

Article

The Diversity Argument for Affirmative Action in Medical School: A Critique

Stephen Kershner

Department of Philosophy, Fenton Hall State University of New York at Fredonia, Fredonia, New York 14063, USA; stephen.kershner@fredonia.edu

Submitted: 9 May 2023, accepted: 21 January 2024, published: 29 April 2024

Abstract: In this article, I argue that medical school admissions should be limited to statistically relevant factors. I argue for it based on two other conclusions: a medical school should maximize quality-adjusted medical services per graduate within the overall optimum spending limit and if this is correct, then a medical school should, other things being equal, select medical students who are better than their competitors. I then explore the implications of this argument for whether a medical school admissions system should be holistic and whether it should consider demographic factors. I also consider and respond to a series of objections to the argument.

Keywords: affirmative action; diversity; sabermetrics; medical school; medical student; doctor; physician

How to cite: Kershner, S. The Diversity Argument for Affirmative Action in Medical School: A Critique. *Journal of Controversial Ideas* **2024**, 4(1), 1; doi:[10.35995/jci04010001](https://doi.org/10.35995/jci04010001).

©2024 Copyright by the authors. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY 4.0) license.



Part One: Theses

Affirmative action is a policy in which a party tries to increase the representation of women and minorities in an area – for example, education or employment – in which they are underrepresented or from which they have been historically excluded. In some cases, the affirmative actor uses preferential treatment or quotas. The US Supreme Court held that universities may not use quota-based or explicit point-based affirmative-action admissions.¹ The same is true regarding hiring or promotion. Still, we should not confuse law and morality.

Diversity is the state of having different types of things. Demographic diversity is the state of having different types of people. Consider different ethnicities, races, religions, sexes, socioeconomic classes, or people with different experiences or ideas.

¹ For quotas, see *Regents of the University of California v. Bakke*, 438 U.S. 265 (1978). For quotas and explicit points, see *Gratz v. Bollinger*, 539 U.S. 244 (2003) and *Grutter v. Bollinger*, 539 U.S. 306 (2003).

The diversity argument for affirmative action for undergraduates focuses on diversity of ideas. First, the courts and universities argue that demographic diversity brings about idea diversity. The Grutter Court favorably cited the Harvard program.² In an amicus brief in that case, Harvard said racial diversity promotes the “robust exchange of ideas which discovers truth out of a multitude of tongues, rather than through any kind of authoritative selection.”³ Second, the courts and universities argue that demographic diversity brings about cross-racial understanding. The Grutter Court said the diversity produces cross-racial understanding and breaks down racial stereotypes.⁴ Third, the courts and universities argue that demographic diversity improves classroom discussion. “[C]lassroom discussion is livelier, more spirited, and simply more enlightening and interesting’ when the students have ‘the greatest possible variety of backgrounds.’”⁵

Fourth, Justice Kennedy said that demographic diversity leads to “the cultivat[ion] [of] a set of leaders with legitimacy in the eyes of the citizenry.”⁶ It appears that he intended that this goal be distinct from the marketplace-of-ideas justification. Perhaps this is also true for Kennedy’s lessening-racial-stereotype justification.

As practiced by Harvard and the University of North Carolina – controversially – the National Association of Scholars claims affirmative action transfers positions from Asian-Americans to non-disadvantaged African-Americans. Studies at Harvard and other colleges have found that up to 80% of slots awarded to African-American and Hispanic students under preferential admissions come at the cost of positions that Asian-Americans would have gotten.⁷ Ron Unz argues that a significant number of the positions come from non-Jewish whites.⁸ Regarding the beneficiaries of the program, one study found that “Harvard is not employing racial preferences in an effort to benefit disadvantaged minority students. Harvard admits more than twice as many non-disadvantaged African-American applicants than disadvantaged African-American applicants.”⁹ Affirmative action also benefits Hispanics and Native Americans. In other contexts, it benefits Asian, gays, and women.

Critics have claimed that the above marketplace-of-ideas justification for affirmative action fails for a number of reasons. First, there is no study that shows that demographic diversity leads to a net increase in academic learning in the classroom. Given that more able students are replaced by less able students, the opposite is likely true.

Second, the broader cost–benefit analysis matters because it is unclear whether affirmative action improves or worsens the market of ideas at universities. In a number of cases, universities have tried to limit discussion of racial matters – especially concerning

² See, 506 U.S. at 335–39.

³ Brief of Harvard University et al. as Amici Curiae Supporting Respondents, *Grutter v. Bollinger*, 539 U.S. 306 (2003) (No. 02-241), 2003 U.S. S. Ct. Briefs LEXIS 189, at 12.

⁴ See 539 U.S. at 330 (quoting *Grutter v. Bollinger*, 137 F. Supp. 2d 821, 850 (E.D. Mich. 2001), rev’d in part and vacated in part, 288 F.3d 732 (6th Cir. 2002), aff’d 539 U.S. 306). [NAS].

⁵ See 539 U.S. at 330 (quoting 137 F. Supp. 2d at 849). [NAS].

⁶ See *Regents of the University of California v. Bakke*, 438 U.S. 265 (1978).

⁷ See [NAS] Data obtained from Harvard in discovery in this case showed that 70.5% of the “underrepresented minority students” it admits are “advantaged.” Expert Report of Richard D. Kahlenberg 28, ECF No. 416-1.

⁸ See Unz, R. (2012, November 28). The Myth of American Meritocracy: How Corrupt Are Ivy League Admissions? *The American Conservative*. “[In 2012], if we aggregate the reported enrollment figures, we discover that 4 percent of all college-age American Jews are currently enrolled in the Ivy League, compared to just 1 percent of Asians and about 0.1 percent of whites of Christian background.”

⁹ See Expert Report of Peter S. Arcidiacono (“Arcidiacono Report”) 8, ECF No. 415-8. [NAS].

affirmative action and black academic performance – by disciplining professors who discuss or research these matters. Notable examples are Joshua Cohen, Charles Negy, Bryan Pesta, and Amy Wax. The Pesta case is especially worthy of note because his university fired him, despite his being tenured, for publishing an article in a peer-reviewed journal. In addition, there have been lecture disruptions and, in some cases, violence in response to conservative intellectuals or professors who research race or race-adjacent issues, or who even comment on racial issues on campus. For example, protesters responded to Heather Mac Donald at McKenna-Claremont College, Charles Murray at Middlebury College, and Brett Weinstein at Evergreen State College by committing acts of intimidation and violence. It is unclear whether affirmative action contributes to the disciplining of professors, and disruption and violence surrounding speeches beyond what contemporary progressivism contributes to it.

Third, schools have not pursued diversity of ideas in a principled manner. For example, they have not sought to admit former felons, Mormons, Orthodox Jews, or polygamists despite their likely having very different ideas from the academic mainstream.

Fourth, the purported difference – emphasized in *Bakke*, *Gratz*, and *Grutter* – between diversity via quotas, points, and, in a holistic admissions system, goals is in practice no difference at all.

Fifth, what actually goes on in these programs is often naked race discrimination. Asian applicants to Harvard scored significantly higher than every other race in both academics and extracurricular rating and comparably on all observable metrics (that is, alumni interviews and teacher recommendations). Yet the staffers on campus who did not meet the applicants gave them lower personality scores. For example, Asian admissions would increase by 19% if the personality score were removed.

Setting aside the above objections, consider how the diversity justification of affirmative-action admissions in medical-school works. Diversity is a forward-looking justification. It differs from backward-looking justifications, such as compensatory (corrective) justice, desert, or ownership.¹⁰ On a side note, I am viewing the present as the limit to the past.

Consider the type of justification that diversity provides. Diversity is not an intrinsic rightness maker. This is because diversity is not a basic right-maker. It is dissimilar to the Categorical Imperative, rights, or utility.

Diversity is also not intrinsically good. It is dissimilar to purported welfarist intrinsic goods such as pleasure, desire-fulfillment, or objective-list goods. An objective-list good is something that by itself makes someone's life go better independent of whether it increases the person's pleasure or fulfills his or her desires.¹¹ Diversity is also dissimilar to purported non-welfarist intrinsic goods such as a person getting what he or she deserves.

