Approaches to Help Understand Disparities

Wassim Tarraf, PhD

RCCN Wokrgroup – Bethesda, Maryland

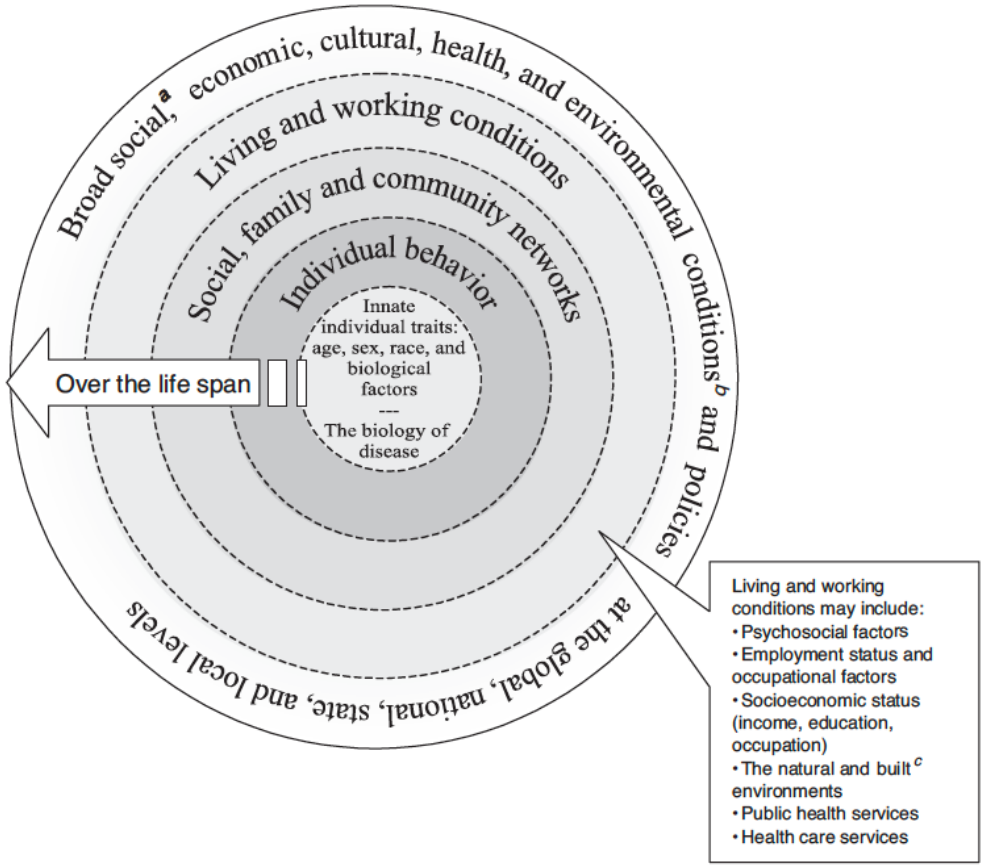
29/06/2022

No Conflicts

Healthcare disparities

- "Differences which systematically and negatively impact less advantaged groups"(1,2)
- "Differences which society has a role in creating, and therefore has the greatest potential to ameliorate" (1)
- "Differences… [in healthcare] that are not due to needs, preferences or appropriateness of intervention"(3)

IOM Model



1.***Dehlendorf, Christine et al***. "Health disparities: definitions and measurements." *American journal of obstetrics and gynecology* vol. 202,3 (2010): 212-3. doi:10.1016/j.ajog.2009.12.003

2.***Braveman P.*** Health disparities and health equity: concepts and measurement. Annu Rev Public Health. 2006;27:167–94;

3.**Smedley B, Stith A, Nelson A**. Unequal Treatment: Confronting Racial and Ethnic Disparities in Healthcare. Washington, D.C: The National Academies Press; 2003

Also see **Fiscella, K., & Sanders, M. R.** (2016). Racial and ethnic disparities in the quality of health care. *Annual review of public health*, *37*, 375-394. and ***Wasserman, J., Palmer, R. C., Gomez, M. M., Berzon, R., Ibrahim, S. A., & Ayanian, J. Z***. (2019). Advancing health services research to eliminate health care disparities. *American journal of public health*, *109*(S1), S64-S69.

