effectively test a natural experiment by observing the outcomes of equivalent applications (in quality) filed by inventors from different demographic groups. The identifying assumption here is that the measure must accurately determine quality and that the text contains all relevant information, as in Proposition 2.[7] The textual information in a patent application is very specific and high-dimensional, so we take an embedding-based approach to capturing quality, as similarly done in Yang (2025).

We first take a large language model (LLM) embedding (semantic vectorization of a text) of the title and abstract of each patent using OpenAI's Ada-002 model. This captures the meaning of the entire body of the text into a $1,536$-dimensional numerical vector. LLM embeddings are very powerful and complete semantic representations of texts in their entirety (Radford et al., 2018; Brown et al., 2020). In the context of patent applications, they capture information related to a variety of patent outcomes, from economic value to citations, encapsulating the essence and quality of the invention (Yang, 2025). Then, we use this embedding vector as a primary input along with other fundamental variables (Cooperative Patent Classification and the number of claims) to secondary ML models, identifying whether an application deserves to be accepted. Specifically, the embedding is fed into a three-layer neural network with activation function *Mish* as proposed by Misra (2019):

$$Mish(x) = x \cdot \tanh(\ln(1 + e^x)). \tag{9}$$

The models predict the binary outcome *Acceptance*, equal to 1 if an application is accepted and 0 otherwise. The net acceptance rate of applications is over 70%, so we optimize the neural network for binary cross-entropy to account for this skewness. Moreover, the models are trained with a three-year rolling window sample, i.e., the predictive model for 2009 is trained on data from 2006 to 2008, ensuring all predictions are out-of-sample. Generally, the predictions are very accurate, with F1 Scores of over 80% across the board. We define *application quality* as an application's predicted chance of acceptance, i.e., how much its contents "deserve" acceptance.

Essentially, the models are trained to take as input the text of the application (i.e., description of the invention) and determine (a monotone transformation of) the objective *application quality*

[7]The quality of a patent application can only be based on its contents; this second part automatically holds.

from its contents.[8] With this measure as a control, we can identify unfair outcomes in the aggregate, again assuming that the model captures textual quality sufficiently.

Generic LLM approaches may be subject to look-ahead bias (e.g., He et al., 2025). Our approach differs from direct querying of ChatGPT, as employed by Jha et al. (2023), Kim, Muhn, and Nikolaev (2023), Hansen and Kazinnik (2023), and others. Instead, as in Yang (2025), we use only the textual embedding model and then train secondary models on the embedding.[9] This general semantic encoding is unlikely to contain granular information related to specific patenting outcomes. Prompting the model can lead to information retrieval from the dataset, but the embedding simply captures the meaning of the text. Moreover, look-ahead biases would likely skew results towards 0: for example, if the model knows all ex-post outcomes, the control then subsumes all variation. In all, it is unlikely that biases in the model could cause incorrect identification of discrimination.

Human perception tests. Next, we describe the human element in C-HML. While the ML quality measure is very accurate, it is imperfect. In general, Causal ML provides strong evidence of causality but does not absolutely rule out confounding factors as shown in Section 2.1. Causal ML techniques, such as our quality measure or CausalBERT in Hochberg et al. (2023), suggest unfair discrimination. C-HML offers, on top of that, more decisive proof of the discriminatory use of name information in our context. C-HML achieves this via a complementary exercise relying on human perceptions. Ultimately, the goal of identification in the case of patenting, job applications, loans, and these other examples is to determine whether the decision-maker is using name/demographic information, for to do so qualifies as discrimination.

Actual demographic information is not available to examiners at the time of review precisely because of discrimination concerns. Racial and gendered discrimination then happens, if at all, through the examiners' perception of names, i.e., s_N. For example, if an inventor is female but has a name that is challenging to identify as female, the examiner would likely not discriminate based on gender.[10] Thus, we have college students individually identify through inspection the "inferred

[8] One may argue that the structural inequity may be present in the training set, causing the model to favor certain applications even though it does not have any information beyond the application texts. This gives a rank-preserving transformation of the true quality as long as the structural inequity and patenting disparities favor the same group of applicants, something we verify in the data later. Additionally, the model replicating bias will lower the significance in the regression results, since more of the bias is controlled. Thus, our estimates present a lower bound.

[9] Their sample only includes publicly traded firms, whereas some of our most important results are derived from cross-sectional variation between individual applicants and firm assignees. Thus, we build new neural network models for each individual category.

[10] Beyond names, several papers have leveraged similar demographic inference noise to determine bias, specifically

gender" of inventors through their reported names.[11] Then, we exploit the difference between this inferred gender and the "real" gender (generated by the USPTO) to reject the alternative hypothesis that inventors do not use names, and the Causal ML results come from a correlation between gender and quality uncaptured by the quality measure. Specifically, if inventors are not inferring gender from names, perceived gender should make no difference, and only the true gender should be related to outcomes. Thus, we can identify whether the examiner is inappropriately using name-based information.

The innovation of C-HML from this literature is to apply perception measures in conjunction with a Causal ML control (*application quality*), as shown in Proposition 3. For example, in the case of gender, we separate the sample into true male and female inventors, then regress outcomes within each sample on perceived gender using the Causal ML control.

This substantially relaxes the identifying assumption relative to perceptions or Causal ML alone. The only case where a spurious correlation can occur is if the "genderability" of a name is significantly correlated with a dimension of quality not captured by the ML measure. The only assumption needed is then the negation of this scenario, which is in most cases true intuitively.

The ML + human learning/perception approach presents significant opportunities in various fields, particularly in identifying biases arising from the judgment of human decision-making, such as in job or loan applications. In contrast to other variables, the impression of a name should have no correlation to the actual content of the application, and its correlation with outcomes demonstrates that the decision-maker is potentially illegally using the name.

Relating to existing methodologies. Finally, we relate C-HML to existing methodologies, including some Causal ML approaches. The ideal way to determine the presence of discrimination is performing a field experiment as in Bertrand and Mullainathan (2004). Such an approach is often effective in several fields, for example, studying discrimination in crowdfunding (Gafni et al., 2021; Bapna and Ganco, 2021; Diep-Nguyen, Price, and Yang, 2025). Because the information available to (unknowing) participants in field experiments can be wholly controlled, it can be cleanly determined

as relating to photos in advertisements (Doleac and Stein, 2013) and policing discrimination (Goncalves and Mello, 2021). These approaches fit within our model: the inferred race from the photo may be considered as s_N, while the actual race is q_c.

[11] Of course, the accuracy of such an approach in actually determining true demographics is somewhat lacking and could be inferior even to external services like genderize.io or NamePrism. However, recall that the purpose is not to extract an accurate estimate of the gender, but instead base off of human perception. The statistical power derives from differences between extremely inaccurate human perceptions and the ML-based benchmark values.

that, for example, hiring managers are changing their decisions based on the name of the applicant. This is not feasible in the case of patenting, as discussed in the introduction.

One alternative approach developed in the computer science literature uses textual analysis techniques to identify treatment effects, known as CausalBERT. Hochberg et al. (2023) pioneer its application and discover a relation between gender and patent citations. Essentially, one trains separate models to simulate outcomes exclusively for each gender, then cross-applies the models on the exact text to approximate counterfactual treatment effects. Our language-embedding-based quality estimation uses the same identifying assumptions they do: the text contains all relevant information and that the ML model captures all relevant textual information.

For example, suppose that there is no unfair discrimination and the observed disparities are a result of a confounding factor (such as social upbringing) driving a relation between gender and patent acceptances. Consider the case where the ML model captures zero idiosyncratic information from the text, and thus always "predicts" the mean. Then, under both the CausalBERT and ML-based-quality methodologies, the results falsely indicate discrimination: in CausalBERT because the sample means are different across genders, and in ML-based-quality because the measure does not control for quality variation. Meanwhile, when the ML is perfect, both methods prove causality.

The key difference in our approach from CausalBERT is our customization of the technique to allow for the analysis of high-dimensional factors, interactions, and continuous independent variables rather than only singular binary treatments. As shown in Proposition 4, such considerations can be crucial. Specifically, CausalBERT and similar methodologies also face several additional challenges resolved by our unconditional quality approach. In particular, training counterfactual models becomes prohibitively time-consuming and difficult to interpret when discrete variables with many possible values (such as race) and interaction effects are involved. For example, the interaction between race and affiliation would produce fifteen distinct groups with CausalBERT, each necessitating its own model to then be cross-applied to all other groups.[12] Whereas earlier techniques are usually restricted to determining a binary treatment effect, our quality control applies to a wide range and variety of disparities. Further, our methodology also allows for determining continuous effects, such as the name frequency results in Section 5.2.

In addition, the perception tests in C-HML are integral because no causal identifying assumption

[12] In cases of such interactions, some groups alone may not contain enough observations to train a distinct, sufficiently accurate ML model in the first place.

can ever hold perfectly. For example, while CausalBERT models may have high accuracy at some times, any imperfection could still admit omitted variable bias. Critically, the assumptions for human perceptions are very likely to hold, whereas sophisticated ML models are also generally very accurate. Combining both of these thus significantly raises the bar for identifying a spurious correlation.

2.3. Practical Accuracy of C-HML: Evidence from Simulation

While our theoretical analysis demonstrates that C-HML always perfectly identifies discrimination under ideal conditions, this approach may fail when some assumptions fail to hold in practice. For example, Proposition 3 assumes that econometricians can condition on the true demographic characteristic q_c. But researchers often only have access to proxies—for instance, inferring demographics using statistical methods as a benchmark to compare perceptions to. Empirically, it is impossible to determine the validity of the estimate ex-post. Thus, we design a simulation study to evaluate C-HML's performance under worst-case conditions.

Simulation design. Our simulation implements the theoretical model from Section 2.1 with the modification that the econometrician observes only a noisy proxy $q_c^{proxy} = q_c + \epsilon_{proxy}$ (where ϵ_{proxy} is normally distributed with variance σ^2_{proxy}) rather than the true value. This creates a realistic scenario where:

1. The examiner makes decisions based on perceived demographics: $y = \beta_1 s_1 + \beta_2 s_N$, where $s_N = q_c + \epsilon_N$ represents the examiner's perception from reading names.

2. The econometrician attempts to estimate β_2 with a ML model, a recreation of s_N, and a proxy q_c^{proxy}, which is an imperfect approximation of true q_c.

We also assume that the ML model is trained on historical data that already contains discrimination, causing it to "inherit" biased patterns: loosely, the training data follows $y_{hist} = \beta_1 q + \beta_2 s_N$, leading to predictions q_{ML} that are correlated with s_N.

For the first simulation, we run $N = 1,000$ independent simulations for the primary estimates, each covering $20,000$ observations, with true discrimination parameter $\beta_2 = 0.5$.

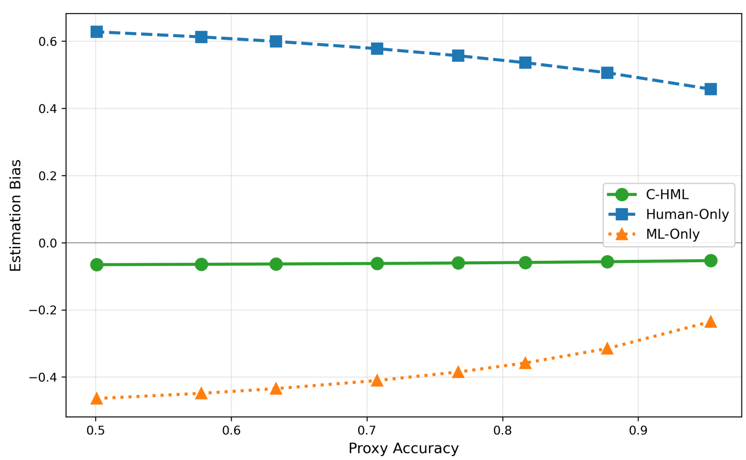
[Insert Table 1 Here]

Table 1 presents the main simulation results. Both the human-only and ML-only approaches exhibit substantial bias. Human-only estimates $\hat{\beta}_2 = 1.036$, an error of 0.536. This is a result of residual variation in true q_c due to the imperfect proxy variable creating omitted variable bias.[13] The ML-only approach similarly fails, with $\hat{\beta}_2 = 0.143$. Despite strong predictive performance for quality, the ML model's training on historically biased data causes q_{ML} to be contaminated with demographic information.[14]

In contrast, C-HML, despite being a composite of these two techniques, successfully recovers a relatively accurate estimate of the true parameter with $\hat{\beta}_2 = 0.442$. Though some assumptions have been weakened, C-HML requires a *much* looser identifying assumption, which is violated to a far lesser degree despite the extreme errors in both other techniques.[15] This strongly suggests that C-HML estimates can be trusted, especially considering that this simulation is essentially an upper bound on potential bias.

We also test to see the degree to which proxy accuracy is relevant, performing experiments while varying the accuracy of the q_c proxy.

Figure 2. Method Performance and Proxy Quality



C-HML's robustness to low proxy quality derives from its dual-signal approach: while proxy error degrades the conditioning strategy, the ML signal provides an alternative (albeit biased) source

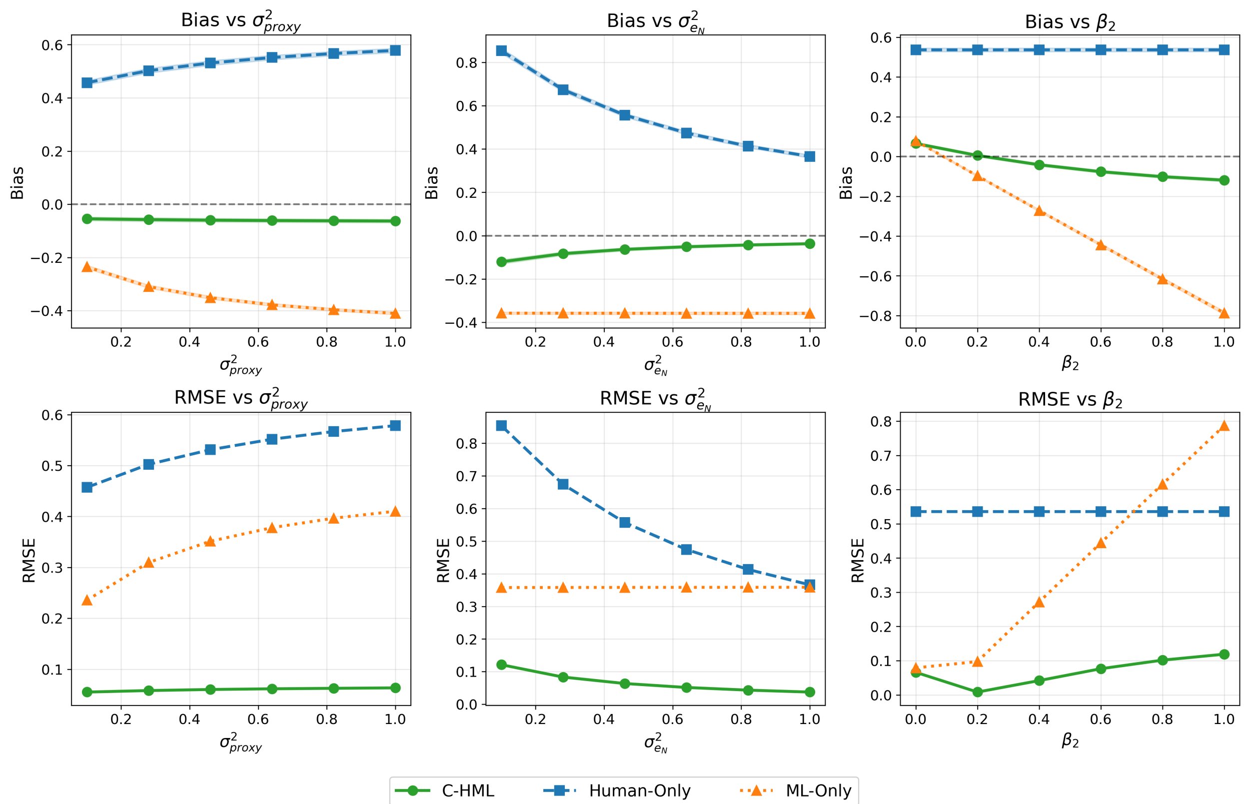
[13] This is essentially a worst-case estimate of the human-only bias, in order to demonstrate the persistent accuracy of C-HML. In reality, human-only tests are unlikely to result in such extreme misestimations, but will always be less accurate than C-HML.