¹⁰ For discussion of the compensatory-justice argument, see Thomson, J. J. (1973). *Preferential Hiring*. *Philosophy & Public Affairs*, 2, 364–384, and Kershnar, S. (2004). *Justice for the Past*. SUNY-Press. For the notion that desert can ground a claim to something, see Sher, G. (1989). *Desert*. Princeton University Press; Mulligan, T. (2018). *Justice and the Meritocratic State*. Routledge; Rachels, J. (1991). *What People Deserve*. In *Justice and Economic Distribution* (pp. 136–148). Prentice Hall; McLeod, O. (1999). *Desert and Institutions*. In *What Do We Deserve? A Reader on Justice and Desert* (pp. 186–195). Oxford University Press; McLeod, O. (1996). *Desert and Wages*. *Utilitas*, 8(2), 205–221. For discussion of an ownership-theory of affirmative action, see Boonin, D. (2011). *Should Race Matter?: Unusual Answers to the Usual Questions*. Cambridge University Press, Kershnar, *Justice for the Past*; Kershnar, S. (2021). *The Sabermetrics of Medical School Admissions*. *International Journal of Applied Philosophy*, 35(1): 45–63.

¹¹ See Parfit, D. (1984). *Reasons and Persons*. Oxford University Press.

If promoting diversity is extrinsically right, extrinsic goodness justifies its extrinsic rightness. That is, diversity's extrinsic rightness depends on extrinsic goodness.

If the diversity argument for affirmative action in medical school is forward-looking, then any good or bad effect it produces affects the program's extrinsic goodness. This is because there is no reason to favor one good result over another so long as both occur in the same amount. So, if affirmative action brings about more classroom learning, creates role models, eliminates harmful stereotypes, improves cultural literacy, and improves medical-group performance, each good result – assuming they are equally good – counts equally toward justifying the program. By the same token, medical-school affirmative action's costs – for example, more death, disability, and pain – counts against the policy.¹² Justice Lewis Powell wrote the plurality opinion in *Regents of the University of California v. Bakke*, 438 U.S. 265 (1978). Powell's purported diversity-of-ideas justification gets added to the sum of affirmative action's good and bad results. It is not an exclusionary reason in that it undermines or excludes other good or bad results.

In addition, the state uses affirmative action in order to produce doctors. The state produces doctors in order to produce medical services (specifically, diagnosis and treatment). As a result, what justifies medical-school affirmative action, if anything, is the efficient production of medical services. This means that the state should use efficient means to accomplish its medical-service goals. For example, if it is more efficient to pay doctors to work in underserved areas rather than to accept applicants from unrepresented groups and hope they work in the underrepresented areas, the state should do the former. The same might be true for choosing medical students who will work as physicians rather than working in other jobs, such as philosophy professor or stay-at-home mother.

With the above background in mind, here are my theses.

- (1) **Efficiency.** A state medical school should maximize quality-adjusted medical services per graduate within the overall optimum spending limit.
- (2) **Means to Efficiency.** A state medical school should, other things being equal, select medical students who are better than their competitors (in terms of a weighted combination of intelligence, knowledge, and probability of working more hours in the state in which the school is located).
- (3) **Admissions.** Medical school admissions should be limited to statistically relevant factors.

In part two, I argue for these theses. In part three, I provide a sabermetric equation – a mathematical equation that relies on statistical data – for admission to medical school. In part four, I discuss statistically validated predictors. In part five, I consider whether medical school admissions committees should value diversity.

In this article, I focus on state medical schools to sidestep the issue of who owns a private medical school. I also focus on affirmative action regarding medical school admissions rather than affirmative action regarding medical-school grading, hiring, or promotion. My conclusions generalize to private medical schools and to grading, hiring, and promotion within medical schools.

I use terms such as 'Asian,' 'black,' and 'white' to refer to people's races. For the purpose of this article, it does not matter whether these races are socially constructed or

¹² For example, consider the gross negligence leading to death, injury, and license suspension of one of the people University of California at Davis chose rather than Alan Bakke. See Martin, D. (2002, August 15). Patrick Chavis, 50, Affirmative Action Figure. *The New York Times*.

not. In most cases, I am referring to Americans of these races. I also refer to 'Hispanic' to refer to people having a shared ethnicity and 'Jew' to refer to people who are of the same race or who share a culture or religion. Nothing rests on what makes someone a Jew. I also use 'doctor' and 'physician' interchangeably.

Part Two: Arguments

My first thesis says the following: *A state medical school should maximize quality-adjusted medical services per graduate.* By quality-adjusted medical service, I mean effectiveness per unit of service. This thesis rests on two premises. The first premise says the following: *A state (professional) school should maximize (efficient) production.*

Two assumptions support the premise. First, *the most plausible moral theories (consequentialism, consent, and fairness) justify maximizing (efficient) production.* Consequentialism says the right maximizes the good. It does so by maximizing efficient production of intrinsic goods. Consent theory says that what is justified is that to which the relevant parties consent. Consider owners. The citizens of a state own the medical school and – I claim – they consent to pay for a state medical school on the basis that it maximizes health within the state.¹³ Let us set aside whether maximum health is an average or a total.¹⁴ Fairness theory says that a state medical school should bring about a fair distribution of health. The fair distribution of health – I claim – is the maximum amount of health or, perhaps, maximum health with greater weight given to those who are more deserving, have worse health, or are worse off in general.¹⁵ Aside from desert – I claim – these other bases for a fairness claim are implausible. Still, the point is that the state medical school should maximize a population's health.

The second assumption says: *If the most plausible moral theories justify maximizing (efficient) production, then a state professional school should maximize production.* These theories are the most plausible theories of what justifies a state professional school's policy. As a result, the state should maximize production.

The second premise says: *If a state professional school should maximize (efficient) production, then a state medical school should maximize quality-adjusted medical services per graduate (within the overall optimum spending limit).* Two assumptions justify this premise. First, *if the first premise is true, then a state medical school should maximize health.* Second, *a state medical school maximizes health by maximizing quality-adjusted medical services per graduate within the optimum spending limit.* Two equations underlie this assumption.

- (1) Total production = (amount of service) x (quality/service)
- (2) Production per graduate = (total production) / (number of graduates)

The second thesis relies on the following premise: *The best way to maximize quality-adjusted services per graduate is by selecting students who are better than their competitors (in terms of a weighted combination of intelligence, knowledge, and probability of working more hours in the state in which the school is located).* This is an empirical claim.

¹³ See Gomez, V. (2022, April 26). As Courts Weigh Affirmative Action, Grades and Test Scores Seen as Top Factors in College Admissions. *Pew Research Center*. [link to the article](#).

¹⁴ For discussion of this issue, see Kershner, S. (2016). Quantifying Health Across Populations. *Bioethics*, 30: 451–461.

¹⁵ The idea for this approach comes from Rawls, J. (1971). *A Theory of Justice*. Harvard University Press.

If the second thesis is true (*A state medical school should, other things being equal, select medical students who are better than their competitors*) and the state can most efficiently select such students through the use of statistically validated factors, then we arrive at the third thesis (*Medical school admissions should be limited to statistically validated factors*). The underlying assumption here is that statistically validated predictors are more accurate than other means of selection. Consider, for example, gestalt judgments. There might be an other-things-being-equal condition here because in some cases the collection of statistically validated predictors might not answer every admissions-related ranking question.

In this article, I use 'statistically validated predictor' to refer to a predictor that is supported by sufficient scientific evidence. As a result, we are epistemically justified in believing it to be accurate. Accuracy is the degree to which a measurement represents the true value of that which it measures.¹⁶ It focuses on the following question: How close is a measurement to the true value? A statistically validated predictor differs from a morally validated predictor. The latter is a predictor we have a sufficient moral reason to use.