Questions of interest

- Does a **risk-exposure** affect a certain outcome **Y, and how**
- Why do **risk-exposure (groups)** differ on a certain **outcome Y ?**
- Do specific risk/protective factors affect/modify **risk-exposure (groups)** differently regarding a certain **outcome Y?**
- How can we manipulate the **risk/protective space** to reduce/eliminate the role of **risk-exposure (groups)** in **outcome Y?**

Standard Regression Techniques

- Interest in linking Exposure (E) and outcome (Y)
- Random sampling from a population of interest
- Probe how Y varies with changes in E (in our context today; dummy indicators of group membership with a reference).

Questions:

- What if y is affected by factors other than E; how should we handle that?
- What is the functional form connecting the two variables?
- Most of us stop there and devise a modeling strategy. BUT
  - How can we attribute the variation between y and E to something other than correlation
  - What if E is confounded by other factors; how should we handle that??

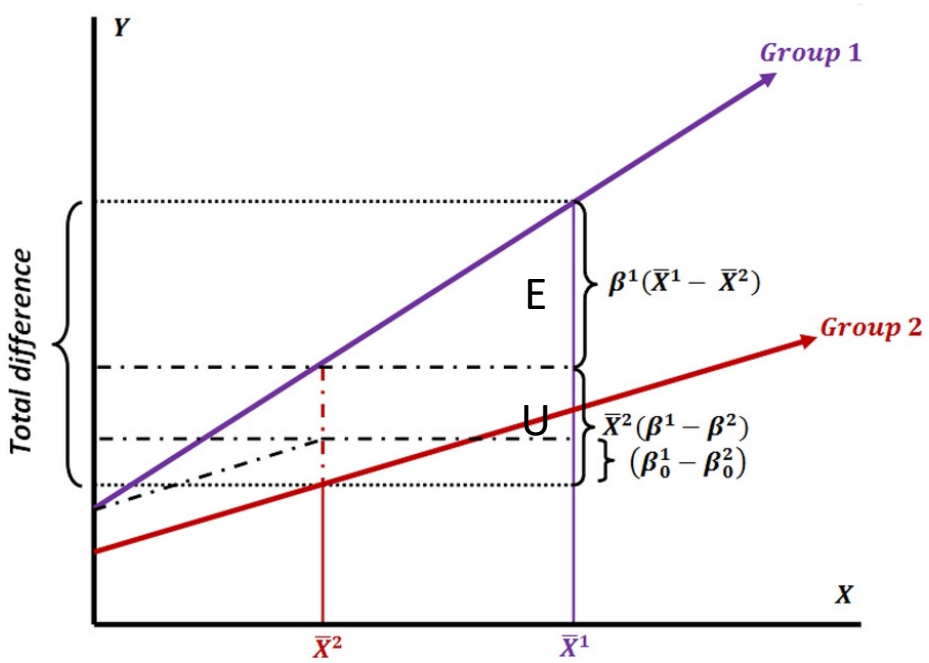
- Standard Regression estimators: Sensitive to dissimilarities in covariates distribution (imbalance between groups) across levels of the exposure (binary, categorical/multinomial or continuous). Not appropriate when we have endogeneity or if there is excessive selection bias.

Correlation does not imply causation

- The rooster and the sun (crows before sunrise), breathing and death (every breathing being dies)
- For two variables X and Y
  - X can cause Y
  - Y can cause X (reverse directional) – violent games and violence
  - Both Y and X are caused by another variable C (common cause) – Sleeping with shoes on causes a headache!
  - X causes Y and Y causes X (bidirectional)
  - X and Y covary coincidentally
- Causal relationship can exist **in the absence of** observed correlation (e.g. confounding, non-linear, etc…).

Petersen, T. (2001). Endogeneity: Methodology. International Encyclopedia of the Social & Behavioral Sciences. N. J. Smelser and P. B. Baltes. Oxford, Pergamon: 4511-4513.