[14] This information laundering effect manifests in the correlation $\rho(q_{ML}, s_N) = 0.651$, as reported in the table.

[15] It may seem like the error is simply averaging out, but this is not the case. It happens that the most likely errors, wherein the Causal ML learns from historical bias patterns and the demographic proxy is inaccurate, lead to opposite effects for the two benchmarks. C-HML is robust to the direction of bias in the fundamental techniques.

of quality control. We also confirm that C-HML superiority is also robust to large changes in other parameters, as seen in Figure 3.

Figure 3. Method Performance and Parameters



These simulation results have three key implications. First, they validate our theoretical framework while highlighting a crucial practical constraint. The human-only approach's failure is not due to model misspecification but rather measurement error—a ubiquitous challenge in discrimination research. Researchers rarely have access to protected characteristics directly and must rely on imperfect proxies, so integrating perception techniques with ML is key.

Second, they confirm that machine learning alone cannot solve discrimination detection. Our results show that ML models trained on biased historical data perpetuate biases, making them unsuitable for identifying contemporary discrimination. This "bias laundering" effect means that even highly accurate ML models can obscure rather than reveal unfair treatment.

Third, and most importantly, they establish C-HML as a practical solution. By leveraging both human perception and machine learning, C-HML alleviates the individual limitations of each approach. The robustness to proxy quality is particularly valuable given inherent empirical challenges in demographic inference.

Notably, any setting where (i) decision-makers infer demographics from imperfect signals like names, (ii) researchers lack access to true demographics, and (iii) ML models are trained on potentially biased historical data will exhibit similar challenges. This includes hiring decisions based on resumes, loan approvals based on applications, and academic admissions based on written materials. The simulation thus suggests that C-HML offers a general framework for detecting discrimination.

2.4. Explaining Discrimination in Review-Based Approvals

Suppose one observes discrimination ex-post, why do decision-makers make unfair decisions in the first place? Such discrimination is surprising because it not only is potentially inefficient but also poses a career risk to the examiners if discovered.

Existing theories of discrimination. The literature has proposed several theories to explain discrimination in decision-making. While foundational, we argue that existing frameworks provide an incomplete picture of the phenomena we document, necessitating a more specific mechanism.

1. Taste-Based Discrimination: The classical theory (Becker, 1957) posits that decision-makers have direct, animus-driven preferences against certain groups. This is less likely to be the primary driver for professional examiners who face significant career risk if such explicit bias is discovered.

2. Statistical Discrimination: Arrow (1973), Phelps (1972), etc., suggest that decision-makers rationally use group membership as a proxy for unobservable quality. However, this becomes unconvincing when an objective and legally proscribed signal of quality (e.g., patent texts) is available for evaluation.

3. Homophily: A related concept is homophily, the tendency for individuals to favor those who are similar to themselves (McPherson, Smith-Lovin, and Cook, 2001). While it could explain some demographic bias, it fails to account for the largest bias we find: the preferential treatment of public firms, with which examiners are unlikely to share a group identity.

4. Cognitive and Processing Biases: A broader class of theories points to cognitive shortcuts. These include stereotyping, where decision-makers overweight representative traits of a group (Bordalo et al., 2016); implicit bias, or unconscious associations that influence judgment (Bertrand, Chugh, and Mullainathan, 2005); and differential attention, where applications from certain groups receive different levels of scrutiny (Bartoš et al., 2016). While these theories correctly identify that examiners operate under cognitive constraints, they are underspecified as explanations for our findings. They do not offer a unified mechanism that explains bias across diverse dimensions (race, gender, affiliation, location) simultaneously. Nor do they make the sharp prediction that discrimination is strongest for borderline, low-quality cases, which we show in later sections.

Correlation neglect as a novel explanation. Building on these cognitive theories, we propose a simple yet systematic explanation in which *unfair* disparities could arise even unintentionally, invoking the concept of correlation neglect. Put simply, examiners may overweight factors such as affiliation, race, or gender because they mistakenly interpret them as independent signals of application strength, leading to "accidental" discrimination.

In our conceptual framework, recall that $q = q_i + q_c$ and that the examiner's job is to estimate the unobservable q. The examiner can read the content of the application, which provides a noisy signal $s_1 = q + \epsilon_1$. If the examiner is perfectly skilled and has full information, $\epsilon_1 = 0$. In practice, this is never the case, and thus, the examiner must seek other signals, namely, $s_N = q_c + \epsilon_N$ in the model. Previously, the examiner's problem of optimally approximating q is considered exogenously solved, and the solution $\beta_1 s_1 + \beta_2 s_N$ is taken as given. Now, we endogenize an examiner's determination of the weights β_1 and β_2. We further assume $(q_i, q_c, \epsilon_1, \epsilon_N)$ are independently and normally distributed with zero mean and variance $(\sigma_i^2, \sigma_c^2, \sigma_{\epsilon_1}^2, \sigma_{\epsilon_N}^2)$, which is not critical for the intuition of the result but admits closed form solutions. The examiner then considers $E[q|s_1, s_N]$.

Proposition 5. *Given s_1 and s_N, the conditional expectation of patent quality q is:*

$$\begin{aligned} E[q|s_1, s_N] &= \beta_1 s_1 + \beta_2 s_N \\ \textit{where} \qquad \beta_1 &= \frac{\sigma^2}{\sigma^2 + \sigma_{\epsilon_1}^2}, \quad \sigma^2 = \sigma_i^2 + \frac{\sigma_c^2 \sigma_{\epsilon_N}^2}{\sigma_c^2 + \sigma_{\epsilon_N}^2}, \\ \beta_2 &= \frac{\sigma_c^2}{\sigma_c^2 + \sigma_{\epsilon_N}^2} \cdot (1 - \beta_1). \end{aligned}$$

Importantly, if σ_{ϵ_1} is small relative to σ_i^2, then β_1 approaches 1 and β_2 approaches 0. Essentially, if the examiner is good and the innate quality of the idea is the most important, then the examiner should almost use only the text of the application in their estimation of q, i.e., $E[q|s_1, s_N] \approx E[q|s_1]$. If this were the case, β_2 would likely be too small to detect empirically.

One reason this may not hold in reality is that the examiner may easily overlook the correlation between the signals s_1 and s_N. This is a well-documented effect known as *correlation neglect*, where agents do not understand that distinct signals contain repetitive information. DeMarzo, Vayanos, and Zwiebel (2003), Eyster and Weizsacker (2011), and Enke and Zimmermann (2019) study the causes and effects of similar biases in various settings.[16] In particular, Chen, Cong, and Li (2025) rationalize apparent ex-post correlation neglect as a result of rational ex-ante choices of information source authentications. In the case of the examiner, one may be incognizant of the fact that s_1 already captures all three factors and instead mistakenly interpret the signals as $s_1' = q_i + \epsilon_1$ and $s_N = q_c + \epsilon_N$. Essentially, the examiner thinks that they are gleaning additional information about quality from the affiliation, name, and other characteristics of the inventor, even though it is already contained in the actual content of the invention.

Proposition 6. *(i) When the examiner misinterprets s_1 as s_1' under correlation neglect assumptions, the conditional expectation of patent quality q is given by:*

$$\begin{aligned} &E[q|s_1', s_N] = \beta_1' s_1' + \beta_2' s_N \\ \textit{where} \qquad &\beta_1' = \frac{\sigma_i^2}{\sigma_i^2 + \sigma_{\epsilon_1}^2}, \quad \beta_2' = \frac{\sigma_c^2}{\sigma_c^2 + \sigma_{\epsilon_N}^2}. \end{aligned}$$

(ii) The following always holds: $\beta_1' < \beta_1, \quad \beta_2' > \beta_2$.

The examiner's correlation neglect results in an overestimation of the importance of s_N, i.e., gendered and racial discrimination. The optimal strategy for a skilled examiner is essentially to use only the textual signal s_1. Thus, a double-blind reviewing process could reduce the risk of bias with minimal impact on approved application quality.

Our correlation neglect mechanism is particularly relevant and provides unique insights to professional decision-making under uncertainty because it

1. **Accounts for context-dependence**: The severity of correlation neglect varies with the quality of the application, and borderline applications create more uncertainty.

2. **Captures realistic information structure of professional review**: Human decision-making in reviewing applications often involves evaluating a direct signal (e.g., the application

[16] Interestingly, Levy and Razin (2015) show potential positive effects of correlation neglect in voting behavior. Because our examiners are otherwise rational agents, this bias is inefficient here.

text) alongside correlated signals (e.g., inventor characteristics). Correlation neglect is well known to hold in such settings yet is understudied in the discrimination literature.

3. **Reflects well-known cognitive constraints**: Examiners process applications under time pressure. The cognitive demand of adjusting for correlation between signals is high, making correlation neglect a natural consequence of bounded rationality.

The correlation neglect framework thus provides a parsimonious explanation for discrimination in patent examination that is consistent with the institutional context and our empirically observed patterns of bias. It suggests double-blind review processes, entirely preventing the observation of correlated demographic signals, may be an effective solution to prevent discrimination.

That said, does correlation neglect lead to distinct, testable predictions? Indeed, the framework yields sharp, testable predictions that we can take to the data. For example, we consider the potential effects of heterogeneity across discriminatory dimensions. Discrimination β_2' is increasing in σ_c^2 and decreasing in $\sigma_{e_N}^2$, so it is informative to consider differences across factors. Particularly, some characteristics may exhibit high σ_c^2, which we deem as **high explanatory power**, or how much q_c affects the true quality. Simultaneously, we term low $\sigma_{e_N}^2$ as **low perception noise**. If it is easier to identify a characteristic, discrimination will also be more pronounced. Thus, the largest discriminatory effects should be seen for characteristics with both high explanatory power and low perception noise. We later test this prediction in the data and confirm it.

At a finer level than characteristic heterogeneity, also consider that a high-quality patent application, characterized by a novel and well-presented idea, is easy for an examiner to evaluate. The signal from the text, s_1, is strong and clear, and the examiner can quickly accept the patent. Conversely, a borderline application is often ambiguous, poorly written, or only incrementally innovative. In situations of high uncertainty, decision-makers may unconsciously fall back on heuristics, stereotypes, or other peripheral cues to guide their judgment. In our model, this means relying more heavily on the demographic signal s_N.

We formalize this by relaxing the assumption that the noise in the examiner's evaluation, ϵ_1, is constant. It is plausible that the variance of this noise, $\sigma_{\epsilon_1}^2$, depends on the intrinsic quality of the application, q_i. Specifically, we assume $\sigma_{\epsilon_1}^2$ is a decreasing function of q_i, as higher-quality ideas are less noisy to evaluate.[17] Let $\sigma_{\epsilon_1}^2(q_i)$ denote this quality-dependent variance, with $\frac{d\sigma_{\epsilon_1}^2}{dq_i} < 0$.

[17]One may also argue for modeling σ_{ϵ_1} as an inverted-U-shaped function of q_i, which may be the case in some

Now, revisit the decision-making process of the correlation-neglecting examiner. The examiner mistakenly treats s_1 and s_N as independent signals for q_i and q_c respectively, leading to the decision rule $y = \beta_1' s_1 + \beta_2' s_N$. The weights then become functions of intrinsic idea quality q_i through ϵ_1.

Proposition 7. *Given correlation neglect, if the variance of the examiner's signal noise, $\sigma^2_{\epsilon_1}(q_i)$, is a decreasing function of intrinsic quality q_i, then the following always hold:*

*(i) The weight on the textual signal, $\beta_1' = \dfrac{\sigma_i^2}{\sigma_i^2 + \sigma^2_{\epsilon_1}(q_i)}$, is **increasing** in q_i.*

*(ii) The weight on the demographic signal, $\beta_2' = \dfrac{\sigma_c^2}{\sigma_c^2 + \sigma^2_{\epsilon_N}}$, is **independent** of q_i.*

*(iii) The relative weight of the demographic signal to the textual signal, $\dfrac{\beta_2'}{\beta_1'}$, is **decreasing** in q_i.*

This provides a second, precise theoretical prediction: discrimination, driven by the examiner's reliance on the demographic signal s_N, is not uniform. For high-quality patents (high q_i), the examiner receives a clear signal from the text (low $\sigma^2_{\epsilon_1}$) and places a large weight β_1' on it. The decision is thus primarily driven by the patent's content. For low-quality patents (low q_i), the examiner is more uncertain (high $\sigma^2_{\epsilon_1}$) and places less weight on the ambiguous textual signal. While the absolute weight on the demographic signal β_2' remains the same, its influence *relative* to the textual signal becomes much larger. Consequently, the examiner's ultimate decision for a low-quality application is more heavily influenced by the inventor's perceived characteristics, resulting in greater measured discrimination for this group. We also later test this prediction in Section 6.3.

3. Institutional Background and Data

3.1. Patenting

When an inventor or firm discovers a new invention, they must file an application with the US Patent and Trademark Office (USPTO).[18] An application consists of a detailed description of the invention, an abstract, specific technical claims, and sometimes figures. In addition, the examiner

settings where extremely low and extremely high quality are obvious. However, patent decisions are discrete and heavily biased towards acceptances, so that only the worst patents are rejected. This means the borderline patents are essentially all at the lower end, while better patents are progressively easier to distinguish.

[18] Patents fall under three categories: utility, design, and plant. This paper focuses on utility patents, which are the most important and by far most common subset.

also receives a limited set of more structured information: the number of claims, the affiliation of the inventor, and the applicant's name.

Once the application is received, it is randomly assigned to an examiner within the art unit (subject area) of the patent. Then, the examiner should determine whether the application meets the three legal standards of acceptance: novelty, utility, and non-obviousness.[19] Importantly, all of these standards pertain only to the content of the application. Thus, the examiner cannot legally consider external factors such as the identity or affiliation of the applicant in making a decision.

3.2. Data and Variables

Our primary data source is the USPTO's PatentsView service, which provides bulk textual data for all utility patents granted after 1976 and all patent applications published after 2001. Firm market capitalization is obtained from CRSP and firm variables from Compustat.

Our ML-based quality measure is as described above in Section 2.2. Our measures of individual inventor and examiner race are constructed using the Bayesian Improved First Name Surname Geocoding (BIFSG) algorithm via the Surgeo Python package. The use of this proxy is necessary because these are not self-reported in the publicly available data. We note that Chernenko and Scharfstein (2023) and Greenwald et al. (2024) recently found that the use of such algorithms may lead to substantially biased results.