Consider, by analogy, how professional sports teams use statistically validated factors to select players (and, also, to select plays). As I mentioned in part one, sabermetrics is the application of statistical analysis to sports to evaluate players (and plays).¹⁷ On one account, it is limited to baseball. As I will use it, though, sabermetrics uses statistical analysis to evaluate applicants for a position, whether in baseball, basketball, or a medical school. What justifies sabermetrics in decisions regarding players also justifies it in decisions regarding medical students.

By analogy, consider how a baseball team decides how to value a player and how much to pay him. Here "Benefit_{Player}," "Cost_{Player}," and "Value_{Player}," stand for a player's benefit, cost, and value. "WAR_{player}" means "wins above replacement per player." It stands for the number of wins a player brings about compared to an average replacement (for his position). The player's value is the difference between his benefit and cost. Benefit_{Player} is the product of the wins he brings about (WAR_{player}) and the cost per win across the league (Cost / WAR)_{general}. The idea is that in general (that is, across the league) it costs a certain amount per win and a player's benefit is equal to the number of wins he brings about multiplied by the cost per win. A player's cost is equal to his salary. Here, then, are the (crude) equations for player-value:

- (3) **Value.** $\text{Value}_{\text{Player}} = \text{Benefit}_{\text{Player}} - \text{Cost}_{\text{Player}}$
- (4) **Benefit.** $\text{Benefit}_{\text{Player}} = \text{WAR}_{\text{player}} \times (\text{Cost} / \text{WAR})_{\text{general}}$
- (5) **Cost.** $\text{Cost}_{\text{Player}} = \text{salary}_{\text{player}}$

Complex equations track how many wins a player brings about, specifically, his wins above replacement (WAR). The WAR equations differ for hitters and pitchers.¹⁸ There are competitor equations.

¹⁶ For the idea for this theory of accuracy, see Trajković, G. (2008). Measurement: Accuracy and Precision, Reliability and Validity. *Encyclopedia of Public Health*, 888–892. [link to the article](#).

¹⁷ See Lewis, M. (2004). *Moneyball: The Art of Winning an Unfair Game*. W. W. Norton & Company.

¹⁸ See Slowinski, P. (2012). WAR for Position Players. *Fan Graph*. Accessed 2 April 2023. [link to the article](#).

The above equations raise some controversies and have some defects.¹⁹ There is a question of whether the baseline should be the average replacement in a league or the player who would actually replace a player on a team. Also, a correct equation would include a player's non-playing contribution to a team's win. Consider, for example, a player's locker-room leadership. Consider, also, other ways in which a player causes his team to win (or lose). For example, New York Giants linebacker Lawrence Taylor paid prostitutes to keep opposing running backs up all night before they played the Giants, thereby degrading the running backs' performance.²⁰ Because this made the opposing team worse, Taylor's payments contributed to his player-value. In addition, some games (for example, playoff games) count more than others (for example, regular-season games). The equations also share some of the purported problems with the counterfactual comparative account of harm, specifically problems with omission, overdetermination, and preemption.²¹ Still, the above equations give us a helpful sketch regarding how to value baseball players. In theory, a similar equation tracks a person's value as a child, spouse, or philosopher.²²

Thus, sabermetrics *should* be used to select medical students. In part three I sketch how this should be done.

Part Three: Sabermetric Equation for Admission to Medical School

If statistically validated factors, and only statistically validated factors, should be used to select a medical student, then – currently – medical students should be selected using the following equation, which gives predicted student value (A, B, and C are weights; MCAT is 'medical college admissions test' and GPA is 'grade point average'):

$$\text{Value} = (A \times \text{MCAT}) + (B \times \text{GPA}) + (C \times \text{Demographic Factors})$$

Perhaps other factors predict performance. Here are two factors that might predict performance (D is a weight):

¹⁹ For discussion of the controversies and defects, see (Kershner, S. (2008). Solving the MVP Problem. *Journal of Social Philosophy*, 39: 141–159 and Kershner, S. 2011. The Most-Valuable-Player Problem Remains Unsolved. *Journal of Philosophy of Sport*, 38: 167–174.

²⁰ See Neuman, W. (2003, November 27). LT 'Sacked' Foes With Hookers. *New York Post*.

²¹ For a discussion of these problems in general, see Feit, N. (2023). *Bad Things: The Nature and Normative Role of Harm*. Oxford University Press. For a discussion of omission-based harm, see Feit, N. (2017). Harming by Failing to Benefit. *Ethical Theory and Moral Practice*, 1–15. For a discussion of the problem of overdetermination and harm, see Feit, N. (2015). Plural Harm. *Philosophy and Phenomenological Research*, 90(2), 361–388.

²² On a side note, why this is not done for philosophers is a mystery. Philosophy is very concerned with racism and sexism, but not with elitism. For example, of the last ten NFL MVPs, nine did not play for teams that in 2022 were ranked in the top 25. See, for example, Patrick Mahomes (2022, 2018, Texas Tech), Aaron Rodgers (2021, 2020, 2014, 2011, California-Berkeley), Lamar Jackson (2019, Louisville), Matt Ryan (2016, Boston College), Cam Newton (2015, Auburn), and Adrian Peterson (2012, Oklahoma). In contrast, Princeton's three assistant professors and three of four of Harvard's assistant professors went to top-ten-ranked philosophy programs. For philosophy department rankings, see Weinberg, J. (2022). Reputational Ranking of Philosophy PhD Programs Updated. *Daily Nous*. Accessed 7 January 2024. [link to the article](#).

- Structured interview score ($D \times$ Structured interview score)²³
- Personality features ($E \times$ personality score)

Examples of personality features include conscientiousness and grit.²⁴

To the extent that statistical validation is unclear, the school might have to make an educated guess. The same might be true of weights. Still, the inclusion of a factor and its weight in an equation should be explicit so that it can be systematically applied and evaluated afterward. In general, a *computer algorithm* – rather than an *admissions committee* – should decide whom to admit because it will more closely follow the relevant equation.

The above equation leaves out the cost per student. Consider, for example, who gets financial aid or one of a limited number of scholarships. Consider, also, a student who costs more because he or she needs costly psychological help or disability services. Still, the above equation is a helpful sketch of a correct equation.

This is similar to how we use studies to discover what medicines to use. This is also similar to how we discover when a medical test (for example, a mammogram) should be given. And this is how teams in Major League Baseball (MLB), the National Basketball Association (NBA), and the National Football League (NFL) select players and, also, discover how much they are worth.²⁵ Holistic judgments regarding medicines, tests, and players make sense only if holistic judgments are as accurate, precise, and reliable as statistically validated factors. If we could show this, though, then the holistic judgment would be statistically validated.

A school should reverse-engineer the weights from past medical-school classes. It should do so by determining which combination of factors and weight per factor best predict medical student success – preferably as a physician but, if necessary, as a student or resident – using past medical-school classes' performance as data.

In addition, the admission should be done via computer algorithm rather than via staffers' holistic judgments. The algorithm would more accurately apply the best equation. If staffers play a role, they should have to explicitly note where and why they chose to break away from the equation so that their judgments can be evaluated.

Medical schools should consider demographic factors when they accurately predict production. The American Medical Association (AMA) reports that in the US in the future, there will be too few physicians.²⁶ Given the AMA's report that female doctors work far fewer hours than men, female doctors worsen this problem. For example, the report states that in the UK, female physicians are six times more likely not to work full-time

²³ For the notion that a structured interview adds validity, see Levashina, J., Hartwell, C., Morgeson, F., & Campion, M. (2014). The Structured Employment Interview: Narrative and Quantitative Review of the Research Literature. *Personnel Psychology*, 67, 241–293. For the notion that structured interviews have less validity with regard to high-complexity jobs, see Huffcutt, A. I., Conway, J. M., Roth, P. L., & Klehe, U.-C. (2004). The Impact of Job Complexity and Study Design on Situational and Behavior Description Interview Validity. *International Journal of Selection and Assessment*, 12, 262–273. [link to the article](#). For the notion that interviews don't correlate with medical success, see Blackman, I., & Darmawan, I. G. (2004). Graduate-Entry Medical Student Variables that Predict Academic and Clinical Achievement. *International Education Journal*, 4, 30–41. It is unclear if this finding is limited to unstructured interviews.