Oaxaca-Blinder/Fairlie Decomposition: The Intuition



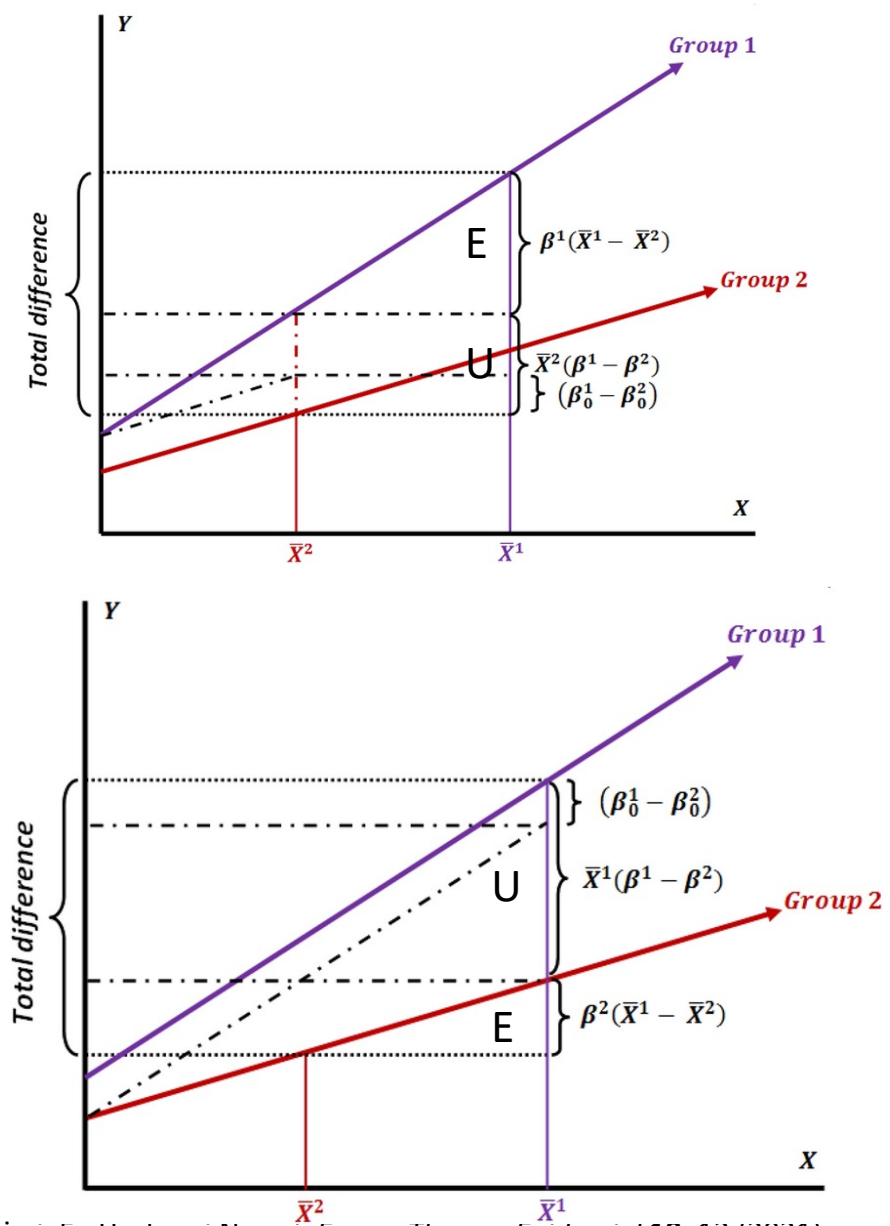
Rahimi, E., Hashemi Nazari, *Emerg Themes Epidemiol* **18,** 12 (2021).

1. The Explained component also referred to as the “endowment effect” (E): Reduction in difference expected if the two groups have the same distribution on the covariates.

2. The Unexplained component, also referred to as the “coefficient effect” (U): Difference not-attributable to the differences in characteristics + effects of unobserved variables (not accounted for in model).

Rahimi, E., Hashemi Nazari, *Emerg Themes Epidemiol* **18,** 12 (2021).