However, this method is appropriate in our use case. Even if the algorithm is inaccurate, our results will be correct when the inaccurate human perceptions match the examiner's perception of race, and BIFSG is simply the more accurate benchmark for comparison. Additionally, simulation evidence in 2.3 shows that C-HML is robust to (independent) error in the demographic proxy (in this case, BIFSG).

Second, the classical measurement error in BIFSG would bias our estimates toward zero, making our findings a conservative lower bound. In the C-HML application, for BIFSG errors to spuriously generate our results, they would need to be systematically correlated with both human perception errors and discriminatory examiner behavior, which is implausible.

Finally, much of the potential error in empirical tests using these algorithms stems from the

[19] Patent Act of 1952, 35 U.S.C. §§ 101–103, and MPEP 2106 and 2107. A technically fourth statutory (subject matter eligibility) requirement is: the patent must be a process, machine, article of manufacture, or composition of matter.

geographical element. For our primary tests, we set the location for all inventors to be Alexandria, Virginia, the location of the USPTO headquarters and examination center. For robustness, we also test without geocoding and obtain largely the same result. Geocoding for inventor location is unreasonable since the examiner's subconscious evaluation of race would not take into account the racial distribution of, say, zip codes in Seattle.

The measure of gender for inventors is created by the USPTO and taken directly from PatentsView, which uses machine learning algorithms to identify the most likely inventor's gender based on first and last name. This proxy functions equivalently as with the race measure. There are also no self-reported or "true" gender demographics available in the bulk patent data. For examiners, we use the third-party API provided by genderize.io for gender inference based on both first and last name.

For most tests, race is measured using a dummy variable measuring the "plurality" race of a patent. For example, if a patent has five listed inventors, three Hispanic, one White, and one Black, then the plurality race is Hispanic. This same logic applies to individual inventors, as SURGEO outputs continuous probabilities for the race and not a discrete binary value. That is, each individual is first assigned the race to which they most likely belong. After this, if the application has multiple inventors, the plurality is taken again.

Human perceptions, i.e., "Maleness," "Whiteness," and "Blackness" of names are determined through reading names by at least two college students and identifying their majority impressions. For gender, only first names are used, while first-name-last-name pairs are identified for race.

4. Racial and Gendered Disparities

The primary motivation for this study lies in the significant disparities found across demographic groups throughout the innovation chain. For example, a wealth of research studies later-stage outcomes by race, such as VC funding (Cook, Marx, and Yimfor, 2022) and citations (Hochberg et al., 2023). Nevertheless, the area of patent application acceptance remains relatively understudied, and is especially important with regard to the survival of young firms and startups and thus, indirectly, social mobility. We hypothesize in this section that patent applications filed by White or male inventors are more likely to be accepted than those filed by minority inventors.

4.1. Suggestive Evidence of the Racial Gap

Table 2 reports the summary statistics of application acceptance and predicted acceptance rate as sorted by (plurality) race. API patents are accepted the most, with an acceptance rate of over 70% across all years, while White inventors receive the second highest rates, at 67.7%. Black inventors receive the worst outcomes of all groups except Indigenous Americans, with only a 60.6% acceptance rate. Comparing instead with predicted acceptance rates (*Application Quality*) in Panel B, the relative rankings of groups by mean rate remain the same. However, the mean predicted acceptance rate for Black inventors is 66.8%, 2% higher than the actual outcomes, whereas both White and API are much closer in terms of difference. These statistics suggest that race may be influencing patent acceptances since the model's predicted rates do not entirely capture the disparity between groups.

[Insert Table 2 Here]

4.2. Identifying Unfair Outcomes

To more rigorously determine this potential effect, we first turn to a standard regression setting, where we are able to control for the objective measure of application quality provided by the predicted acceptance rate from the machine learning model. Specifically, we estimate the following regression for application-time (i, t):[20]

$$Accepted_{i,t} = \sum_{R \in \{Races\}} \beta_n \cdot R_{i,t} + \theta \cdot Quality_{i,t} + \gamma \cdot Controls_{i,t} + \delta_t + \epsilon_{i,t}. \tag{10}$$

Table 3 reports these regression results. The relation between race and patent acceptance remains significant at the 1% level under all specifications for the groups White, Black, and API. Black inventors are systematically discriminated against, while White and API inventors receive an advantage. Importantly, these coefficients do drop by half or more when introducing the application quality control. Thus, the disparity in overall success is driven by two orthogonal, roughly equally important factors: first, Black and Hispanic inventors, on average, produce lower "quality" inventions

[20] Due to collinearity, the set $\{Races\}$ cannot contain all six race variables. In general, Hispanic inventors as a group tend to roughly rank in the middle of the races, in terms of group size and also potential discrimination (i.e., the difference between predicted rates and real acceptances). Thus, the Hispanic group is treated as the benchmark and reported coefficients from the regression are to be viewed as relative to the Hispanic group. That is, when we say that being White grants an unfair advantage, this is relative to Hispanic inventors.

than White and API inventors (where quality is determined as per the USPTO's metrics for patent acceptance), and that, *ceteris paribus*, they still experience lower acceptance rates.

[Insert Table 3 Here]

Overall, controlling for application quality removes almost all other confounding factors that could be driving a spurious relation. For example, if inventors were disproportionately sorted into industries with different aggregate acceptance rates by race, then summary statistics would show a disparity in successful innovation outcomes. This effect is controlled for because the application quality measure accounts for sectors and matches the aggregate acceptance rates across industries (see Yang, 2025, Figure 1). Of course, there may still be confounding factors driving this relation. One primary source of variation is found in examiners, who may be individually harsher, more lenient, or busier (Farre-Mensa, Hegde, and Ljungqvist, 2020; Shu, Tian, and Zhan, 2022). If White inventors were to receive more lenient examiners, this could also drive the results. However, this scenario would still qualify as unfair discrimination, given the USPTO's claims of true random assignment.

4.3. The Gendered Patenting Gap

In addition to race, another important dimension of discrimination is gendered bias. Like race, gender is not officially reported in the patent application process, but the name of the inventor can allow the examiner to (either subconsciously or consciously) form their own determination of the inventor's likely gender. We define gender as the majority gender among inventors of a certain application.[21]

Similar to race, disparities exist both in application quality and in aggregate acceptance rates. Panels A and B of Table 4 report these summary statistics. Specifically, the mean acceptance rate for female majority patents is only 63.8%, whereas the male acceptance rate is 69.0%. The gap in quality (predicted rate from text) is smaller, from 65.2% to 69.6%.[22]

[Include Table 4 Here]

[21] Patents with equal male and female inventorship are classified as female. The results are robust when they are classified as male too.

[22] Quality accounts for many dimensions of variation in acceptance rate, including field, and does not necessarily reflect inventor skill. For example, patent applications filed in a more competitive field are, by definition, lower "quality."

Next, we perform the same quality-controlled test as in Section 4.2. Specifically, we estimate:

$$Accepted_{i,t} = \beta \cdot Male_{i,t} + \gamma \cdot Controls_{i,t} + \delta_t + \epsilon_{i,t}, \tag{11}$$

where Male is a binary variable equal to 1 if the inventors are majority male and 0 otherwise.

Panel C of Table 4 reports the results. Across all specifications, applications filed by male inventors are much more likely to be accepted. Controlling for quality and year-fixed effects, the effect is still nearly 2 percentage points. Economically, this bias is stronger than the gap between Hispanic and Black or Hispanic and White but slightly weaker than that between Black and White.

4.4. Confirming Causality with C-HML

The above results demonstrate the persistence of substantial disparities between male and female and races. Nevertheless, as discussed in Section 2, the ML techniques are potentially subject to bias. In this section, we seek to show that examiners are directly inferring (and discriminating based on) gender and race from *inventor names*. As demonstrated in the model, we exploit the difference between perceived gender and race (from names) and their true values.

For gender, name is reviewed and labeled as male, female, or ambiguous. The labeling is done after brief inspection, without extended thought or access to outside materials such as the internet. This recreates examiners' evaluations, who are very unlikely to Google the racial distribution of an inventor's name. Then, each name is labeled separately by a second reviewer without knowledge of the first review outcome for robustness. Results presented are from the first review; second and subsequent reviews do not generate qualitative or statistical differences in results.

Discrimination based on perceived gender. First, we seek to rigorously show causality for the gendered discrimination finding. As our "true" value, we use the USPTO provided data, which uses ML to accurately infer inventor gender.[23] Meanwhile, we also manually label the $15,000$ most common inventor names as obviously male, obviously female, or not obviously gendered. By labeling inventors using the impressions their names make on average humans, we extract the examiner's beliefs about gender. In general, human perceived gender often agrees with true gender, but the Pearson correlation coefficient is only around 71%. While the name-based perception is far from

[23]This measure generated by the USPTO for ex-post studies is not available to the examiner at the time of review.

random, it also often disagrees with the true value, providing the statistical power we need.[24]

Essentially, we seek to isolate and remove the underlying correlation of gender and application quality, instead determining if the examiner is biased. In this case, we first split the sample into two: male inventors, and female inventors. Then, within each sample, we perform a similar gendered discrimination regression to Section 4.3, using *perceived* instead of actual gender, and taking only names that are identified as strongly gendered. If the examiner is using only the legal standards for acceptance, then perceived gender, conditional on actual gender, should have no effect on outcomes.

Panel A of Table 5 reports the regression results. In both samples, the perceived "Maleness" of a name has a significant effect on acceptances, even after controlling for quality. Among female inventors, those with male-sounding names experience an increase of 2 percentage points in acceptance rate, and a similar effect is seen among male inventors. The significance of these coefficients is strong evidence that examiners are not simply examining other aspects of the application correlated with gender but instead are making judgments based on an impression of gender derived from the inventor's name.

[Insert Table 5 Here]

Note that many individuals are ambiguously gendered to the average reader, and are dropped from the above test to isolate the effect of strongly gendered names. These failures to label gender at all result both from gender-neutral names such as "Jordan" but also from cultural differences: American readers are unable to identify names such as Sung-Min as either male or female. Thus, we also perform a test by splitting the sample into the following two subsamples. First, inventors with names obviously gendered to the average American English-speaking reader. Second, inventors without obviously gendered names.

Then, we again perform a gendered discrimination regression. Panel B of Table 5 reports the results. Although males experience an advantage in both samples, the effect is universally twice as large in inventors with obviously gendered names.[25] This difference is further evidence that examiners tend to reject applications with female-sounding names *ceteris paribus*.

[24]In untabulated results, the manually-labeled gender measure produces slightly stronger results than the USPTO gender as a regressor, confirming that the labeling is relevant.

[25]In order to not over-estimate the difference between samples, we lean towards ambiguity when labeling names. Thus, some names that may seem female or male to certain examiners are labeled ambiguous, which could be the cause for the persistence of the effect in ambiguously gendered names. Regardless, our result effectively provides a lower bound on the actual impact.

Discrimination based on perceived race. Next, we apply similar techniques for the case of race. Racial discrimination requires a more complicated treatment because there are many potential races, and the human perception error is unclear and may often be correlated with true race. For example, it is much more likely to falsely label a White name as Black than an API name as White. Thus, we perform two tests. First, to identify if White inventors experience an advantage (relative to others). Second, to determine what disadvantage (if any) Black inventors face. For White inventors, we similarly split the sample into White and non-White inventors, then regress acceptances on perceived Whiteness.

Panel A of Table 6 shows the results. Even among White inventors, name Whiteness has a significant effect on success. The effect is even more pronounced among non-White inventors, with a White-sounding name leading to a 2.2 percentage point increase in acceptance rate *controlling for quality*. As with gender, the perceived Whiteness of a name should have no underlying link with economic variables. Thus, the examiner must also infer race from names.

[Insert Table 6 Here]

Similarly, we perform the tests for perceived Blackness. For race, we label the $20,000$ first-name-surname pairs comprised of the $4,000$ most common ambiguous, White, Black, API, and Hispanic names.[26] Panel B of Table 6 reports these regressions. Among non-Black inventors, there is a small effect that vanishes when controlling for quality. However, among Black inventors, the perceived Blackness of a name has a very significant negative association with outcomes—Black inventors with more Black-sounding names are more than 3% less likely to receive acceptances. This effect is measured after controlling for race and application quality and thus shows that examiners discriminate substantially against inventors with "Black names."

In all, the evidence suggests that examiners (whether consciously or unconsciously) make biased decisions based on name-related information. Thus, a double-blind procedure could reduce unfair discrimination.

5. Novel Discrimination Factors and Their Relative Importance

Beyond race and gender, could other factors also lead to unfair outcomes? Examiners could discriminate (intentionally or unintentionally) based on any relevant individual characteristics. Thus,

[26] Race, unlike gender, is not easily identifiable from first names, so surnames are used.

bias across non-traditional dimensions could also be economically significant. In fact, the degree of bias could increase with the relative importance of the factor in the first place—meaning that certain characteristics could generate even more extreme disparities.

5.1. Applicant Types and Affiliation Bias

One of the most important non-demographic traits of an inventor is *firm affiliation*; essentially, whether or not the patent is associated with a company and what that company is. This could be another source of bias. Under correlation neglect rather than taste-based discrimination, a rational examiner would mistakenly try to use all information available, leading to many potential discriminatory factors beyond demographics like firm affiliation.

For example, an examiner may subconsciously give more credence to claims made by inventors from Apple or Microsoft than John Smith from Arkansas, even if the two file exactly identical applications.[27] This can potentially be incredibly harmful in terms of social mobility since poorer entrepreneurs and inventors will obviously file applications as individuals.

To test for the presence of such effects, we define three types of patent applicants: public firms, private firms, and individual inventors. Public firms are all assignees that can be matched to a publicly traded PERMNO in CRSP. Private firms are assignees that are not matched to public firms or government agencies. Individuals are simply inventors filing for a patent under their own name, that will belong entirely to themselves if accepted.

Panels A and B of Table 7 report the aggregate acceptance rates and mean application quality across these applicant types. Notably, the differences here are much larger than in gender and race, while the sizes are as expected: public firms have the lion's share of their applications accepted, at 75.1%, with far less accepted for private firms and individuals (66.7% and 60.1%). Again, differences in quality/expected acceptance also exist, but are notably several percentage points lower than the gaps in actualized outcomes. In the case of applicant types, it is expected that there will be large gaps in quality irrespective of unfair treatment; after all, if NVIDIA were producing worse patent applications than individuals, that would be cause for significant concern.

The size of the gap indicates that affiliation may precisely be the type of characteristic with *high*

[27]Whether such a bias is efficient is not immediately clear, but it seems rather inappropriate in the context of patent applications. Unlike, for example, credit score in loan or mortgage applications, the identity of the inventor should be irrelevant in such a technical context.

explanatory power discussed in Section 2.4. Importantly, firm affiliation also has essentially zero *perception noise* since it is written explicitly on the application if the inventor is from a firm or an individual. Thus, we ought to see very large discriminatory effects for affiliation.

[Insert Table 7 Here]

We estimate the following quality-controlled regression:

$$Accepted_{i,t} = \beta_1 \cdot Public_{i,t} + \beta_2 \cdot Private_{i,t} + \gamma \cdot Controls_{i,t} + \delta_t + \epsilon_{i,t}. \tag{12}$$

The results are shown in Panel C of Table 7. In all specifications, private firms experience an advantage relative to individuals, and public firms an even larger advantage. In fact, the gap between public and private firms is twice as large as that between private firms and individuals. Overall, public firms are associated with an improvement of 6.6 percentage points in acceptance rate *with the same quality of application*. Indeed, it is much larger than for either gender or race, confirming the theoretical predictions under correlation neglect.[28] The persistence of such a large effect controlling for quality is concerning because the ability of individuals to succeed in innovation is critical to entrepreneurship.