²⁴ For the role of grit, see Duckworth, A. (2016). *Grit: The Power of Passion and Perseverance*. Scribner.

²⁵ See Mizels, J., Erickson, B., & Chalmers, P. (2022). Current State of Data and Analytics Research in Baseball. *Current Reviews in Musculoskeletal Medicine*, 15(4), 283–290. [link to the article](#).

²⁶ See Association of American Medical Colleges (2021). *The Complexities of Physician Supply and Demand: Projections From 2019 to 2034*. IHS Markit: Robeznieks, A. (2022). Doctor Shortages Are Here – and They'll Get Worse If We Don't Act Fast. Accessed 13 April 2023. [link to the article](#).

than male physicians (42% versus 7%), while in the US, 25% of female physicians do not work full-time.²⁷ An admissions system should consider demographic factors in order to accurately predict production.

Medical school admissions officers should not consider factors that have not been statistically validated. Currently, they do consider such factors, in particular the following: charitable work, essays,²⁸ interview (unstructured),²⁹ major,³⁰ physician shadowing,³¹ and recommendations.³² These factors are neither statistically validated nor given explicit weights. Medical schools should not consider them.

The third thesis that I give in part one (*Medical school admissions should be limited to statistically validated factors*) rules out holistic review of medical-school applicants. The Association of American Medical Colleges (AAMC) says that holistic review of medical school applicants is nearly universal and, also, morally required. Here is what the it says are the core principles regarding holistic admissions: “Schools consider each applicant’s potential contribution to both the school and the field of medicine, allowing them the flexibility to weigh and balance the range of criteria needed in a class to achieve their institutional mission and goals.”³³ Currently, holistic admissions give the qualitative factors so much weight that between 2019 and 2022, more than 15% of applicants with

²⁷ For the idea that having many female physicians is causing a problem in the UK, see Pemberton, M. (2013, March 24). Part-time Women Doctors Are Creating a Timebomb. *The Telegraph*. For female veterinarians are more likely to work part-time and work fewer hours, see Gates, R. (2014, September 30). Commentary: Why Women Are Leaving the Veterinary Profession. *dvm360*; Burns, K., & Larkin, M. (2013, May 20). The Gender Gap: Why Do Female Veterinarians Earn Less than Male Veterinarians. *Journal of the American Veterinary Medical Association*. On one account, medical error is the third most frequent cause of death. See Makary, M., & Daniel, M. (2016). Medical error – the Third Leading Cause of Death in the US. *The British Medical Journal*, 353; Anderson, J., & Abrahamson, K. (2017). Your Health Care May Kill You: Medical Errors. *Studies in Health Technology and Informatics*, 234, 13–17.

²⁸ One review of the literature found that more work was needed to see if these predicted success in medical school. See Ferguson, E., James, D., & Madeley, L. (2002). Factors Associated with Success in Medical School: Systematic Review of the Literature. *British Medical Journal*, 324(7343), 952–957. [link to the article](#). A later study found a correlation, but probably not with the things that an admissions officer would look for. See Pennebaker, J. W., Chung, C. K., Frazee, J., Lavergne, G. M., & Beaver, D. I. (2014). When Small Words Foretell Academic Success: The Case of College Admissions Essays. *PLoS ONE*, 9(12), e115844. [link to the article](#).

²⁹ See Dana, J., Dawes, R., & Peterson, N. (2013). Belief in the Unstructured Interview: The Persistence of an Illusion. *Judgment And Decision Making*, 8(5), 512–520. [link to the article](#). More generally, see Dana, Jason (2017, April 8). The Utter Uselessness of Job Interviews. *The New York Times*, [link to the article](#).

³⁰ There is no effect in osteopathic medicine. See Dixon, D. (2012). Prediction of Osteopathic Medical School Performance on the Basis of MCAT Score, GPA, Sex, Undergraduate Major, and Undergraduate Institution. *Journal of Osteopathic Medicine*, 112(4), 175–181. [link to the article](#). However, one study found that an undergraduate’s institution affected the predictive power of his or her MCAT score. See Zeleznik, C., Hojat, M., & Veloski, J. J. (1987). Predictive Validity of the MCAT as a function of undergraduate institution. *Journal of Medical Education*, 62(3), 163–169. [link to the article](#).

³¹ I cannot find any evidence that this predicts success in medical school or as a physician.

³² See DeZee, K. J., Magee, C. D., Rickards, G., Artino, A. R., Jr., Gilliland, W. R., Dong, T., McBee, E., Paolino, N., Cruess, D. F., & Durning, S. J. (2014). What Aspects of Letters of Recommendation Predict Performance in Medical School? Findings from One Institution. *Academic Medicine*, 89(10), 1408–1415. [link to the article](#). The study found, “Conclusions: LORs have limited value to admission committees, as very few LOR characteristics predict how students perform during medical school.” See, also, Ferguson et al., Factors Associated with Success in Medical School.

³³ See Association of American Medical Colleges. (2023). Holistic Review. Accessed 23 March 2023. [link to the article](#).

the highest levels of both GPA and MCAT scores were rejected by all medical schools to which they applied.³⁴

On the above sabermetric account, selection criteria include academic metrics, but not applicant experiences. Both sabermetric and holistic theories tell admissions officers to consider demographic attributes, but they differ in how to use them. The AAMC claims that medical schools should use holistic admissions and that considering race and ethnicity is part of holistic admissions. Considering these factors, the AAMC claims, is part of holistically considering how each applicant would contribute to the school learning environment and to the profession. If the above sabermetric theory of medical school is correct, admissions officers should not holistically consider how each applicant would contribute to the school learning environment and to the profession. They should also not vary weight per criterion in deciding whom to admit.

What does a sabermetric theory of medical school admissions tell us regarding what factors an admissions system should look at? What does the theory tell us about diversity? We now turn to these issues.

Part Four: Statistically Validated Predictors

Admission to medical school matters. The need to select high quality medical students can be seen in the combination of the following facts:

- (1) Medical error brings about a large amount of harm.
- (2) Doctors are rarely held accountable.
- (3) Nearly every medical student graduates.

The below statistics indicate the amount of harm and graduation rate:

- (1) **Percentage.** Medical error brings about 10% of all US deaths.³⁵
- (2) **Rank.** Medical error is the third leading cause of death.
- (3) **Number.** Medical error brings about more than 250,000 deaths per year.
- (4) **Accountability.** The state and hospitals rarely hold physicians accountable for misconduct.
- (5) **Graduation.** More than 19 out of 20 medical students graduate.³⁶
- (6) **Shortage.** According to the AAMC, the US faces a shortage of between 37,800 and 124,000 physicians by 2034. This includes shortfalls in (a) primary care (17,800–48,000) and (b) specialty care (21,000–77,100).³⁷

Consider accountability.³⁸ Doctors disciplined or banned by hospitals often keep clean licenses: From 2001 to 2011, nearly 6,000 doctors had their clinical privileges restricted

³⁴ See Association of American Medical Colleges. (2022). Brief for Amici Curiae Association of American Medical Colleges et al in Support of Respondents. *Amici Curiae Students for Fair Admission, Inc. v. President and Fellows of Harvard College, Students for Fair Admissions, Inc. v. University of Carolina, et al.* Accessed 8 July 2023: [link to the article](#).

³⁵ The first three facts come from applications of Makary & Daniel, Medical Error – the Third Leading Cause of Death in the US: 353.

³⁶ Because 97% of admitted medical student admitted become physicians, the admissions committee is deciding who becomes a physician. See Association of American Medical Colleges Graduation Rates and Attrition Factors for U.S. Medical School Students.

³⁷ See Association of American Medical Colleges, *The Complexities of Physician Supply and Demand: Projections From 2019 to 2034*.