Oaxaca-Blinder/Fairlie Decomposition: The Intuition



Pros

- Useful for making attributions about sources of differences/disparities.
- Actionable as it points to specific factors or cluster of factors that can be targeted for change.
- Accessible with functional implementation across statistical software.

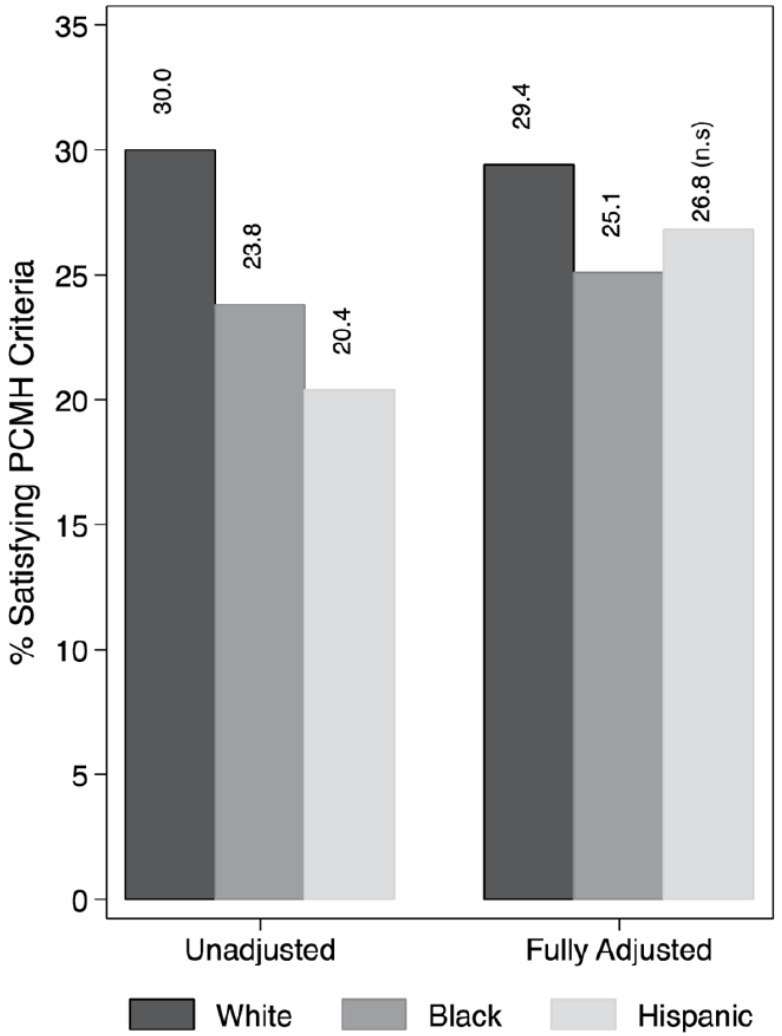
Cons

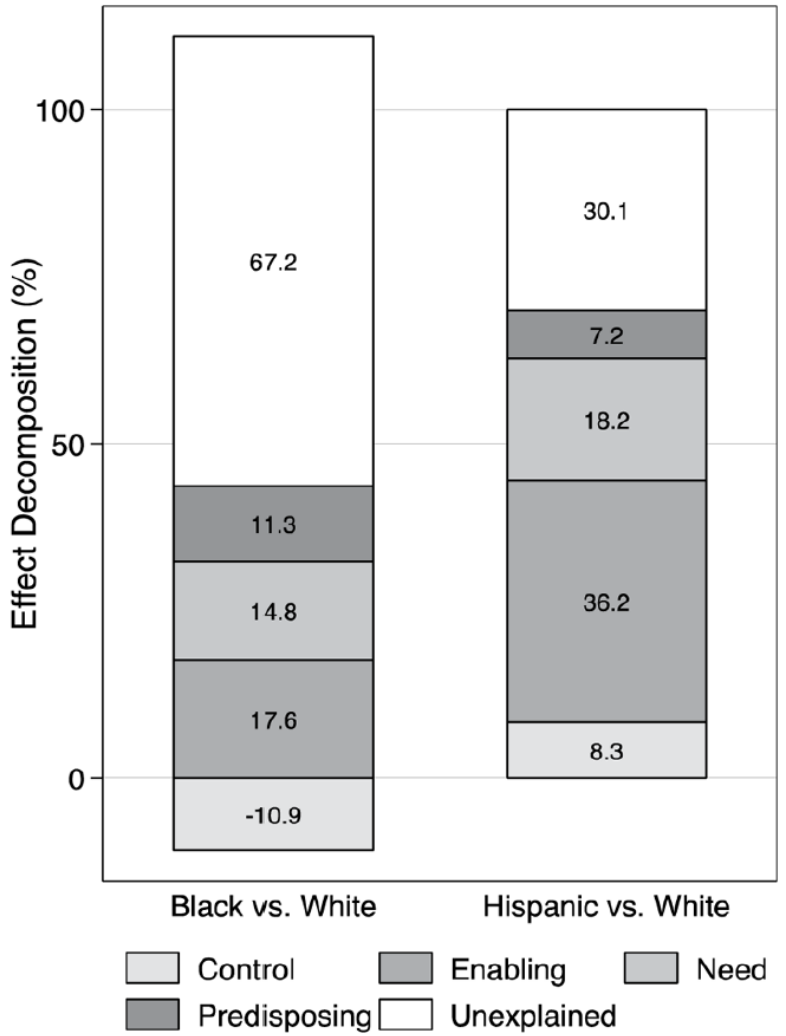
- Strong assumptions that might not be reflective of realistic disparities scenarios; omitted variables might be particularly problematic.
- Other statistical issues; several proposed solutions.

Rahimi, E., Hashemi Nazari, *Emerg Themes Epidemiol* **18,** 12 (2021)

The Blinder–Oaxaca decomposition for linear regression models. *The Stata Journal*, *8*(4), 453-479; Firpo, S. Identifying and measuring economic discrimination. IZA World of Labor 2017: 347 doi: 10.15185/izawol.347

Oaxaca-Blinder/Fairlie Decomposition





Tarraf, W., et al. (2017). J Immigr Minor Health **19**(6): 1271-1280. also see Tarraf, W., et al. (2012). Med Care **50**(3): 233-242.

What is causal inference?

- Methods used to establish cause and effect
- Impact of exposures (“events”, “programs”, “policies”, “categories”, etc…) on outcomes
- Examples:
  - Does diabetes increase the incidence of ADRD
  - Does Medicaid eligibility expansion increase use of preventive services?
  - Do minoritized groups receive lower quality care
- Includes a range of modeling techniques: “Potential outcome models “; “ Graphical Models “; “ Difference in Difference models“, “Regression discontinuity “, “Instrumental Variables models “; “Panel data”; to name a few (see book treatment references at end).
- **Focus is Matching techniques for causal inference**