Note. This "bias" may not be economically inefficient. For example, a patent for a new type of screen technology could, on average, generate more total welfare in the possession of Apple or Microsoft than an individual inventor. If this is the case, such outcomes could be optimal for welfare.

5.2. Other Dimensions of Bias

Broadening our examination of discrimination beyond demographic groups and homophily effects also opens the door to several other dimensions. Beyond affiliation with a powerful institution, as discussed above, many other "soft" variables are also available to the examiner. For example, we have already shown that examiners discriminate by race and gender via inventor names. However, there could be a third dimension: bias against uncommon names. For example, an examiner could prefer a patent filed by "John" rather than "Percival," even though both are most common among White men. Because name frequency over the entire population is not representative of inventors in particular,

[28] We note that human perceptions and affiliation homophily are irrelevant here because affiliations are directly observed in the applications and examiners are government officials.

we define Name Frequency as the natural logarithm of the total occurrences of a particular first name over all inventors plus one. Then, we estimate the following regression:

$$Accepted_{i,t} = \beta \cdot Name\ Frequency_{i,t} + \gamma \cdot Controls_{i,t} + \delta_t + \epsilon_{i,t}. \tag{13}$$

Panel A of Table 8 reports the results. The effect is quite pronounced with the most common names experiencing an advantage of over 6% relative to uncommon names. However, this could also be a result of name frequency simply proxying for male or White inventors, both of which are over-represented. Thus, we perform the same regressions controlling for both race and gender.

Panel B of Table 8 shows the controlled regression. Interestingly, the effect remains as strong or grows even more significant when controlling for race and gender.

[Insert Table 8 Here]

Another potential source of bias lies in an inventor's location, which is also reported on the patent application. An examiner may (correctly) understand that the average innovation produced in San Francisco is more impactful than that produced in rural areas yet still overweight the importance of the location in their judgement. To test for such an effect, we classify American inventors as either from large cities or not, where a large city is defined as a top 50 most populous city in the United States. Then, we estimate the following regression:

$$Accepted_{i,t} = \beta \cdot Large\ City_{i,t} + \gamma \cdot Controls_{i,t} + \delta_t + \epsilon_{i,t}. \tag{14}$$

Panel C of Table 8 shows that examiners do favor inventors from big cities, but the effect is relatively small in magnitude when controlling for application quality. Examiners are less biased against inventors from rural areas or smaller cities than against those without firm affiliations. This is consistent with the relatively low importance of inventor location. For example, firm affiliation is a stronger predictor of quality in the first place. Thus, examiners exhibit more significant beneficial treatment to such inventors than those from large cities. Nevertheless, bias along this dimension persists and would potentially disappear via double-blind reviewing.

5.3. Examiner Characteristics and Homophily

Homophily and its variants are often believed to be a driver for irrational or discriminatory behavior (Alsan, Garrick, and Graziani, 2019; Cullen and Perez-Truglia, 2023; Diep-Nguyen, Price, and Yang, 2025, e.g.,). More broadly, a hitherto unexplored dimension in our study is the potential role of examiner characteristics. Even though the examiners at USPTO are randomly assigned, it is still possible that a particular type of examiner dominates the population and discriminate against particular groups of applicants. For example, a net preference for White inventors may simply be the manifestation of a majority of examiners being White and all examiners prefer their own race. After all, examiner unreliability is well-documented in Farre-Mensa, Hegde, and Ljungqvist (2020) and related literature that use assignment as an instrument and by Yang (2025), who find that examiners make decisions not just based on content but also presentation.

Many of our findings so far suggest that *correlation neglect* is driving at least part of the results. To confirm this, we test whether bias still persists after accounting for homophily. Specifically, we perform the standard quality-controlled regressions conditional on examiner gender and race.[29]

[Insert Table 9 Here]

Panel A of Table 9 shows the results for male and female examiners. Across both samples, the preference for male inventors persists, at 1.6 percentage points (quality-controlled) for male examiners and 1.8 percentage points (quality-controlled) for female examiners.[30] Clearly, the gendered bias result is not a result of homophily.

Panel B of Table 9 shows the same tests for race. The interpretation here is slightly more nuanced, as some clear homophily effects do emerge. For example, the White preference for Whites is the largest, while API examiners have a larger preference for API inventors. Nevertheless, Black examiners still exhibit against Black inventors, which means it is again impossible for homophily to fully explain the core bias results. This suggests the need for an alternative, supplementary explanation, such as correlation neglect.

The presence of homophily at all is nonetheless interesting, and raises the question of what other

[29]Unlike for patent inventors, examiner gender is not available in the USPTO dataset. Thus, we use the API from genderize.io, a third-party service that uses similar machine learning approaches to identify gender. Specifically, we give genderize.io examiners' full names, and run the regressions only for examiners labeled definitively as either male or female (and not "unknown").

[30]The difference between these two coefficients is not significant.

dimensions of homophily might be present. In Section 5.2, we find that name frequency, beyond race, affects outcomes. Could examiners also prefer inventors with the same first or last names, as seen in Mao, Lu, and Wang (2025)? We regress success on whether the examiner has the same first name or last name as the inventor, controlling for a race match between the examiner and inventor.

[Insert Table 10 Here]

Table 10 shows the regression results. Interestingly, though a first name match does not affect the outcome, a last name match *controlling for race* is predictive of success, with an increase in chances of 2.4 percentage points.[31] Ultimately, this is further evidence both that examiners are unreliable and that some biases, but not all of them, can driven by a combination of homophily and correlation neglect.

6. Heterogeneity and Intersectionality in Patenting Disparities

6.1. Cross-Sectional Disparities

In Section 4, we demonstrate the persistence of causally unfair outcomes. However, such examiner biases may not be homogeneous. Many soft, unregulated factors including, but not limited to, race and gender, may influence examiner decisions. For example, firm affiliation plays an even larger role in driving biased outcomes.

In terms of correlation neglect, examiners may believe that if the patent is filed by a well-known firm such as Apple or Microsoft, it likely contains novel, useful technology. An examiner may learn this firm affiliation before reading the text of the patent or the name of the inventor, thus "overriding" demographic biases. It is also possible that examiners prioritize reading the inventor's name first.

6.1.1 Racial Discrimination

If affiliation overrides demographic biases, the disparities between races should be most pronounced for individual inventors, less so for private firms, and minimal or non-existent for applications filed by public firms. That is, affiliation with a big name firm may be powerful enough to push the application over the line alone, before the examiner's racial biases even have a chance to come into

[31] The statistical power in last names independent of race comes when the name is not indicative of race.

play. To test this hypothesis, we again split the sample into the three cross-sections: public firms, private firms, and individuals.

Public firms. In Table 11, we report, for the subsample of publicly traded firms, the same summary statistics and regressions as for the full sample. Interestingly, many of the trends seen in the full sample do not apply for public firms. Across races, there is much less variation, both in acceptance rate and application quality.[32] In the regression setting, White inventors still experience some advantages, but the effects are much smaller and only appear in a few specifications.[33] Clearly, the discrimination effect is much less prominent for applications filed by public firms.

[Insert Table 11 Here]

Private firms. Table 12 shows the same tables for the subset of applications filed by private firms. These represent a middle ground between big public firms and individual applicants—an LLC or Corp. will reassure the examiner, but not to the level that NVIDIA or AMD would. Here, although the acceptance rates look somewhat similar across races, White inventors experience a significant unfair advantage when controlling for quality. Overall, the discrimination effect is much stronger than found in public firms, but still weaker than in the entire sample, especially when comparing the relative pairs of White-Black and API-Black.

[Insert Table 12 Here]

Individual inventors. Individuals are the most likely to experience an unfair process in patenting: for example, many inventors may lack a patent lawyer. They are also, unfortunately, the most vulnerable. While an inventor at Apple is unlikely to lose their job simply because of a failed patent application, an aspiring entrepreneur may find themselves in financial ruin after unsuccessfully trying to protect their intellectual property. If individuals are discriminated against on the basis of race, this would further exacerbate the problem and have significant implications for socioeconomic mobility.

Table 13 reports the results for individuals. For individuals, even the summary statistics show that acceptance rates are not at all close across races. While Black inventors experience an average

[32]The closing of the gap in quality could be a result of several factors. For example, greater access to patent lawyers when working under a major company, or talent identification in the hiring process.

[33]The observation count of inventors identified as Indigenous American by SURGEO is too low to draw conclusive results from, but the evidence generally suggests a moderate level of discrimination.

rate of 54.3%, 60% of White patents are granted, a difference of 5.7 percentage points. While there is a difference in quality between these two groups, it is significantly lower, at only 2.2 percentage points. When controlling for application quality, White and API inventors experience a major advantage relative to the Hispanic group, and the Black inventors have a major disadvantage relative to Hispanic inventors and, of course, White/API. The magnitude of the effect is even greater than in the full sample. The true extent of the discrimination effect is, in fact, masked by the presence of public firms in the sample.

[Insert Table 13 Here]

In general, such cross-sectional variation is critical to acknowledge. In this case, failing to filter by affiliation bias would lead one to under-estimate the discriminatory effect. However, if, for example, for inventors affiliated with public firms, the discriminatory effect were reversed (i.e., preference for Black inventors), then a study of the entire sample would reveal no discrimination even though individuals face unfair outcomes. Thus, considering multiple dimensions simultaneously is critical.

6.1.2 Gendered Discrimination

The natural question that then arises is whether gendered discrimination follows the same patterns. Ex-ante, the result is unclear: the backing of a public firm could have an ameliorating effect on the bias, or the discrimination could persist across all applicant types. Thus, we conduct the same quality-controlled regression of application acceptance and gender for all three applicant types (Public Firms, Private Firms, and Individual Inventors).

Table 14 reports these regression results. Interestingly, the effect remains large and statistically significant in all three categories. While there is a slightly larger discrimination effect for individual female inventors, the difference is negligible. Essentially, female inventors experience a substantial disadvantage *irrespective* of their affiliation with a public firm, conceptually consistent with previous results showing that female employees at innovating firms experience unfair treatment throughout the innovation chain and not only in patenting (Waldfogel, 2023; Chien and Grennan, 2024).

[Include Table 14]

Note that only a simultaneous analysis of gender and race with identical techniques as presented in this paper can demonstrate and quantify differences between gendered and racial discrimination.

There are several possible explanations for this effect in particular. For example, gendered bias could be "stronger," not in terms of magnitude but rather persistence.

6.2. Effects of Intersectionality and Bias Interactions

One benefit of applying a uniform methodology to varying dimensions of discrimination is to compare their magnitudes and identify cross-sectional differences, as done above. However, we can also investigate the effects of intersectionality between groups. For example, would membership in more than one disadvantaged group ameliorate or compound the discrimination effect? The answer is ex-ante unclear, but it is rather unlikely that the interaction would be simply additive.

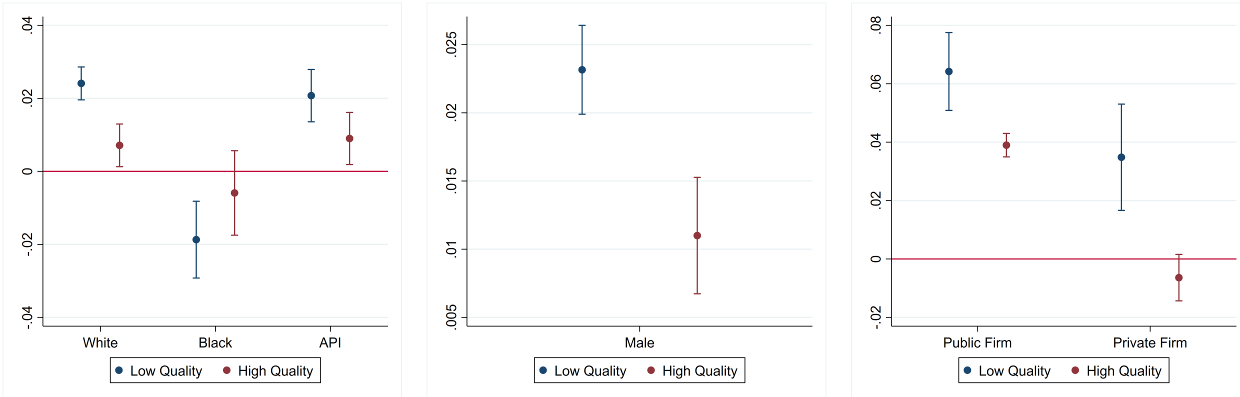
Specifically, we are particularly interested in the intersectionality between race and gender, which is an important field. To identify the subtle interactions between the two, we perform two interaction tests. First, we filter on race, then perform gender regressions. Second, we filter by gender and perform race regressions.

Table 15 shows that conditional on inventor gender, racial discrimination persists but with substantially varying effects. The advantage experienced by White and API inventors is greater for female than male inventors, while the disadvantage experienced by Black inventors decreases. In fact, API females enjoy a much more significant advantage (3.5%), which even exceeds that of White females (1.9%). Among males, White and API inventors have roughly equal advantages (1.4%) and Black inventors a disadvantage (-1.4%).

[Insert Table 15 Here]

Table 16 displays the gender disparity conditional on inventor race. Again, there are substantial cross-sectional differences in discrimination. Gender bias is strongest among applications filed by Hispanic (2.7%) and White (2.5%) inventors, followed by Black inventors (1.4%). Notably, gender disparity is almost entirely mitigated for the API population (0.7%). Recall that a large portion of discrimination is driven entirely by name-inferred gender. Since API names are often harder to accurately identify as male or female for the average American, the effect is less present.

[Insert Table 16 Here]



(A) Race Disparities **(B)** Gender Disparities **(C)** Applicant Type Disparities

Figure 4. Disparities in High- vs Low-Quality Applications.

6.3. Quality-Heterogeneous Discrimination

If examiners are discriminating, would the effects be homogeneous? Section 2.4 predicts that examiners experiencing correlation neglect will rely more heavily on the external characteristics for low-quality applications. To test for this effect, we examine the effects of the three dimensions of discrimination (race, gender, and applicant type) in relation to application quality.

Specifically, we separate the sample into below- and above-median quality applications and test for discrimination in both halves of the sample. We consider the most restrictive regression settings in which fixed effects and application quality controls are included, and standard errors are clustered by year. The coefficients and their 95% confidence intervals for the high-quality and low-quality samples are plotted in Figures 4A to 4C.

Figure 4A shows that there are substantial differences in racial disparities based on application quality. White and API applicants' advantages are substantially higher for low-quality applications. Racial disparities are substantially lower among high-quality applications, with some statistically insignificant from zero.[34]

Similarly, Figure 4B reveals that the gender disparity is also significantly greater for low-quality patent applications. Figure 4C finds that public firms enjoy substantial advantages for all applications filed, with a greater advantage for low-quality applications (relative to individual inventors). Private firms only experience an advantage for low-quality applications.

This confirms the predictions generated under the correlation neglect model and strongly suggests

[34]The difference in the proportional effect is greater because low-quality applications have lower predicted acceptances.

the presence of correlation neglect in examination review.

7. Conclusion

We introduce Causal Human+Machine Learning (C-HML), combining advanced language models and human perception tests to address fundamental limitations of Causal ML for texts and provide robust causal identification. Using patenting effectively as our laboratory, we find evidence of discrimination by USPTO against underrepresented groups and individuals.