³⁸ The following facts come from Eisler, P., & Hansen, B. (2013, August 20). Thousands Of Doctors Practicing Despite Errors, Misconduct. *USA TODAY*.

or taken away by hospitals and other medical institutions for misconduct involving patient care. But 52% – more than 3,000 doctors – never were fined or had their licenses restricted, suspended, or revoked by a state medical board. State medical board discipline differs from that given by hospitals or medical institutions.

Even the most severe misconduct goes unpunished. Nearly 250 of the doctors sanctioned by health care institutions were cited as an “immediate threat to health and safety,” yet their licenses still were not restricted or taken away. About 900 were cited for incompetence, malpractice, negligence, or substandard care, and kept practicing with no licensure action.

Doctors with the worst malpractice records keep treating patients. Among the nearly 100,000 doctors who made payments to resolve malpractice claims from 2001 to 2011, roughly 800 were responsible for 10% of all the dollars paid and their total payouts averaged about \$5.2 million per doctor. Yet fewer than one in five faced any sort of licensure action by their state medical boards.

Next consider what predicts successful grades, board scores, and clinical grades. First, consider studies regarding MCATs:

- (1) **Predictor.** MCATs are medium to large predictors of medical student performance.³⁹
- (2) **Grades.** MCATs outpredict undergraduate grades (uGPA).⁴⁰
- (3) **MCAT and Grades.** MCATs and grades together predict grades better than MCAT scores alone.
- (4) **Fair.** The MCAT is fair to underrepresented minorities. It does not underestimate their grades.⁴¹

Next consider other predictors:

- (1) **Board Scores.** MCATs predict board scores.⁴²

³⁹ See Hanson, J., Busche, K., Elks, M., Jackson-Williams, L., Liotta, R., Miller, C., Morris, C., Thiessen, B., & Yuan, K. (2022). The Validity of MCAT Scores in Predicting Students' Performance and Progress in Medical School: Results From a Multisite Study. *Academic Medicine*, 97(9), 1374–1384. [link to the article](#); Murphy, B. (2022). MCAT Scores and Medical School Success: Do They Correlate? *American Medical Association*. The MCAT is a reasonably strong predictor of medical performance. It predicts better than undergraduate GPA, although a combination of both is better at predicting grades in medical school. See Julian, E. R. (2005). Validity of the Medical College Admission Test for Predicting Medical School Performance. *Academic Medicine*, 80(10), 910–917. [link to the article](#). For a similar result with regard to osteopathic medicine, see Dixon, D. (2012). Prediction of Osteopathic Medical School Performance on the Basis of MCAT Score, GPA, Sex, Undergraduate Major, and Undergraduate Institution. *Journal of Osteopathic Medicine*, 112(4), 175–181. [link to the article](#). Dixon found that MCAT and uGPA were correlated with success in osteopathic medicine with the former (at least part of it) having the highest predictive value. For an argument that uGPA is a better predictor of medical school grades, see Salvatori, P. (2001). Reliability and Validity of Admissions Tools Used to Select Students for the Health Professions. *Advances in Health Sciences Education: Theory and Practice* 6(2), 159–175. [link to the article](#).

⁴⁰ See Hanson, J., Busche, K., Elks, M., Jackson-Williams, L., Liotta, R., Miller, C., Morris, C., Thiessen, B., & Yuan, K. (2022). The Validity of MCAT Scores in Predicting Students' Performance and Progress in Medical School: Results From a Multisite Study. *Academic Medicine*, 97(9), 1374–1384; Murphy, MCAT Scores and Medical School Success.

⁴¹ See Murphy, MCAT Scores and Medical School Success.

⁴² See, for example, Saguil, A., Dong, T., Gingerich, R. J., Swygert, K., LaRochelle, J. S., Artino, A. R., Jr., Cruess, D. F., & Durning, S. J. (2015). Does the MCAT Predict Medical School and PGY-1 Performance? *Military Medicine*, 180(4 Suppl), 4–11. [link to the article](#). On some accounts, the MCAT outpredicts undergraduate GPA on STEP scores. See Julian, Validity of the Medical College Admission Test for Predicting Medical School Performance. For the claim that MCAT scores significantly associate with Step 1 and 2 scores, see Gauer, J. L., Wolff, J. M., & Jackson, J. B. (2016). Do MCAT Scores Predict

- (2) **Clinical Grades.** Board scores predict clinical grades. This is true even when researchers control for the MCAT.
- (3) **Graduation.** The vast majority of medical school students graduate: (a) 98% graduation rate for students with respectable MCATs (510–513); (b) 94% graduation rate for students with low MCATs (498–501).⁴³
- (4) **Job Performance.** Do clinical grades predict performance as a physician? *I cannot find a direct study on this, but there is indirect evidence for it.*⁴⁴

USMLE Scores? An Analysis On 5 Years Of Medical Student Data. *Medical Education Online*, 21, 31795. [link to the article](#). See, also, Ogunyemi, D., & Taylor-Harris, D. (2005). Factors that Correlate with the U.S. Medical Licensure Examination Step-2 Scores in a Diverse Medical Student Population. *Journal of the National Medical Association*, 97(9), 1258–1262; Callahan, C. A., Hojat, M., Veloski, J., Erdmann, J. B., & Gonnella, J. S. (2010). The Predictive Validity of Three Versions of the MCAT in Relation to Performance in Medical School, Residency, and Licensing Examinations: A Longitudinal Study of 36 Classes of Jefferson Medical College. *Academic Medicine*, 85(6), 980–987. [link to the article](#); Donnon, T., Paolucci, E. O., & Violato, C. (2007). The Predictive Validity of the MCAT for Medical School Performance and Medical Board Licensing Examinations: A Meta-Analysis of the Published Research. *Academic Medicine*, 82(1), 100–106. [link to the article](#); Swanson, D. B., Case, S. M., Koenig, J., & Killian, C. D. (1996). Preliminary Study of the Accuracies of the Old and New Medical College Admission Tests for Predicting Performance on USMLE Step 1. *Academic Medicine*, 71(1 Suppl), S25–27. [link to the article](#); Basco, W. T., Jr., Way, D. P., Gilbert, G. E., & Hudson, A. (2002). Undergraduate Institutional MCAT scores as predictors of USMLE step 1 performance. *Academic Medicine*, 77(10 Suppl), S13–16. [link to the article](#). There is a debate as to whether standardized test scores outpredict grades. Some researchers assert that they do. See Kuncel, N. R., & Hezlett, S. A. (2007). Standardized Tests Predict Graduate Students' Success. *Science*, 315(5815), 1080–1081. [link to the article](#). Other researchers deny this. See Borghans, Lex, Golsteyn, Bart, H. H., Heckman, James J., & Humphries, John Eric. "What Grades and Achievement Measure," IZA DP No. 10356, November 2016.

⁴³ The five-year graduate rate is 94%. The eight-year graduation rate is 97%. See Association of American Medical Colleges, "Graduation Rates and Attrition Factors for U.S. Medical School Students," *Analysis in Brief* 14:5 (2014): 1–2. [link to the article](#). Compare this to the military academies. The graduation rate for the Naval Academy is 86% and for West Point it is 83%. See Prudente, T. (2014, December 1). Superintendent: Fewer Dropouts Seen at Naval Academy. *Capital Gazette*. The average West Point student is smarter than the average medical student. The former has an average SAT of 1285 and thus an average IQ of 129. A dated study found that the average physician has an IQ of 125. See Perry, W., & Crean, R. D. (2005). A Retrospective Review of the Neuropsychological Test Performance of Physicians Referred for Medical Infractions. *Archives of Clinical Neuropsychology*, 20(2), 161–170. [link to the article](#).