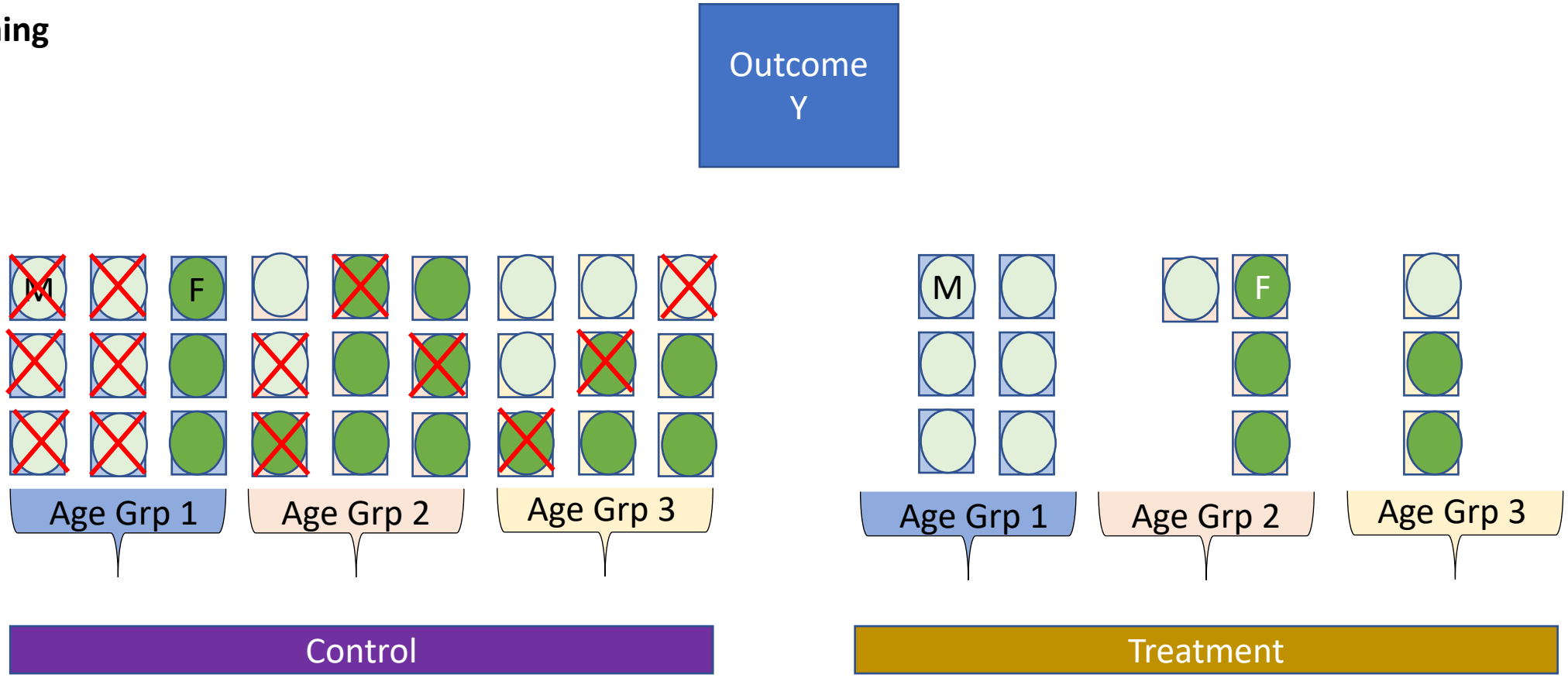
Matching

- **Randomized Experiments (RE):** Key benefit of it is treated and control groups are randomly different on observed and unobserved characteristics
- **Matching methods:** focus on how to replicate RE conditions using observed variables, when randomized experiments "prohibitive" (e.g. you can't randomize **the racialization process**).
- Matching (see process below; and Stuart 2010): Approximate distribution of observed variables in groups (could be extended to multiple groups and to continuous variables).
  - Complements regression techniques
  - Overcome weaknesses of regression and selection techniques in the absence of sufficient overlap in model covariables
  - Has straightforward diagnostics to assess performance

Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical science: a review journal of the Institute of Mathematical Statistics*, *25*(1), 1.

Matching under simple conditions

Age and Sex Differences
1:1 Matching



Process

(1) **Define a distance measure** -> to estimate closeness

- Propensity scores predicted from a model; conditional probability of being in the "treatment" (vs. control) group (assumption is we have a binary grouping).
- Reduces matching problem (comparison of group members) to a single "index".
- Generating the PS:
  - Use of substantive knowledge and previous literature guidance is critical
  - As many variables as warranted but not strongly predictive of the treatment or strongly affected by it.
  - Avoid variables that are not associated with outcomes (unnecessary inflation in variance).