Our study is the first in economics to systematically investigate patenting disparities across multiple dimensions and their interactions/intersectionality. We document persistent discrimination towards female and Black inventors in patenting approvals. Interactions between these biases are particularly concerning: discrimination compounds for vulnerable groups, with individual Black female inventors facing the steepest barriers. We also identify a novel "affiliation bias" in innovation that favors, in ways unrelated to invention quality, applicants from public and private firms to an extent more significant than demographic discrimination. Homophily does not fully account for the disparities we document. Instead, a parsimonious theoretical model based on correlation neglect rationalizes all our findings: examiners may unintentionally discriminate by mistakenly treating correlated signals as independent information. Correlation neglect predicts heterogeneous discrimination, with more pronounced gaps for low quality applications, which our data corroborates.

If the policy goal is to circumvent this persistent roadblock, improving the education and living environments for better development of skills in under-represented innovators and entrepreneurs is not sufficient. Government agencies could consider adopting double-blind patent applications to address discrimination, perhaps in conjunction with broader incentives and social programs to rectify structural disparities.

More broadly, C-HML offers a scalable and rigorous toolkit for auditing and mitigating unfair bias in a wide range of critical economic domains involving decision-makers relying on human perceptions, applications/requests by individuals involving texts, and allocation of scarce resources (e.g., loan applications, hiring decisions, academic papers, and grant evaluations). In many such cases, if field experiments are infeasible, C-HML promises to distinguish correlations between demographics and outcomes from human bias, which is critical to providing solutions to such previously unresolved issues.

References

Aghion, P., U. Akcigit, A. Bergeaud, R. Blundell, and D. Hémous. 2019. Innovation and top income inequality. *The Review of Economic Studies* 86:1–45.

Aghion, P., U. Akcigit, A. Hyytinen, and O. Toivanen. 2017. The social origins of inventors. Working Paper, National Bureau of Economic Research.

Akcigit, U., J. Grigsby, and T. Nicholas. 2017. The rise of american ingenuity: Innovation and inventors of the golden age. Working Paper, National Bureau of Economic Research.

Alsan, M., O. Garrick, and G. Graziani. 2019. Does diversity matter for health? Experimental evidence from oakland. *American Economic Review* 109:4071–111.

Altonji, J. G., and R. M. Blank. 1999. Race and gender in the labor market. *Handbook of Labor Economics* 3:3143–259.

Arrow, K. J. 1973. The theory of discrimination. In *Discrimination in Labor Markets*, 3–33. Princeton University Press.

Athey, S., and G. W. Imbens. 2015. Machine learning methods for estimating heterogeneous causal effects. Working Paper.

Bailey, M., R. Cao, T. Kuchler, J. Stroebel, and A. Wong. 2018. Social connectedness: Measurement, determinants, and effects. *Journal of Economic Perspectives* 32:259–80.

Bapna, S., and M. Ganco. 2021. Gender gaps in equity crowdfunding: Evidence from a randomized field experiment. *Management Science* 67:2679–710.

Bartoš, V., M. Bauer, J. Chytilová, and F. Matejka. 2016. Attention discrimination: Theory and field experiments with monitoring information acquisition. *American Economic Review* 106:1437–75.

Becker, G. S. 1957. *The economics of discrimination.* University of Chicago Press.

Bell, A., R. Chetty, X. Jaravel, N. Petkova, and J. Van Reenen. 2019. Who becomes an inventor in america? The importance of exposure to innovation. *Quarterly Journal of Economics* 134:647–713.

Bernstein, S., A. Korteweg, and K. Laws. 2017. Attracting early-stage investors: Evidence from a randomized field experiment. *Journal of Finance* 72:509–38.

Bertrand, M. 2011. New perspectives on gender. In *Handbook of Labor Economics*, vol. 4, 1543–90. Elsevier.

Bertrand, M., D. Chugh, and S. Mullainathan. 2005. Implicit discrimination. *American Economic Review* 95:94–8.

Bertrand, M., and E. Duflo. 2017. Field experiments on discrimination. *Handbook of Economic Field Experiments* 1:309–93.

Bertrand, M., E. Kamenica, and J. Pan. 2015. Gender identity and relative income within households. *Quarterly Journal of Economics* 130:571–614.

Bertrand, M., and S. Mullainathan. 2004. Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *American Economic Review* 94:991–1013.

Blau, F. D., and L. M. Kahn. 2017. The gender wage gap: Extent, trends, and explanations. *Journal of Economic Literature* 55:789–865.

Bloom, N., M. Schankerman, and J. Van Reenen. 2013. Identifying technology spillovers and product market rivalry. *Econometrica* 81:1347–93.

Bordalo, P., K. Coffman, N. Gennaioli, and A. Shleifer. 2016. Stereotypes. *Quarterly Journal of Economics* 131:1753–94.

Brouillette, J.-F. 2025. The lost marie curies and foregone economic growth. Working Paper, HEC Montréal.

Brown, T., B. Mann, N. Ryder, et al. 2020. Language models are few-shot learners. *Advances in Neural Information Processing Systems* 33:1877–901.

Chen, J., L. W. Cong, and S. Li. 2025. Endogenous misinformation and source authentication. *Working Paper* .

Chernenko, S., and D. S. Scharfstein. 2022. Racial disparities in the paycheck protection program. Working Paper.

———. 2023. The limits of algorithmic measures of race in studies of outcome disparities. Working Paper, Purdue Krannert, HBS.

Chernozhukov, V., W. K. Newey, and R. Singh. 2022. Automatic debiased machine learning of causal and structural effects. *Econometrica* 90:967–1027.

Chetty, R., N. Hendren, M. R. Jones, and S. R. Porter. 2020. Race and economic opportunity in the United States: An intergenerational perspective. *Quarterly Journal of Economics* 135:711–83.

Chetty, R., N. Hendren, P. Kline, and E. Saez. 2014. Where is the land of opportunity? The geography of intergenerational mobility in the United States. *Quarterly Journal of Economics* 129:1553–623.

Chien, C. V., and J. Grennan. 2024. Unpacking the innovator-inventor gap: Evidence from engineers. Working Paper, Berkeley Law, Emory.

Chien, C. V., J. Grennan, and J. Sandvik. 2025. Small-scale mentoring, large-scale innovation: Evidence from a superstar firm. Working Paper, Berkeley Law, Emory, UArizona.

Coluccia, D. M., G. Dossi, and S. Ottinger. 2023. Racial discrimination and lost innovation. Working Paper 5.

Cook, L. D., J. Gerson, and J. Kuan. 2022. Closing the innovation gap in pink and black. *Entrepreneurship and Innovation Policy and the Economy* 1:43–66.

Cook, L. D., and C. Kongcharoen. 2010. The idea gap in pink and black. Working Paper, National Bureau of Economic Research.

Cook, L. D., M. Marx, and E. Yimfor. 2022. Funding black high-growth startups. Working Paper, National Bureau of Economic Research.

Cullen, Z., and R. Perez-Truglia. 2023. The old boys' club: Schmoozing and the gender gap. *American Economic Review* 113:1703–40.

Dell, M. 2025. Deep learning for economists. *Journal of Economic Literature* 63:5–58. doi: 10.1257/jel.20241733.

DeMarzo, P. M., D. Vayanos, and J. Zwiebel. 2003. Persuasion bias, social influence, and unidimensional opinions. *Quarterly Journal of Economics* 118:909–68.

Derenoncourt, E., and C. Montialoux. 2021. Minimum wages and racial inequality. *Quarterly Journal of Economics* 136:169–228.

Diep-Nguyen, H., M. K. Price, and J. Yang. 2025. Racial and gender favoritism in crowdfunding: Evidence from the field. Working Paper, Purdue Krannert, Alabama, Nanyang Technological.

Doleac, J. L., and L. C. Stein. 2013. The visible hand: Race and online market outcomes. *The Economic Journal* 123:F469–92.

Drucker, P. 2014. *Innovation and entrepreneurship*. Routledge.

Ellison, G., E. L. Glaeser, and W. R. Kerr. 2010. What causes industry agglomeration? Evidence from coagglomeration patterns. *American Economic Review* 100:1195–213.

Enke, B., and F. Zimmermann. 2019. Correlation neglect in belief formation. *Review of Economic Studies* 86:313–32.

Eyster, E., and G. Weizsacker. 2011. Correlation neglect in financial decision-making. Working Paper, LSE Economics, Humboldt University.

Fairlie, R., A. Robb, and D. T. Robinson. 2022. Black and white: Access to capital among minority-owned start-ups. *Management Science* 68:2377–400.

Fairlie, R., and D. T. Robinson. 2023. Racial differences in access to capital for innovative start-ups. *Entrepreneurship and Innovation Policy and the Economy* 2:149–66.

Farre-Mensa, J., D. Hegde, and A. Ljungqvist. 2020. What is a patent worth? Evidence from the US patent "lottery". *Journal of Finance* 75:639–82.

Frakes, M. D., and M. F. Wasserman. 2017. Is the time allocated to review patent applications inducing examiners to grant invalid patents? Evidence from microlevel application data. *Review of Economics and Statistics* 99:550–63.

Fryer, Roland G., J., and S. D. Levitt. 2004. The causes and consequences of distinctively black names. *Quarterly Journal of Economics* 119:767–805.

Gafni, H., D. Marom, A. Robb, and O. Sade. 2021. Gender dynamics in crowdfunding (Kickstarter): Evidence on entrepreneurs, backers, and taste-based discrimination. *Review of Finance* 25:235–74.

Goldin, C. 2014. A grand gender convergence: Its last chapter. *American Economic Review* 104:1091–119.

Goldin, C., and C. Rouse. 2000. Orchestrating impartiality: The impact of "blind" auditions on female musicians. *American Economic Review* 90:715–41.

Gompers, P., J. Lerner, and D. Scharfstein. 2005. Entrepreneurial spawning: Public corporations and the genesis of new ventures, 1986 to 1999. *Journal of Finance* 60:577–614.

Gompers, P. A., W. Gornall, S. N. Kaplan, and I. A. Strebulaev. 2020. How do venture capitalists make decisions? *Journal of Financial Economics* 135:169–90.

Goncalves, F., and S. Mello. 2021. A few bad apples? Racial bias in policing. *American Economic Review* 111:1406–41.

Greenwald, D. L., S. T. Howell, C. Li, and E. Yimfor. 2024. Regulatory arbitrage or random errors? Implications of race prediction algorithms in fair lending analysis. *Journal of Financial Economics* 157:103857–.

Griffith, R., S. Redding, and J. V. Reenen. 2004. Mapping the two faces of R&D: Productivity growth in a panel of OECD industries. *Review of Economics and Statistics* 86:883–95.

Hacamo, I., and K. Kleiner. 2022. Forced entrepreneurs. *Journal of Finance* 77:49–83.

Hansen, A. L., and S. Kazinnik. 2023. Can ChatGPT decipher fedspeak? Working Paper, Federal Reserve Board.

He, S., L. Lv, A. Manela, and J. Wu. 2025. Chronologically consistent large language models. Working Paper, WashU Olin.

Hoberg, G., and A. Manela. 2025. The natural language of finance. *Available at SSRN 5119322* .

Hochberg, Y., A. Kakhbod, P. Li, and K. Sachdeva. 2023. Inventor gender and patent undercitation: Evidence from causal text estimation. Working Paper, National Bureau of Economic Research.

Hsieh, C.-T., E. Hurst, C. I. Jones, and P. J. Klenow. 2019. The allocation of talent and US economic growth. *Econometrica* 87:1439–74.

Hunt, J., J.-P. Garant, H. Herman, and D. J. Munroe. 2012. Why don't women patent? Working Paper, National Bureau of Economic Research.

Jensen, K., B. Kovács, and O. Sorenson. 2018. Gender differences in obtaining and maintaining patent rights. *Nature biotechnology* 36:307–9.

Jha, M., J. Qian, M. Weber, and B. Yang. 2023. ChatGPT and corporate policies. Working Paper, GSU Robinson, Chicago Booth.

Khetan, V., R. Ramnani, M. Anand, S. Sengupta, and A. E. Fano. 2022. Causal bert: Language models for causality detection between events expressed in text. In *Intelligent Computing: Proceedings of the 2021 Computing Conference, Volume 1*, 965–80. Springer.

Kim, A. G., M. Muhn, and V. V. Nikolaev. 2023. Bloated disclosures: Can ChatGPT help investors process information? Working Paper, Chicago Booth.

Kline, P., E. K. Rose, and C. R. Walters. 2022. Systemic discrimination among large US employers. *The Quarterly Journal of Economics* 137:1963–2036.

Koffi, M. 2025. Innovative ideas and gender (in) equality. *American Economic Review* 115:2207–36.

Lerner, J. 2002. 150 years of patent protection. *American Economic Review* 92:221–5.

Levy, G., and R. Razin. 2015. Correlation neglect, voting behavior, and information aggregation. *American Economic Review* 105:1634–45.

Li, Z., and S. Liu. 2023. Gender difference in innovation recognition: A textual analysis approach. *Available at SSRN 4527066* .

Mao, K., H. Lu, and S. J. Wang. 2025. The effect of incidental name similarity on favoritism in the chinese financial market. *Scientific Reports* 15:13077–.

McPherson, M., L. Smith-Lovin, and J. M. Cook. 2001. Birds of a feather: Homophily in social networks. *Annual Review of Sociology* 27:415–44.

Misra, D. 2019. Mish: A self regularized non-monotonic activation function. Working Paper, arXiv.

Phelps, E. S. 1972. The statistical theory of racism and sexism. *American Economic Review* 62:659–61.

Radford, A., K. Narasimhan, T. Salimans, et al. 2018. Improving language understanding by generative pre-training. Working Paper, OpenAI.

Sampat, B., and H. L. Williams. 2019. How do patents affect follow-on innovation? Evidence from the human genome. *American Economic Review* 109:203–36.

Schuster, W. M., M. Marcowitz-Bitton, and D. R. Gerhardt. 2022. The gender gap in academic patenting. *UC Davis L. Rev.* 56:759–.

Shu, T., X. Tian, and X. Zhan. 2022. Patent quality, firm value, and investor underreaction: Evidence from patent examiner busyness. *Journal of Financial Economics* 143:1043–69.

Tabakovic, H., and T. G. Wollmann. 2018. From revolving doors to regulatory capture? Evidence from patent examiners.

Veitch, V., D. Sridhar, and D. Blei. 2020. Adapting text embeddings for causal inference. In *Conference on Uncertainty in Artificial Intelligence*, 919–28. PMLR.

Waldfogel, J. 2023. The welfare effect of gender-inclusive intellectual property creation: Evidence from books. NBER Working Paper.

Wang, X., L. Wu, and L. M. Hitt. 2024. Social media alleviates venture capital funding inequality for women and less connected entrepreneurs. *Management Science* 70:1093–112.

Yang, S. 2025. Understanding innovation quality and success with large-language models. Working Paper.

Table 1. C-HML Simulation Results: Mean Estimates and Bias

This table reports the results from the C-HML simulation comparing three estimation methods: Human-Only (conditioning on proxy q_c), ML-Only (using machine learning predictions), and C-HML (combining human evaluation with ML predictions). The simulation generates data with true discrimination parameter $\beta_2 = 0.5$. Mean Estimate shows the average estimated coefficient across simulations. Bias is calculated as the difference between the mean estimate and the true parameter value. RMSE is the root mean squared error. Diagnostic statistics include the correlation between the proxy and true q_c, the proxy classification accuracy, and the correlation between ML predictions and human signals.