⁴⁴ The studies are mixed on whether medical school performance correlates with performance in residency. See Yindra, K. J., Rosenfeld, P. S., & Donnelly, M. B. (1988). Medical school Achievements As Predictors Of Residency Performance. *Journal of Medical Education*, 63(5), 356–363. [link to the article](#); Daly, K. A., Levine, S. C., & Adams, G. L. (2006). Predictors for Resident Success in Otolaryngology. *Journal of the American College of Surgeons*, 202(4), 649–654. [link to the article](#); Stohl, H. E., Hueppchen, N. A., & Bienstock, J. L. (2010). Can Medical School Performance Predict Residency Performance? Resident Selection and Predictors of Successful Performance in Obstetrics and Gynecology. *Journal of Graduate Medical Education*, 2(3), 322–326. <https://doi.org/10.4300/jgme-d-09-00101.1>; Hayden, S. R., Hayden, M., & Gamst, A. (2005). What Characteristics of Applicants to Emergency Medicine Residency Programs Predict Future Success as an Emergency Medicine Resident? *Academic Emergency Medicine*, 12(3), 206–210. [link to the article](#); Metro, D., Talarico, J., Patel, R., & Wetmore, A. (2005). The Resident Application Process and Its Correlation to Future Performance as a Resident. *Anesthesia and Analgesia*, 100, 502–505. [link to the article](#); Dirschl, D., Campion, E., & Gilliam, K. (2006). Resident Selection and Predictors of Performance: Can We Be Evidence Based? *Clinical Orthopaedics and Related Research*, 449, 44–49. [link to the article](#); Borowitz, S. M., Saulsbury, F. T., & Wilson, W. G. (2000). Information Collected during the Residency Match Process Does Not Predict Clinical Performance. *Archives of Pediatrics & Adolescent Medicine*, 154(3), 256–260. [link to the article](#); Bell, J. G., Kanellitsas, I., & Shaffer, L. (2002). Selection of Obstetrics and Gynecology Residents on the Basis of Medical School Performance. *American Journal of Obstetrics and Gynecology*, 186(5), 1091–1094. [link to the article](#). It is also not clear whether residency selection criteria correlate with residency performance. See Dirschl, D. R., Dahners, L. E., Adams, G. L., Crouch, J. H., & Wilson, F. C. (2002). Correlating Selection Criteria with

- (5) **IQ.** IQ is probably the strongest predictor of job performance.^{45,46} Consider, also, conscientiousness. On one estimate, for medium complexity jobs, the two explain 36% of performance.
- (6) **SAT/LSAT.** SAT & LSAT are strongly g-loaded.⁴⁷ An IQ score attempts to measure general intelligence (g).

Some recent studies show MCATs are the best predictors of medical student outcomes.⁴⁸

Also, intelligence is likely relevant. Smarter or more knowledgeable physicians will on average produce more high-quality services than others. IQ is one of the best predictors of economic success and income.⁴⁹ In general, it has a high degree of correlation with job performance at high, medium, and low complexity.⁵⁰ On at least one interpretation, IQ

Subsequent Performance as Residents. *Clinical Orthopaedics and Related Research*, (399), 265–271. [link to the article](#); Olawaiye, A., Yeh, J., & Withiam-Leitch, M. (2006). Resident selection Process And Prediction Of Clinical Performance in an Obstetrics and Gynecology Program. *Teach Learn Med*, 18(4), 310–315. [link to the article](#); and Naylor, R. A., Reisch, J. S., & Valentine, R. J. (2008). Factors Related to Attrition in Surgery Residency Based on Application Data. *Archives of Surgery*, 143(7), 647–651. [link to the article](#).

⁴⁵ IQ is one of the best predictors of economic success and income. See Murray, C. (1998). *Income Inequality and IQ*. American Enterprise Institute; Kuncel, N., & Hezlett, Sarah. (2010). Fact and Fiction in Cognitive Ability Testing for Admissions and Hiring Decisions. *Current Directions in Psychological Science*, 19, 339–345; Kuncel, N. R., Hezlett, S. A., & Ones, D. S. (2001). A Comprehensive Meta-analysis of the Predictive Validity of the Graduate Record Examinations: Implications for Graduate Student Selection and Performance. *Psychological Bulletin*, 127(1), 162–181. [link to the article](#). For an interpretation indicating that medical school success is strongly predicted by IQ, see Hughs, P. (2002). Can We Improve on How We Select Medical Students? *Journal of the Royal Society of Medicine*, 95(1), 18–22.

⁴⁶ If one wants people who will succeed in a variety of positions and at both analytical and creative tasks, testing should focus on standardized tests, IQ, and intelligence. Consider, for example, the incredible accomplishments of mathematically precocious 13-year-olds. See Lubinski, D., Benbow, C. P., & Kell, H. J. (2014). Life Paths and Accomplishments of Mathematically Precocious Males and Females Four Decades Later. *Psychological Science*, 25(12), 2217–2232. [link to the article](#). Emotional intelligence appears to be independent of cognitive abilities that are traditionally part of medical school admissions. See Carr, S. E. (2009). Emotional Intelligence in Medical Students: Does It Correlate with Selection Measures? *Medical Education*, 43(11), 1069–1077. [link to the article](#); Leddy, J. J., Moineau, G., Puddester, D., Wood, T. J., & Humphrey-Murto, S. (2011). Does an Emotional Intelligence Test Correlate With Traditional Measures Used to Determine Medical School Admission? *Academic Medicine*, 86(10). [link to the article](#) J. There is some disagreement about whether emotional intelligence correlates with academic performance. See Chew, B. H., Zain, A. M., & Hassan, F. (2013). Emotional Intelligence and Academic Performance in First and Final Year Medical Students: A Cross-sectional Study. *BMC Medical Education*, 13, 44. [link to the article](#); Austin, E. J., Evans, P., Magnus, B., & O'Hanlon, K. (2007). A preliminary study of empathy, emotional intelligence and examination performance in MBChB students. *Medical Education*, 41(7), 684–689. [link to the article](#). Also, there are some problems with using tests to measure emotional intelligence. First, two of the different tests do not correlate with one another. Second, the trait and ability measures appear not to be tightly connected. See Brannick, M. T., Wahi, M. M., Arce, M., Johnson, H. A., Nazian, S., & Goldin, S. B. (2009). Comparison of Trait and Ability Measures of Emotional Intelligence in Medical Students. *Medical Education*, 43(11), 1062–1068. [link to the article](#).

⁴⁷ My guess is that MCATs relate to IQ similar to how other standardized tests do so. Note, however, that I cannot find a study that finds that the two correlate.

⁴⁸ See Murphy, MCAT Scores and Medical School Success.

⁴⁹ See Murray, *Income Inequality and IQ*. For the notion that IQ is a good predictor of grades, performance at work, and many other aspects of success in life," see Nisbett, R. E., Aronson, J., Blair, C., Dickens, W., Flynn, J., Halpern, D. F., & Turkheimer, E. (2012). Intelligence: New Findings and Theoretical Developments. *American Psychologist*, 67(2), 130–159. [link to the article](#).

⁵⁰ See Kuncel & Hezlett, Fact and Fiction in Cognitive Ability Testing for Admissions and Hiring Decisions; Kuncel et al., A Comprehensive Meta-analysis of the Predictive Validity of the Graduate Record

strongly predicts success in medical school and as a physician.⁵¹ To the extent that other personality factors predict quantity or quality of medical services, this premise should be modified to include these features. Consider, for example, grit, conscientiousness, or emotional intelligence.⁵²

Part Five: Diversity and Group Performance

The argument from group performance in favor of affirmative action consists of two premises. First, *physician-diversity maximizes the good*. Second, *if physician-diversity maximizes the good, then medical schools should value an applicant's contribution to diversity when deciding whom to admit*.

The first premise (*Physician-diversity maximizes the good*) rests on several assumptions. First, *if, other things being equal, physician-diversity causes people to be healthier and there are no overriding effects, then physician-diversity maximizes the good*. Second, *other things being equal, physician-diversity causes people to be healthier*. Consider, for example, the evidence – discussed below – that diverse groups function better in medicine. Consider, also, the evidence – again discussed below – that diverse groups function better in areas outside of medicine, such as finance. Third, *there are no overriding effects*. An overriding effect might be present, specifically when physicians are not part of a team. Perhaps, though, they are still part of a medical team that consists of the wider medical community because it shapes how, when, and where a patient gets medical care.