Method	Mean Estimate ($\hat{\beta}_2$)	Bias ($\hat{\beta}_2 - \beta_2$)
C-HML	0.442	-0.058
Human-Only	1.036	0.536
ML-Only	0.143	-0.357
Statistics:		
Proxy accuracy (corr):		0.816
Proxy classification accuracy:		80.4%
Correlation(q_{ML}, s_N):		0.651

Note: Results based on $1,000$ simulations with $n = 20,000$ observations each. True parameters: $\beta_1 = 1.0$, $\beta_2 = 0.5$. Variance parameters: $\mathrm{Var}(q_i) = 1.0$, $\mathrm{Var}(q_c) = 1.0$, $\mathrm{Var}(\epsilon_1) = 0.5$, $\mathrm{Var}(\epsilon_N) = 0.5$, Var(proxy error) $= 0.5$.

Table 2. Patent Application Acceptance and Race

This table reports the summary statistics of the acceptance rates of all patent applications filed, sorted by inventor race. Acceptances is a binary variable, equal to 1 for granted applications and 0 for rejected applications. Application Quality is the chance of acceptance as calculated by a neural network trained on all input variables. Race is determined by the plurality race calculated based on the Surgeo BIFSG algorithm, geo-coded to Alexandria, Virginia.

Panel A: Average Acceptance Rates by Race.

	Acceptances		
	Mean	SD	N
Plurality Race			
White	0.677	0.468	2,536,203
API	0.707	0.455	718,773
Hispanic	0.649	0.477	78,664
Black	0.606	0.489	37,177
Mixed	0.669	0.471	4,678
Indigenous American	0.613	0.490	75

Panel B: Average Predicted Acceptance Rate by Race.

	Application Quality							
	Mean	SD	10 Pct.	25 Pct.	Median	75 Pct.	90 Pct.	N
Plurality Race								
White	0.684	0.194	0.400	0.568	0.727	0.836	0.900	2,536,203
API	0.716	0.183	0.451	0.613	0.760	0.857	0.913	718,773
Hispanic	0.668	0.205	0.365	0.537	0.712	0.831	0.897	78,664
Black	0.635	0.219	0.311	0.491	0.675	0.812	0.890	37,177
Mixed	0.698	0.197	0.418	0.589	0.744	0.849	0.912	4,678
Indigenous American	0.672	0.216	0.400	0.522	0.714	0.863	0.912	75

Table 3. Regression of Application Acceptance on Race

This table reports the results of the regressions of application acceptance on inventor race. Inventor race is defined as the average race across inventors as identified by SURGEO. Application Quality is the neural-network predicted chance of application success. T-statistics are reported in parentheses, with standard errors clustered at the year level where indicated. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels, respectively.

	Acceptances					
	(1)	(2)	(3)	(4)	(5)	(6)
White	0.028***	0.028***	0.032***	0.014***	0.014***	0.016***
	(16.62)	(11.34)	(17.70)	(8.90)	(7.54)	(9.47)
Black	-0.043***	-0.043***	-0.036***	-0.016***	-0.016***	-0.014***
	(-14.69)	(-7.93)	(-7.43)	(-5.99)	(-4.37)	(-3.66)
API	0.058***	0.058***	0.053***	0.018***	0.018***	0.016***
	(33.11)	(11.53)	(12.54)	(10.69)	(5.79)	(5.42)
Indigenous American	-0.036	-0.036	-0.031	-0.039	-0.039	-0.035
	(-0.67)	(-0.64)	(-0.55)	(-0.77)	(-0.81)	(-0.71)
Mixed	0.020***	0.020***	0.016***	-0.005	-0.005	-0.007
	(2.85)	(3.30)	(2.65)	(-0.81)	(-1.17)	(-1.49)
Year-Fixed Effects	-	-	Y	-	-	Y
Clustered by Year	-	Y	Y	-	Y	Y
Application Quality Control	-	-	-	Y	Y	Y
N	3,375,570	3,375,570	3,375,570	3,375,570	3,375,570	3,375,570
Adjusted R^2	0.001	0.001	0.014	0.118	0.118	0.121

Table 4. Patent Application Acceptance and Gender

This table reports, in Panel A and B, the summary statistics of the acceptance rates of all patent applications filed, sorted by majority inventor gender. Acceptances is a binary variable, equal to 1 for granted applications and 0 for rejected applications. Predicted acceptance rate is the chance of acceptance as calculated by a neural network trained on all input variables. Gender is obtained from USPTO data. Patents with equal male and female inventorship are classified as female. Panel C reports the regression of application acceptance on inventor gender, where Application Quality is the neural-network predicted chance of application success. T-statistics are reported in parentheses, with standard errors clustered at the year level where indicated. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels, respectively.

Panel A: Average Acceptance Rates by Gender.

	Acceptances		
	Mean	SD	N
Majority Gender			
Female	0.638	0.481	389,362
Male	0.690	0.463	2,912,399

Panel B: Average Predicted Acceptance Rate by Gender.

	Predicted Acceptance Rate							
	Mean	SD	10 Pct.	25 Pct.	Median	75 Pct.	90 Pct.	N
Majority Gender								
Female	0.652	0.208	0.345	0.514	0.692	0.819	0.892	389,362
Male	0.696	0.190	0.420	0.585	0.739	0.843	0.904	2,912,399

Panel C: Regression of Application Acceptance on Gender

	Acceptances					
	(1)	(2)	(3)	(4)	(5)	(6)
Male	0.052***	0.052***	0.055***	0.016***	0.016***	0.018***
	(65.56)	(22.16)	(26.91)	(21.25)	(10.61)	(13.05)
Year-Fixed Effects	-	-	Y	-	-	Y
Clustered by Year	-	Y	Y	-	Y	Y
Application Quality Control	-	-	-	Y	Y	Y
N	3,301,761	3,301,761	3,301,761	3,301,761	3,301,761	3,301,761
Adjusted R^2	0.001	0.001	0.015	0.117	0.117	0.120

Table 5. Discrimination and Gendered Names

This table reports the results of the regressions of application acceptance on inventor gender. Table A reports regressions on perceived gender, sorted into two groups by actual gender. Table B performs regressions on actual gender, sorted by whether or not the inventor's name is obviously gendered. All variables are as defined in above tables. T-statistics are reported in parentheses, with standard errors clustered at the year level where indicated. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels, respectively.

Panel A: Perceived Name Gender by Actual Gender

	Acceptances			
	(1) Female	(2) Male	(3) Female	(4) Male
Perceived Male-ness	0.077***	0.059***	0.018***	0.011***
	(19.16)	(15.81)	(6.53)	(5.39)
Year Fixed Effects	Y	Y	Y	Y
Clustered by Year	Y	Y	Y	Y
Application Quality Control	-	-	Y	Y
N	129,978	1,405,341	129,978	1,405,341
Adjusted R^2	0.018	0.013	0.127	0.111

Panel B: Actual Gender by Gender Ambiguity

	Acceptances			
	(1) Ambiguous Names	(2) Gendered Names	(3) Ambiguous Names	(4) Gendered Names
Male	0.033***	0.078***	0.012***	0.024***
	(8.95)	(36.99)	(3.91)	(13.51)
Year Fixed Effects	Y	Y	Y	Y
Clustered by Year	Y	Y	Y	Y
Application Quality Control	-	-	Y	Y
N	429,052	1,535,319	429,052	1,535,319
Adjusted R^2	0.014	0.015	0.108	0.115

Table 6. Preference for Whites and Name Perception

This table reports the results of the regressions of application acceptance on perceived race. Table A reports regressions on name Whiteness, sorted into two groups by actual race (White vs. non-White), while Panel B reports regressions on name Blackness, sorted by Black vs. non-Black. All variables are as defined in above tables. T-statistics are reported in parentheses, with standard errors clustered at the year level where indicated. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels, respectively.

Panel A: Name Whiteness by True Race

	Acceptances			
	(1) White	(2) Non-White	(3) White	(4) Non-White
Perceived White-ness	0.009***	0.031***	0.005***	0.022***
	(4.74)	(5.48)	(3.05)	(5.18)
Year Fixed Effects	Y	Y	Y	Y
Clustered by Year	Y	Y	Y	Y
Application Quality Control	-	-	Y	Y
N	669,756	530,584	669,756	530,584
Adjusted R^2	0.012	0.014	0.110	0.108

Panel B: Name Blackness by True Race

	Acceptances			
	(1) Black	(2) Non-Black	(3) Black	(4) Non-Black
Perceived Black-ness	-0.046***	-0.003	-0.033***	-0.002
	(-5.40)	(-1.11)	(-4.79)	(-1.04)
Year Fixed Effects	Y	Y	Y	Y
Clustered by Year	Y	Y	Y	Y
Application Quality Control	-	-	Y	Y
N	17,633	1,182,707	17,633	1,182,707
Adjusted R^2	0.015	0.012	0.134	0.108

Table 7. Patent Application Acceptance and Applicant Type

This table reports, in Panels A and B, the summary statistics of the acceptance rates of all patent applications filed, sorted by applicant type (public firm, private firm, or individual inventor). Acceptances is a binary variable, equal to 1 for granted applications and 0 for rejected applications. Predicted acceptance rate is the chance of acceptance as calculated by a neural network trained on all input variables. Panel C reports the regression of acceptance on applicant type. Individual inventor is omitted from the regression due to collinearity. Application Quality is the neural-network predicted chance of application success. T-statistics are reported in parentheses, with standard errors clustered at the year level where indicated. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels, respectively.

Panel A: Average Acceptance Rates by Applicant Type.

	Acceptances		
	Mean	SD	N
Applicant Type			
Public	0.751	0.433	1,465,572
Private	0.667	0.471	846,644
Individual	0.601	0.490	1,075,552

Panel B: Average Predicted Acceptance Rate by Applicant Type.

	Predicted Acceptance Rate							
	Mean	SD	10 Pct.	25 Pct.	Median	75 Pct.	90 Pct.	N
Applicant Type								
Public	0.746	0.152	0.533	0.668	0.782	0.859	0.906	1,465,572
Private	0.655	0.172	0.415	0.553	0.685	0.785	0.850	846,644
Individual	0.618	0.190	0.347	0.491	0.643	0.767	0.848	1,075,552

Panel C: Regression of Application Acceptance on Applicant Type.

	Acceptances					
	(1)	(2)	(3)	(4)	(5)	(6)
Public Firms	0.150***	0.150***	0.122***	0.075***	0.075***	0.066***
	(255.92)	(11.15)	(12.56)	(130.94)	(12.50)	(13.33)
Private Firms	0.066***	0.066***	0.033**	0.030***	0.030***	0.021***
	(98.59)	(3.90)	(2.47)	(47.46)	(3.65)	(2.70)
Year-Fixed Effects	-	-	Y	-	-	Y
Clustered by Year	-	Y	Y	-	Y	Y
Application Quality Control	-	-	-	Y	Y	Y
N	3,387,768	3,387,768	3,387,768	3,387,768	3,387,768	3,387,768
Adjusted R^2	0.019	0.019	0.026	0.112	0.112	0.115

Table 8. Inventor Name Frequency, Location and Application Acceptances

This table reports the results of the regressions of application acceptance on inventors' name frequency and location. Name frequency is the natural logarithm of the total occurrences of a name over all patent applications filed plus one. All variables are as defined in the above tables. A big city is any of the 50 most populous cities in the United States, and an application is from a big city if any of its inventors is from a big city. T-statistics are reported in parentheses, with standard errors clustered at the year level where indicated. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels, respectively.

Panel A: Regression of Acceptance on Name Frequency.

	Acceptances					
	(1)	(2)	(3)	(4)	(5)	(6)
Name Frequency	0.006***	0.006***	0.008***	0.004***	0.004***	0.005***
	(38.95)	(5.49)	(11.58)	(24.54)	(6.04)	(9.43)
Year-Fixed Effects	-	-	Y	-	-	Y
Clustered by Year	-	Y	Y	-	Y	Y
Application Quality Control	-	-	-	Y	Y	Y
N	3,301,703	3,301,703	3,301,703	3,301,703	3,301,703	3,301,703
Adjusted R^2	0.000	0.000	0.014	0.117	0.117	0.120

Panel B: Name Frequency Regression with Race and Gender Controls.

	Acceptances					
	(1)	(2)	(3)	(4)	(5)	(6)
Name Frequency	0.008***	0.008***	0.009***	0.004***	0.004***	0.005***
	(44.06)	(9.90)	(17.76)	(23.76)	(8.75)	(13.48)
Year-Fixed Effects	-	-	Y	-	-	Y
Clustered by Year	-	Y	Y	-	Y	Y
Application Quality Control	-	-	-	Y	Y	Y
N	3,301,703	3,301,703	3,301,703	3,301,703	3,301,703	3,301,703
Adjusted R^2	0.003	0.003	0.017	0.117	0.117	0.120

Panel C: Regression of Acceptance on Inventor Location.

	Acceptances					
	(1)	(2)	(3)	(4)	(5)	(6)
Large City	0.019***	0.019***	0.012***	0.008***	0.008**	0.005*
	(28.76)	(6.65)	(5.85)	(13.30)	(2.59)	(2.07)
Year-Fixed Effects	-	-	Y	-	-	Y
Clustered by Year	-	Y	Y	-	Y	Y
Application Quality Control	-	-	-	Y	Y	Y
N	2,231,139	2,231,139	2,231,139	2,231,139	2,231,139	2,231,139
Adjusted R^2	0.000	0.000	0.013	0.127	0.127	0.130

Table 9. Discrimination and Examiner Demographics

This table tests for the presence of examiner gender by regressing application acceptance on inventor gender and race subsamples based on the gender/race of the examiner. All variables are as defined in the above tables. T-statistics are reported in parentheses, with standard errors clustered at the year level where indicated. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels, respectively.

Panel A: Gender.

	Male Examiners		Female Examiners	
	(1)	(2)	(3)	(4)
Male	0.048***	0.016***	0.064***	0.018***
	(19.00)	(8.80)	(33.88)	(12.61)
Year-Fixed Effects	Y	Y	Y	Y
Clustered by Year	Y	Y	Y	Y
Application Quality Control	-	Y	-	Y
N	2,241,079	2,241,079	851,187	851,187
R^2	0.015	0.114	0.016	0.128

Panel B: Race.

	White Examiners		Black Examiners		API Examiners	
	(1)	(2)	(3)	(4)	(5)	(6)
White	0.032***	0.016***	0.013	-0.000	0.040***	0.016
	(16.26)	(8.58)	(0.82)	(-0.00)	(3.43)	(1.57)
Black	-0.037***	-0.013***	-0.101**	-0.080**	-0.025	-0.021
	(-7.60)	(-3.41)	(-2.76)	(-2.30)	(-1.62)	(-1.39)
API	0.051***	0.016***	0.025	0.000	0.081***	0.033**
	(13.83)	(6.17)	(1.33)	(0.02)	(6.28)	(2.91)
Year Fixed Effects	Y	Y	Y	Y	Y	Y
Clustered by Year	Y	Y	Y	Y	Y	Y
Application Quality Control	-	Y	-	Y	-	Y
N	3,091,581	3,091,581	33,320	33,320	79,249	79,249
Adjusted R^2	0.015	0.121	0.009	0.093	0.018	0.128

Table 10. Examiner-Inventor Name Homophily and Application Success

This table tests for name-based homophily by regressing application success on an indicator for whether the examiner and inventor share the same first or last name. The "Same Name" variable is a binary indicator for a name match, either first or last name. T-statistics are reported in parentheses, with standard errors clustered at the year level. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels, respectively.

	First Name Match		Last Name Match	
	(1)	(2)	(3)	(4)
Same Name	-0.007	-0.004	0.046***	0.024***
	(-0.90)	(-0.58)	(7.30)	(4.09)
Year-Fixed Effects	Y	Y	Y	Y
Clustered by Year	Y	Y	Y	Y
Application Quality Control	-	Y	-	Y
Matching Race Control	Y	Y	Y	Y
N	3,213,215	3,213,215	3,213,215	3,213,215
Adjusted R^2	0.014	0.120	0.014	0.120

Table 11. Public Firms: Application Acceptance and Race

This table reports, in Panel A, the summary statistics of the acceptance rates and Application Quality of all patent applications filed by public firms, sorted by inventor race. All variables are as defined in Table 2. In Panel B, t-statistics are reported in parentheses, with standard errors clustered at the firm level. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels, respectively.

Panel A: Average Acceptance Rates and Application Quality by Race.

	Application Acceptances		Application Quality		
	Mean	SD	Mean	SD	N
Plurality Race					
White	0.751	0.432	0.750	0.167	986,560
API	0.754	0.431	0.760	0.161	403,499
Black	0.755	0.430	0.758	0.164	29,348
Hispanic	0.749	0.434	0.750	0.171	10,562
Mixed	0.744	0.436	0.774	0.151	2,185
Indigenous American	0.600	0.503	0.798	0.232	20

Panel B: Regression of Application Acceptance on Race.

	Acceptances					
	(1)	(2)	(3)	(4)	(5)	(6)
White	-0.001	-0.000	0.013***	0.004	0.004	0.008**
	(-0.15)	(-0.03)	(2.82)	(1.58)	(0.95)	(2.12)
Black	-0.008	-0.005	-0.000	-0.001	-0.001	-0.001
	(-0.78)	(-0.56)	(-0.03)	(-0.18)	(-0.08)	(-0.08)
API	-0.001	0.005	0.007	-0.002	-0.001	0.003
	(-0.09)	(0.44)	(1.02)	(-0.77)	(-0.18)	(0.65)
Indigenous American	-0.124	-0.129	-0.200**	-0.167*	-0.168*	-0.191**
	(-1.07)	(-1.14)	(-2.02)	(-1.78)	(-1.90)	(-2.29)
Mixed	-0.005	-0.003	-0.005	-0.020**	-0.020	-0.014
	(-0.36)	(-0.19)	(-0.43)	(-2.05)	(-1.58)	(-1.36)
Firm & Year Fixed Effects	-	-	Y	-	-	Y
Firm Size Control	-	Y	Y	-	Y	Y
Application Quality Control	-	-	-	Y	Y	Y
N	1,301,539	1,301,539	1,301,539	1,301,539	1,301,539	1,301,539
Adjusted R^2	0	0.002	0.053	0.094	0.094	0.111

Table 12. Private Firms: Application Acceptance and Race

This table reports, in Panel A, the summary statistics of the acceptance rates and Application Quality of all patent applications filed by private firms, sorted by inventor race. All variables are as defined in Table 2. In Panel B, t-statistics are reported in parentheses, with standard errors clustered at the firm level. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels, respectively.

Panel A: Average Acceptance Rates and Application Quality by Race.

	Application Acceptances		Application Quality		
	Mean	SD	Mean	SD	N
Plurality Race					
White	0.667	0.471	0.673	0.189	688,473
API	0.667	0.471	0.673	0.187	145,726
Black	0.647	0.478	0.660	0.191	20,588
Hispanic	0.640	0.480	0.659	0.195	9,470
Mixed	0.694	0.461	0.677	0.187	993
Indigenous American	0.524	0.512	0.595	0.210	21

Panel B: Regression of Application Acceptance on Race.

	Acceptances					
	(1)	(2)	(3)	(4)	(5)	(6)
White	0.020***	0.020***	0.022***	0.010***	0.010***	0.011***
	(5.93)	(6.13)	(6.51)	(3.20)	(3.42)	(3.68)
Black	-0.008	-0.008	-0.003	-0.008	-0.008	-0.006
	(-1.31)	(-0.95)	(-0.38)	(-1.36)	(-1.08)	(-0.89)
API	0.020***	0.020*	0.012	0.010***	0.010	0.007
	(5.65)	(1.89)	(1.23)	(2.96)	(1.35)	(0.96)
Indigenous American	-0.124	-0.124	-0.132	-0.077	-0.077	-0.079
	(-1.20)	(-1.05)	(-1.13)	(-0.78)	(-0.64)	(-0.67)
Mixed	0.046***	0.046**	0.038**	0.033**	0.033**	0.031**
	(3.03)	(2.92)	(2.38)	(2.28)	(2.38)	(2.20)
Year-Fixed Effects	-	-	Y	-	-	Y
Clustered by Year	-	Y	Y	-	Y	Y
Application Quality Control	-	-	-	Y	Y	Y
N	865,271	865,271	865,271	865,271	865,271	865,271
Adjusted R^2	0.000	0.000	0.015	0.085	0.085	0.089

Table 13. Individual Inventors: Application Acceptance and Race

This table reports, in Panel A, the summary statistics of the acceptance rates and Application Quality of all patent applications filed by individual inventors, sorted by inventor race. All variables are as defined in Table 2. In Panel B, t-statistics are reported in parentheses, with standard errors clustered at the firm level. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels, respectively

Panel A: Average Acceptance Rates and Application Quality by Race.

	Application Acceptances		Application Quality		
	Mean	SD	Mean	SD	N
Plurality Race					
White	0.600	0.490	0.619	0.203	861,170
API	0.629	0.483	0.650	0.197	169,548
Black	0.543	0.498	0.581	0.213	28,728
Hispanic	0.500	0.500	0.551	0.223	17,145
Mixed	0.543	0.498	0.601	0.217	1,500
Indigenous American	0.676	0.475	0.645	0.181	34

Panel B: Regression of Application Acceptance on Race.

	Acceptances					
	(1)	(2)	(3)	(4)	(5)	(6)
White	0.058***	0.058***	0.055***	0.026***	0.026***	0.027***
	(19.73)	(14.09)	(14.88)	(9.58)	(8.66)	(10.26)
Black	-0.043***	-0.043***	-0.041***	-0.018***	-0.018***	-0.017***
	(-9.12)	(-6.09)	(-5.98)	(-4.09)	(-3.88)	(-3.43)
API	0.087***	0.087***	0.084***	0.029***	0.029***	0.027***
	(27.84)	(16.00)	(17.12)	(9.76)	(6.70)	(7.19)
Indigenous American	0.134	0.134	0.147	0.080	0.080	0.085
	(1.59)	(1.64)	(1.74)	(1.02)	(1.16)	(1.19)
Mixed	0.001	0.001	0.001	-0.016	-0.016	-0.018
	(0.06)	(0.04)	(0.05)	(-1.34)	(-1.21)	(-1.33)
Year-Fixed Effects	-	-	Y	-	-	Y
Clustered by Year	-	Y	Y	-	Y	Y
Application Quality Control	-	-	-	Y	Y	Y
N	1,078,125	1,078,125	1,078,125	1,078,125	1,078,125	1,078,125
Adjusted R^2	0.002	0.002	0.005	0.123	0.123	0.126

Table 14. Application Acceptance and Gender by Applicant Type

This table reports the results of the regressions of application acceptance on inventor gender, sorted into three groups: public firms, private firms, and individual inventors. All variables are as defined in above tables. T-statistics are reported in parentheses, with standard errors clustered at the year level where indicated. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels, respectively.

Panel A: Public Firms.

	Acceptances					
	(1)	(2)	(3)	(4)	(5)	(6)
Male	0.046***	0.044***	0.032***	0.017***	0.017***	0.015***
	(9.96)	(9.20)	(8.90)	(6.31)	(6.19)	(6.89)
Firm & Year Fixed Effects	-	-	Y	-	-	Y
Firm Size Control	-	Y	Y	-	Y	Y
Application Quality Control	-	-	-	Y	Y	Y
N	1,273,966	1,273,966	1,273,966	1,273,966	1,273,966	1,273,966
Adjusted R^2	0.001	0.003	0.046	0.093	0.093	0.109

Panel B: Private Firms.

	Acceptances					
	(1)	(2)	(3)	(4)	(5)	(6)
Male	0.046***	0.046***	0.050***	0.016***	0.016***	0.018***
	(28.25)	(17.12)	(18.57)	(10.11)	(7.21)	(8.08)
Year-Fixed Effects	-	-	Y	-	-	Y
Clustered by Year	-	Y	Y	-	Y	Y
Application Quality Control	-	-	-	Y	Y	Y
N	844,672	844,672	844,672	844,672	844,672	844,672
Adjusted R^2	0.001	0.001	0.016	0.085	0.085	0.089

Panel C: Individual Inventors.

	Acceptances					
	(1)	(2)	(3)	(4)	(5)	(6)
Male	0.070***	0.070***	0.069***	0.019***	0.019***	0.019***
	(47.00)	(29.94)	(24.97)	(13.45)	(10.31)	(10.08)
Year-Fixed Effects	-	-	Y	-	-	Y
Clustered by Year	-	Y	Y	-	Y	Y
Application Quality Control	-	-	-	Y	Y	Y
N	1,040,739	1,040,739	1,040,739	1,040,739	1,040,739	1,040,739
Adjusted R^2	0.002	0.002	0.006	0.123	0.123	0.125

Table 15. Patent Application Acceptance and Race: Sorted by Gender

This table reports the results of the regressions of application acceptance on inventor race, sorted into two groups by inventor gender. All variables are as defined in above tables. T-statistics are reported in parentheses, with standard errors clustered at the year level where indicated. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels, respectively.

Panel A: Female Inventors.

	Acceptances					
	(1)	(2)	(3)	(4)	(5)	(6)
White	0.022***	0.022***	0.024***	0.017***	0.017***	0.019***
	(4.37)	(3.64)	(4.22)	(3.50)	(3.13)	(3.62)
Black	-0.038***	-0.038***	-0.033***	-0.007	-0.007	-0.004
	(-4.75)	(-3.44)	(-3.08)	(-0.93)	(-0.80)	(-0.50)
API	0.113***	0.113***	0.104***	0.037***	0.037***	0.035***
	(22.15)	(10.75)	(11.18)	(7.65)	(4.67)	(4.33)
Indigenous American	-0.144	-0.144	-0.149	-0.047	-0.047	-0.047
	(-1.20)	(-0.87)	(-0.86)	(-0.42)	(-0.29)	(-0.28)
Mixed	0.042**	0.042*	0.035	0.019	0.019	0.019
	(2.02)	(1.93)	(1.57)	(0.96)	(0.93)	(0.94)
Year-Fixed Effects	-	-	Y	-	-	Y
Clustered by Year	-	Y	Y	-	Y	Y
Application Quality Control	-	-	-	Y	Y	Y
N	389,362	389,362	389,362	389,362	389,362	389,362
Adjusted R^2	0.009	0.009	0.022	0.131	0.131	0.135

Panel B: Male Inventors.

	Acceptances					
	(1)	(2)	(3)	(4)	(5)	(6)
White	0.024***	0.024***	0.029***	0.012***	0.012***	0.014***
	(13.49)	(8.83)	(13.59)	(7.18)	(5.85)	(7.07)
Black	-0.039***	-0.039***	-0.033***	-0.017***	-0.017***	-0.014***
	(-12.29)	(-7.60)	(-6.92)	(-5.47)	(-4.28)	(-3.63)
API	0.052***	0.052***	0.047***	0.016***	0.016***	0.014***
	(27.34)	(13.10)	(14.99)	(9.12)	(6.91)	(6.50)
Indigenous American	0.005	0.005	0.010	-0.033	-0.033	-0.028
	(0.08)	(0.07)	(0.16)	(-0.57)	(-0.58)	(-0.48)
Mixed	0.013	0.013*	0.010	-0.008	-0.008	-0.009*
	(1.64)	(1.87)	(1.54)	(-1.15)	(-1.58)	(-1.81)
Year-Fixed Effects	-	-	Y	-	-	Y
Clustered by Year	-	Y	Y	-	Y	Y
Application Quality Control	-	-	-	Y	Y	Y
N	2,912,399	2,912,399	2,912,399	2,912,399	2,912,399	2,912,399
Adjusted R^2	0.001	0.001	0.014	0.113	0.113	0.117

Table 16. Patent Application Acceptance and Gender: Sorted by Race

This table reports the results of the regressions of application acceptance on inventor gender, sorted into groups by inventor race. All variables are as defined in above tables. T-statistics are reported in parentheses, with standard errors clustered at the year level where indicated. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels, respectively.

Panel A: White, Black, and API Inventors.

	Acceptances					
	(1) White	(2) Black	(3) API	(4) White	(5) Black	(6) API
Male	0.083***	0.077***	0.020***	0.025***	0.014*	0.007***
	(43.70)	(9.06)	(8.34)	(15.07)	(1.85)	(3.46)
Year Fixed Effects	Y	Y	Y	Y	Y	Y
Clustered by Year	Y	Y	Y	Y	Y	Y
Application Quality Control	-	-	-	Y	Y	Y
N	2,499,887	35,862	684,287	2,499,887	35,862	684,287
Adjusted R^2	0.015	0.015	0.014	0.119	0.160	0.114

Panel B: Hispanic, Indigenous, and Mixed Inventors.

	Acceptances					
	(1) Hispanic	(2) Indigenous	(3) Mixed	(4) Hispanic	(5) Indigenous	(6) Mixed
Male	0.078***	0.307	0.053**	0.027***	0.142	-0.005
	(12.16)	(1.38)	(2.39)	(4.19)	(0.59)	(-0.21)
Year Fixed Effects	Y	Y	Y	Y	Y	Y
Clustered by Year	Y	Y	Y	Y	Y	Y
Application Quality Control	-	-	-	Y	Y	Y
N	77,230	70	4,422	77,230	70	4,422
Adjusted R^2	0.018	0.039	0.009	0.139	0.145	0.137

Appendix A: Theoretical Notation and Proofs

Table A.1. Theoretical Notation

Symbol	Description	Status at decision time
q	Latent, legally relevant application quality (merit); $q = q_i + q_c$.	Latent
q_i	Idiosyncratic idea/content component of quality.	Latent
G, R	Binary indicators (e.g., G=1 female, R=1 non-White).	Latent
$G \times R$	Gender–race interaction.	Latent
q_c	Quality component explained by demographics	Latent
s_1	Textual signal from the application; $s_1 = q + \epsilon_1$.	Observed
s_1'	Mis-specified textual signal under correlation neglect; $s_1' = q_i + \epsilon_1$.	Modeling construct
s_N	Signal from perceived demographics (e.g., from names); $s_N = q_c + \epsilon_N$.	Observed (perception)
y	Examiner decision for acceptance.	Outcome
β_1, β_2	Weights on s_1 and s_N in $y = \beta_1 s_1 + \beta_2 s_N$.	Latent (reduced-form behavior)
β_1', β_2'	Weights under correlation neglect in $y = \beta_1' s_1' + \beta_2' s_N$.	Latent (reduced-form behavior)
$\gamma_G, \gamma_R, \gamma_{GR}$	Reduced-form outcome coefficients on $G, R, G\times R$ with $\gamma_G = \beta_2\beta_G$, $\gamma_R = \beta_2\beta_R$, $\gamma_{GR} = \beta_2\beta_{GR}$.	Latent
ϵ_1, ϵ_N	Noise in s_1 (or s_1') and s_N.	Latent
σ_i^2, σ_c^2	Variances of q_i and q_c.	Parameter (assumed/belief)
$\sigma_{\epsilon_1}^2, \sigma_{\epsilon_N}^2$	Variances of ϵ_1 and ϵ_N.	Parameter (assumed/belief)
σ^2	$\sigma^2 = \sigma_i^2 + \dfrac{\sigma_c^2\sigma_{\epsilon_N}^2}{\sigma_c^2 + \sigma_{\epsilon_N}^2}$.	Parameter (derived)
q_{ML}	Econometrician's proxy for quality; $q_{\text{ML}} = q + \epsilon_{\text{ML}}$.	Not available to examiner
$\epsilon_{\text{ML}}, \sigma_{\text{ML}}^2$	ML measurement error and its variance.	Not available to examiner
q_c^{proxy}	Proxy for q_c used by researcher; $q_c^{\text{proxy}} = q_c + \epsilon_{\text{proxy}}$.	Econometrician only
$\epsilon_{\text{proxy}}, \sigma_{\text{proxy}}^2$	Proxy error and its variance.	Parameter (econometrician only)

"Observed" means directly available to the examiner at decision time; "Observed (perception)" is the examiner's inference from cues (e.g., names) and may be noisy. "Latent" are structural objects not directly seen by the examiner. "Outcome" is the realized choice. Parameters are primitives/beliefs, not observed. The econometrician observes y ex-post and constructs q_{ML} and q_c^{proxy}.