The maximization of the good might not just be through better performing teams. It might also come about because underrepresented minority doctors function as role models, increase cultural competence, decrease harmful stereotypes, or provide more services to disadvantaged communities. In the aggregate these effects might also cause underrepresented minorities to invest efficiently – by investing more – in university education. Investing in a university education competes with investing in learning a trade.

One concern about the benefits listed in the above paragraph is that they are hard to quantify. A second concern is that they bring about costs that outweigh them. Among possible costs are that doctors admitted through affirmative action function as negative role models, decrease cultural competence, increase harmful stereotypes, or provide lower quality service to disadvantaged communities. In the aggregate these effects might cause underrepresented minorities to overinvest in university education. Still – I claim – we do not know which of these benefits or costs occur, or whether their sum is good or bad.

The second premise (*If physician-diversity maximizes the good, then medical schools should value an applicant's contribution to diversity when deciding whom to admit*) rests on two assumptions. First, *if physician-diversity maximizes the good, then medical schools should bring about an efficient amount of physician-diversity*. The idea here is that state medical schools are interested in maximizing the good – or, perhaps, the prudential good for the state's citizens – and that this is done by maximizing the citizens' health. Alternatively, the focus might be on residents rather than citizens. Second,

Examinations. For a comparison of IQ and grit, see Reznick, B. (2016). Why IQ Matters more than Grit: How to Think about Intelligence – a Conversation with an IQ Researcher. Vox.

⁵¹ See Hughs, Can We Improve on How We Select Medical Students?

⁵² For grit, see Duckworth, *Grit: The Power of Passion and Perseverance*.

if medical schools should bring about an efficient amount of physician-diversity, then medical schools should value diversity-contribution in admissions.

One response to the above argument – and the most important one – is that if diversity maximizes the good, then a medical school should bring it about via statistically validated predictors. That is, sabermetrics should tell us, on average, how much a medical-school applicant will contribute to medical-team diversity. That is, the above argument is consistent with my theses. The above argument does not support holistic admissions because we have no reason to believe that a holistic admissions system will outperform a sabermetric system.

A second response assumes that sufficient physician-diversity maximizes the good. Still, we need to know whether admission based on individual performance alone achieves sufficient diversity. That is, we need to know whether admission based on purely academic measures brings about too much, too little, or the right amount of diversity. Consider, for example, Chinese, Indian, Japanese, Korean, Nigerian, and Pakistani students, female students, and students from lower socioeconomic status who would have been accepted via an admissions equation that focuses solely on academic ability. If we do not know whether individual-performance-based admission achieves sufficient diversity, then medical schools should not value physician-diversity-contribution in admissions. This is because we do not know if we are trading off too much, too little, or the right amount of individual performance for diversity.

To see this, consider the diagram below. The chart illustrates a possible world in which a medical school trades off individual performance for diversity. The x-axis tells us the average amount of diversity-contribution per medical school student. The y-axis tells us the average amount of individual performance per medical school student. The curved lines – indifference curves in the graph – reflect efficient tradeoffs between the two factors. The straight line is the budget line. It tells us the efficient cost, that is, what the state medical school should spend on students.

The efficient amount of individual performance and diversity-contribution is where the most efficient tradeoff curve that intersects with the efficient budget does so.⁵³ This tells us the efficient amount of (average) individual performance and diversity-contribution. This point is A. It consists of the Y_0 amount of individual performance and X_0 amount of diversity-contribution.

The problem is that without knowing whether a class admitted purely on the basis of academic ability is to the right or left of X_0 , we do not know whether an admissions committee should pursue more diversity-contribution.

The individual-performance data is as follows.⁵⁴ Asians and white medical-school matriculants score significantly higher than blacks and Hispanics with regard to that which predicts grades in medical school. Consider MCAT scores and undergraduate GPA. Asians and whites perform better in medical school. Specifically, Asians and white medical students do significantly better than black and Hispanic medical students on the following: attrition, medical school grades, board scores (general), board scores (specific), and clinical grades.

⁵³ The diagram comes from Marlymimran. Consumer Equilibrium: Indifference Curves & Budget Line. *Quizlet*. Accessed 23 October 2023. [link to the article](#).

⁵⁴ See Association of American Medical Colleges, 2022 Facts: Applicants and Matriculants Data. [link to the article](#).

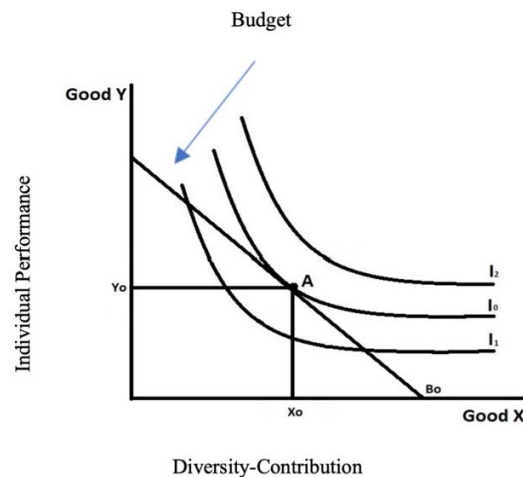


Figure 1: Efficient Diversity-Contribution.

The predictors' ability to predict physician-performance is less clear. If factors such as attrition, medical school grades, board scores (general), board scores (specific), and clinical grades predict physician-performance – and I cannot find a *recent* study addressing this – then Asian and white physicians likely outperform black and Hispanic physicians. If IQ is the best predictor of job performance – and a number of specialists assert that it is – then Asian and white physicians likely outperform black and Hispanic physicians. Based on a California study, it is likely that state medical boards discipline black and Hispanic physicians more than they discipline Asian and white physicians.⁵⁵

In addition, if the various tests – for example, classroom grades, board scores, and clinical grades – do not predict who is going to be a better doctor, then it is unclear why they should be used to determine who goes into what specialty or who gets what residency. That is, if they do not predict physician-performance, then relying on them is a mistake. And if an equation better predicts physician-performance than holistic criteria – and holistic criteria subtract rather than add to prediction-accuracy – then the equation alone should be used. For example, an equation for a specialty or residency – different from the above one because aimed more at physician-performance than medical-school performance – might involve one or more of the following: IQ score, personality score (for example, testing for conscientiousness or grit), and, also, MCATs, medical school grades and board scores.

A third response is that evidence that physician-diversity causes greater health is weak because it is very limited. For example, the AMA's brief cites two types of studies:

Study #1: Pain

Black physicians in a lab are more likely than others to accurately assess Black patients' pain tolerance and prescribe the correct amount of pain medication as a result.

Study #2: Babies

For high-risk Black newborns, having a Black physician is tantamount to a miracle drug: it more than doubles the likelihood that the baby will live.

⁵⁵ See Rogers, Patrick. (2017, January). Demographics of Disciplinary Action by the Medical Board of California (2003–2013). California Research Bureau.

The AMA cites only these two effects. One of the main studies supporting the first claim – response to Black patients’ pain – is merely a lab study. As a result, we might question the extent to which the pattern occurs in the actual practice of medicine. This suggests that we lack strong evidence that greater physician-diversity makes a population healthier. The second study is striking. Without more, though, it is weak evidence for the claim that physician-diversity increases health.

The studies in other areas – for example, finance and law review editorship – are stronger evidence. Still, in the absence of studies that involve diversity being traded off for significantly reduced individual performance, they are not strong evidence for how medical schools should select students. Some of the studies showing benefits from diversity include Asians. This suggests that merit-contribution is what is driving some of the diversity-contribution.⁵⁶ In addition, this evidence is mixed. Some of the meta-studies show that more diverse teams do not perform better.

A fourth response – and one that is highly controversial – says that markets tell us that demographic diversity does not significantly improve team performance. Consider markets involving friends, marriage, movies, and professional sports. The market-argument tells us that physician-diversity (probably) does not maximize group performance, at least when there is significant cost in terms of individual performance.

Consider, for example, professional sports. Specifically, consider the 2022 champions in three major sports. The Colorado Avalanche won the 2022 NHL championship. Its starting lineup was entirely white. The Golden State Warriors won the 2022 NBA championship. Its starting lineup was entirely black. In the 2023 Super Bowl, the Kansas City Chiefs defeated the Philadelphia Eagles. The Chiefs’ defensive starting lineup consisted of ten black players, one white player, and no Asians. The Eagles’ starting lineup consisted of eleven black players, no white players, and no Asians. I focus on defense-lineups because defense-lineups function as a team within a team. The NHL and NBA winners did not start a single Asian or Jewish player. Nor did Super Bowl defenses. Apparently, neither the owners nor the coaches think that greater diversity would improve the success of professional sports teams.

In fact, the number of Jews in professional sports is strikingly low. Only 1 of 450 players in the NBA is Jewish.⁵⁷ Only 5 of 1,690 players in the NFL is Jewish.⁵⁸ And only 1 of 578 UFC fighters is Jewish.⁵⁹ The percentages are 0.2%, 0.3%, and 0.2% respectively for a group that constitutes 2.4% of the American population.⁶⁰

Next, consider social markets. For example, 75% of whites have only white friends. Only 10.2% of marriages are interracial.⁶¹ These markets have been around for millennia.

⁵⁶ See Xu, K. (2021). *An Inconvenient Minority: The Attack on Asian American Excellence and the Fight for Meritocracy*. Diversion Books.

⁵⁷ See Zilgitt, J. (2022, November 4). Wizards’ Deni Avdija, NBA’s Lone Jewish Player, Addresses Kyrie Irving Controversy. *USA Today*.

⁵⁸ See Gurvis, J. (2022, September 9). Jewish Players to Watch as NFL’s 2022-2023 Season Kicks Off. *The Times of Israel*.

⁵⁹ See Spiro, A. (2022, December 16). UFC’s only Israeli fighter believes all Jews ‘should know how to defend themselves’. *The Times of Israel*.

⁶⁰ 2.4% of the US is Jewish. See Center, P. R. (2021). The Size of the U.S. Jewish Population. *Pew Research Center*. Accessed 11 May 2023. [link to the article](#).

⁶¹ The percentage of married-couple households that are interracial or interethnic grew across the United States from 7.4 to 10.2 percent from 2000 to 2012-2016. See Rico, B., Kreider, R., & Anderson, L. (2018). Growth in Interracial and Interethnic Married-Couple Households. *United States Census Bureau*. Accessed 9 July 2023. [link to the article](#). See, also, Swanson, E. (2014, September 3). Do Most White

If diverse marriages, friends, and neighborhoods brought about efficiency gains, one would think that these gains would be recognized and taken advantage of. This is similar to how, in theory, free markets tend to eliminate discrimination because firms do not last if they exclude customers or pay more for labor than their competitors.⁶² The fact that this is not true for social markets is a reason to be wary of the claim that demographic diversity increases team efficiency.

There is also a theoretical reason to think that social teams should disvalue diversity. If the preferences of spouses, friends, or community members are closer, then, on average, members' preferences will be frustrated less by agreed-upon solutions to conflicting preferences. In addition, if social team members' preferences are closer, then – on average – their transaction costs are lower. Specifically, they have to spend less on arriving at and enforcing an agreement.⁶³ The greater altruism tracking shared genes – whether in a family, clan, or race – further adds to the theoretical cost of diversity. Greater altruism lowers the cost to a member whose preference is frustrated.

Note this theoretical reason is in tension with the claim that markets tend to eliminate discrimination. Strictly speaking, though, they tend to eliminate inefficient discrimination and this sort of discrimination is efficient. It is efficient because it maximizes preference satisfaction.

Nor does Hollywood value diversity as much as we would expect, especially given that the people who work in the industry are very progressive.⁶⁴ One might think – although this is unclear – that the underrepresentation of women and minorities reflects a difference in the production of profits.⁶⁵ A similar pattern is true of actors.⁶⁶ This is true regarding the lack of actor-diversity in award-winning roles and in top money-making roles.

The sports and Hollywood markets have a couple of advantages in discovering the contribution-value of (demographic) diversity. First, they allow for clearcut measurement of production, for example, wins or profits. In sports, this is true for both individuals and teams. Second, most workers in both fields are progressive, so we would not expect that explicit racism or sexism to play a significant role.

Whether medical teams are similar to sports teams, Hollywood teams, or social teams is unclear. Some of the meta-studies show a diversity-based improvement in performance. Others show no such improvement. In addition, there might be informational benefits of diverse teams that outweigh the theoretical costs. Perhaps, on average, demographically diverse teams constitute a more robust micro-market of ideas than do demographically homogenous teams. By analogy, consider the information advantage of capitalism over socialism.

Americans Really Only Have White Friends? Let's Take a Closer Look. *Huffington Post*. Only 0.4% of whites have a black spouse; see Roland G. Fryer, Jr. (2007). Guess Who's Been Coming to Dinner? Trends in Interracial Marriage over the 20th Century. *Journal of Economic Perspectives*, 21(2): 71–90. It is worth noting that the proportion of new marriages that are interracial has increased from 3% in 1967 to 19% in 2019. See Parker, K., & Barroso, A. (2021). In Vice President Kamala Harris, We Can See How America Has Changed. *Pew Research Center*. Accessed 25 February 2024. [link to the article](#).

⁶² More specifically, markets tend to get rid of inefficient discrimination. See Becker, G. (1971). *The Economics of Discrimination*. University of Chicago Press.

⁶³ See Epstein, R. (1995). *Forbidden Grounds: The Case against Employment Discrimination Laws*. Harvard University Press.

⁶⁴ See Marks, A. (2023, January 2). Hollywood Actually Got Less Diverse Last Year. *Rolling Stone*.

⁶⁵ Ibid.

⁶⁶ For this and the next sentence, see Snow, S. Data Shows the Roots of Hollywood's Diversity Problem. *Shane Snow*. [link to the article](#).

If – and this is a big if – professional sports, social markets, and Hollywood are reasonably efficient, then it is odd that they would ignore an obvious contributor to team success: demographic diversity. This assumption is controversial, and I have not provided much support for it. Still, markets do not seem to value demographic diversity much, at least when we consider sports, marriage, or friendship teams. Perhaps these teams should sacrifice individual ability – regarding, for example, a cornerback, friend, or wife – for diversity-based gain. Alternatively, perhaps these teams are disanalogous to medical teams. In the absence of a reason to believe this, though, the case for diversity-contribution to teams weakens.

In conclusion, then, the theses are compatible with the sabermetric equation for medical school admissions taking demographic factors into account. Consider, for example, the predicted shortage of doctors and the frequency with which female doctors do not work full time. The theses are also compatible with the average underrepresented minority doctor making a (comparative) diversity-contribution to health that is greater than the (comparative) individual-performance-contribution that his or her competitor would have provided. Still, none of these points supports holistic admissions.

In general, medical schools should be wary of explicitly valuing diversity because they do not know whether purely academic-ability-based admissions would generate sufficient (demographic) diversity and, also, what the tradeoff is between diversity-based contribution and individual-performance-based contribution. In addition, we should be wary of the evidence in support of there being a diversity-based contribution to a medical team's performance. This is because the evidence in the medical context is minimal and the evidence outside the medical context is mixed and does not properly account for ability-related loss. This is also because competitive markets – especially ones that can measure individual and team production – do not value diversity much if at all.

Part Six: Conclusions

In this article, I argued that medical school admissions should be limited to statistically relevant factors. I argued for it based on two other conclusions: a medical school should maximize quality-adjusted medical services per graduate within the overall optimum spending limit and if this is correct, then a medical school should, other things being equal, select medical students who are better than their competitors. I then explored the implications of this argument for whether a medical school admissions system should be holistic and whether it should consider demographic factors.

Funding: This research received no external funding.

Conflicts of Interest: The author declares no conflict of interest.