Proof of Proposition 1.

Let $\beta_2 = 0$. Then $y = \beta_1(q + \epsilon_1) = \beta_1 q_i + \beta_1 q_c + \beta_1 \epsilon_1$. Then

$$corr(y, q_c) = \frac{\mathrm{Cov}(y, q_c)}{\sqrt{\mathrm{Var}(y)\,\mathrm{Var}(q_c)}}. \tag{15}$$

By the independence of q_c from q_i and ϵ_1, we have

$$corr(y, q_c) = \frac{\beta_1 \,\mathrm{Cov}(q_c, q_c)}{\sqrt{\mathrm{Var}(y)\,\mathrm{Var}(q_c)}} = \beta_1 \sqrt{\frac{\mathrm{Var}(q_c)}{\mathrm{Var}(y)}} = \beta_1 \frac{\sigma_{q_c}}{\sigma_y} \tag{16}$$

Clearly, this takes non-zero values if $\beta_1 \neq 0$. Since $\beta_2 = 0$, β_1 must be non-zero as otherwise all applications are identically of 0 quality. Since $corr(y, q_c) \neq 0$, we have $\hat{\beta}_2 \neq 0$ and $\hat{\beta}_2 \neq \beta_2$. □

Proof of Proposition 2.

We first rewrite y in the form

$$y = \beta_1 q + \beta_2 q_c + (\beta_1 \epsilon_1 + \beta_2 \epsilon_N) = \beta_1(q_{ML}) + \beta_2 q_c + u, \tag{17}$$

$$u = \beta_1 \epsilon_1 + \beta_2 \epsilon_N - \beta_1 \epsilon_{ML}. \tag{18}$$

Now, let r_{q_c} be the residual obtained from regressing q_c on q_{ML}, i.e.,

$$r_{q_c} = q_c - \gamma q_{ML}, \gamma = \frac{\mathrm{Cov}(q_{ML}, q_c)}{\mathrm{Var}(q_{ML})} \tag{19}$$

Then, by the Frisch-Waugh-Lovell theorem, the expression for $\hat{\beta}_2$ in the estimated regression is given by

$$\hat{\beta}_2 = \frac{\mathrm{Cov}(r_{q_c}, y)}{\mathrm{Var}(r_{q_c})}. \tag{20}$$

Recall that $y = \beta_1 q_{ML} + \beta_2 q_c + u$ and that $\mathrm{Cov}(r_{q_c}, q_{ML}) = 0$ by definition. Then, we can rewrite

$$\hat{\beta}_2 = \frac{\beta_2 \,\mathrm{Var}(r_{q_c}) + \mathrm{Cov}(r_{q_c}, u)}{\mathrm{Var}(r_{q_c})} = \beta_2 + \frac{\mathrm{Cov}(r_{q_c}, u)}{\mathrm{Var}(r_{q_c})}. \tag{21}$$

Using the fact that $r_{q_c} = q_c - \gamma q_{ML}$, we have $\mathrm{Cov}(r_{q_c}, u) = \mathrm{Cov}(q_c, u) - \gamma\,\mathrm{Cov}(q_{ML}, u)$. u is a sum of errors independent from q_c, so

$$\operatorname{Cov}(r_{q_c}, u) = -\gamma \operatorname{Cov}\left(q + \epsilon_{ML}, \beta_1(\epsilon_1 - \epsilon_{ML}) + \beta_2\epsilon_N\right) = \gamma\beta_1 \operatorname{Var}(\epsilon_{ML}). \tag{22}$$

Additionally, note that

$$\operatorname{Var}(r_{q_c}) = \operatorname{Var}(q_c) - \frac{\operatorname{Cov}(q_{ML}, q_c)^2}{\operatorname{Var}(q_{ML})}. \tag{23}$$

Thus, with some simplification, we have that

$$\hat{\beta}_2 - \beta_2 = \frac{\beta_1 \operatorname{Cov}(q_{ML}, q_c) \operatorname{Var}(\epsilon_{ML})}{\operatorname{Var}(q_{ML}) \operatorname{Var}(q_c) - \operatorname{Cov}(q_{ML}, q_c)^2}. \tag{24}$$

Note that $\operatorname{Cov}(q_{ML}, q_c) = \operatorname{Var}(q_c) = \sigma^2_{q_c}$, so

$$\hat{\beta}_2 - \beta_2 = \frac{\beta_1 \sigma^2_{q_c} \sigma^2_{ML}}{\sigma^2_{q_{ML}} \sigma^2_{q_c} - \sigma^4_{q_c}} = \beta_1 \frac{\sigma^2_{ML}}{\sigma^2_{q_{ML}} - \sigma^2_{q_c}} = \beta_1 \frac{\sigma^2_{ML}}{\sigma^2_{q_i} + \sigma^2_{ML}}, \tag{25}$$

the desired expression. □

Proof of Proposition 3.

We will apply the Frisch-Waugh-Lovell theorem to find $\hat{\beta}_2$. Again, let

$$r_{s_N} = s_N - \gamma q_{ML}, \gamma = \frac{\operatorname{Cov}(q_{ML}, s_N)}{\operatorname{Var}(q_{ML})}. \tag{26}$$

Then,

$$\hat{\beta}_2 = \frac{\operatorname{Cov}(r_{s_N}, y)}{\operatorname{Var}(r_{s_N})}. \tag{27}$$

Before proceeding, we recall that $s_N = q_c + \epsilon_N$, so with q_c held constant, ϵ_N is the source of all variation in s_N. Consequently,

$$\operatorname{Cov}(s_N, q_{ML}) = \operatorname{Cov}(q_c + \epsilon_N, q_{ML}) = \operatorname{Cov}(\epsilon_N, q_{ML}) = 0. \tag{28}$$

This is the numerator of the regression coefficient of s_N on q_{ML}, so $\gamma = 0$ and r_{s_N} simply equals s_N.

Returning to the expression for $\hat{\beta}_2$, we have

$$\hat{\beta}_2 = \frac{\mathrm{Cov}(s_N, y)}{\mathrm{Var}(s_N)} = \frac{\beta_1 \,\mathrm{Cov}(s_N, s_1) + \beta_2 \,\mathrm{Var}(s_N)}{\mathrm{Var}(s_N)}. \tag{29}$$

Because q_c is constant, $\beta_1 \,\mathrm{Cov}(s_N, s_1)$ in the numerator is equal to $\beta_1 \,\mathrm{Cov}(\epsilon_N, q_i + \epsilon_1) = 0$. Thus,

$$\hat{\beta}_2 = \beta_2 \frac{\mathrm{Var}(s_N)}{\mathrm{Var}(s_N)} = \beta_2. \tag{30}$$

□

Proof of Proposition 4. Substituting the true model,

$$y = \beta_1 q + \gamma_G \, G + \gamma_R R + \gamma_{GR} \, (G \times R) + \beta_1 \epsilon_1 + \beta_2 \epsilon_N. \tag{31}$$

The term $\beta_1 q$ is completely taken out by q_{ML}, since we assume $q_{ML} = q$. Thus, the remaining terms are given by

$$\eta = \beta_2 \big(\beta_R \, R + \beta_{GR} \, (G \times R)\big) + \beta_1 \epsilon_1 + \beta_2 \epsilon_N. \tag{32}$$

By the omitted variable bias formula, the estimated coefficient $\hat{\beta_G}$ on G is given by:

$$\hat{\gamma}_G = \gamma_G + \frac{\mathrm{Cov}\left(G, \ \gamma_R \, R + \gamma_{GR} \, (G \times R)\right)}{\mathrm{Var}(G)}. \tag{33}$$

Since G is binary (so that $G^2 = G$), we have:

$$\mathrm{Cov}(G, \ G \times R) = \mathrm{Cov}(G, R). \tag{34}$$

Thus, the estimated coefficient becomes:

$$\hat{\gamma}_G = \gamma_G + (\gamma_R + \gamma_{GR}) \, \frac{\mathrm{Cov}(G, R)}{\mathrm{Var}(G)}. \tag{35}$$

The "true" coefficient is γ_G, so the bias is given by

$$\hat{\gamma}_G - \gamma_G = (\gamma_R + \gamma_{GR}) \frac{\mathrm{Cov}(G, R)}{\mathrm{Var}(G)}. \tag{36}$$

There is no restriction on the potential magnitude of $\gamma_R + \gamma_{GR}$, so $\hat{\gamma}_G$ could be over-estimated, under-estimated, or of a different sign than the desired γ_G. □

Proof of Proposition 5.

Because all four normal variables $(q_i, q_c, \epsilon_1, \epsilon_N)$ are independent and of zero mean, the system $(q, s_1, s_N)^T$ follows a multivariate normal distribution. Therefore, the conditional expectation of q on $(s_1, s_N)^T$ is given by

$$E[q|s_1, s_N] = \mathrm{Cov}[q, (s_1, s_N)]\mathrm{Var}[(s_1, s_N)^T]^{-1} \cdot (s_1, s_N)^T. \tag{37}$$

Again using the independence of $(q_i, q_c, \epsilon_1, \epsilon_N)$, we solve

$$\mathrm{Cov}[q, (s_1, s_N)] = (\sigma_i^2 + \sigma_c^2, \sigma_c^2). \tag{38}$$

Similarly,

$$\mathrm{Var}[(s_1, s_N)^T] = \begin{bmatrix} \sigma_i^2 + \sigma_c^2 + \sigma_{\epsilon_1}^2 & \sigma_c^2 \\ \sigma_c^2 & \sigma_c^2 + \sigma_{\epsilon_N}^2 \end{bmatrix}. \tag{39}$$

Taking the inverse of this and plugging in, we obtain after simplification,

$$\mathrm{Cov}[q, (s_1, s_N)] \cdot \mathrm{Var}[(s_1, s_N)^T]^{-1} = \left(\frac{\sigma_i^2 + \frac{\sigma_c^2 \sigma_{\epsilon_N}^2}{\sigma_c^2 + \sigma_{\epsilon_N}^2}}{\sigma_i^2 + \frac{\sigma_c^2 \sigma_{\epsilon_N}^2}{\sigma_c^2 + \sigma_{\epsilon_N}^2} + \sigma_{\epsilon_1}^2}, \frac{\sigma_c^2}{\sigma_c^2 + \sigma_{\epsilon_N}^2} \cdot \frac{\sigma_{\epsilon_1}^2}{\sigma_i^2 + \frac{\sigma_c^2 \sigma_{\epsilon_N}^2}{\sigma_c^2 + \sigma_{\epsilon_N}^2} + \sigma_{\epsilon_1}^2} \right).$$

For ease of notation, we write

$$\beta_1 = \frac{\sigma^2}{\sigma^2 + \sigma_{\epsilon_1}^2}, \quad \sigma^2 = \sigma_i^2 + \frac{\sigma_c^2 \sigma_{\epsilon_N}^2}{\sigma_c^2 + \sigma_{\epsilon_N}^2}$$

such that

$$E[q|s_1, s_N] = \left(\beta_1, \frac{\sigma_c^2}{\sigma_c^2 + \sigma_{\epsilon_N}^2} \cdot (1 - \beta_1) \right) \cdot (s_1, s_N)^T. \tag{40}$$

$\square$

Proof of Proposition 6.

Similarly to the proof of Proposition 5, the conditional expectation of q on $(s_1', s_N)^T$ is given by

$$E[q|s_1', s_N] = \text{Cov}[q, (s_1', s_N)]\text{Var}[(s_1', s_N)^T]^{-1} \cdot (s_1', s_N)^T. \tag{41}$$

The expressions for both matrices are simpler in this case:

$$\text{Cov}[q, (s_1', s_N)] = (\sigma_i^2, \sigma_c^2) \tag{42}$$

$$\text{Var}[(s_1', s_N)^T] = \begin{bmatrix} \sigma_i^2 + \sigma_{\epsilon_1}^2 & 0 \\ 0 & \sigma_c^2 + \sigma_{\epsilon_N}^2 \end{bmatrix}. \tag{43}$$

In the same manner as before, we solve

$$\text{Cov}[q, (s_1', s_N)] \cdot \text{Var}[(s_1', s_N)^T]^{-1} = \left(\frac{\sigma_i^2}{\sigma_i^2 + \sigma_{\epsilon_1}^2}, \frac{\sigma_c^2}{\sigma_c^2 + \sigma_{\epsilon_N}^2} \right), \tag{44}$$

such that

$$E[q|s_1', s_N] = \left(\frac{\sigma_i^2}{\sigma_i^2 + \sigma_{\epsilon_1}^2}, \frac{\sigma_c^2}{\sigma_c^2 + \sigma_{\epsilon_N}^2} \right) \cdot (s_1', s_N)^T. \tag{45}$$

Note that $\beta_2' = \beta_2 \cdot \frac{1}{1-\beta_1}$. Since $\beta_1 \in (0,1)$, $\beta_2' > \beta_2$. Meanwhile, $\beta_1 = \frac{\sigma^2}{\sigma^2 + \sigma_{\epsilon_1}^2}$ and $\beta_1' = \frac{\sigma^2 - x^2}{\sigma^2 + \sigma_{\epsilon_1}^2 - x^2}$, where $x^2 = \frac{\sigma_c^2 \sigma_{\epsilon_N}^2}{\sigma_c^2 + \sigma_{\epsilon_N}^2} < \sigma^2$, and thus $\beta_1' < \beta_1$. □

Proof of Proposition 7.

(i) Since $\sigma_{\epsilon_1}^2(q_i)$ is decreasing in q_i, the denominator of β_1' is decreasing in q_i, which makes the fraction β_1' an increasing function of q_i.

(ii) The expression for β_2' does not contain any function of q_i.

(iii) The ratio is $\frac{\beta_2'}{\beta_1'} = \left(\frac{\sigma_c^2}{\sigma_c^2 + \sigma_{\epsilon_N}^2} \right) \left(\frac{\sigma_i^2 + \sigma_{\epsilon_1}^2(q_i)}{\sigma_i^2} \right)$. The first term is a positive constant and the second term is a decreasing function of q_i. Thus, the ratio is decreasing. □