(2) **Do the matching** -> use distance measure from 1

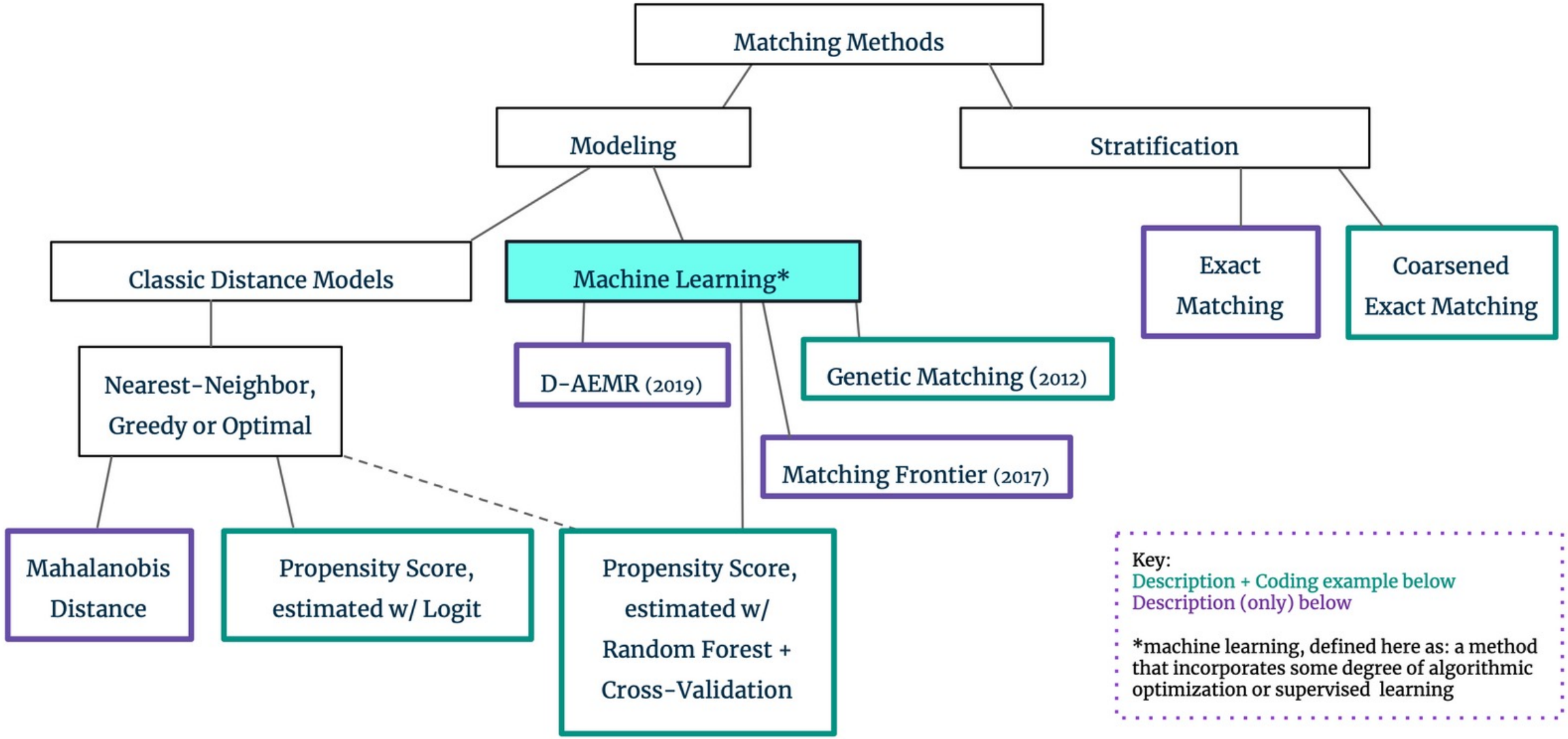
(3) **Diagnose the quality of matching**

- Numerical and graphical diagnostics
  - Standardized Mean Difference (SMD) between matched groups on the distance measure
  - Summaries and SMD of covariables included in the PS model
  - Coverage and overlap of the score distribution

(4) **Estimate effect** (diff. between groups) after matching (crude vs. model based).

Stuart, E. A. (2010). *Statistical science: a review journal of the Institute of Mathematical Statistics*, *25*(1), 1. Austin PC. Multivariate Behav Res. 2011 May;46(3):399-424.

Matching Methods



See: https://humboldt-wi.github.io/blog/research/applied_predictive_modeling_19/matching_methods/

Propensity Score Matching

Pros

- Works well with large # of confounders
- Refocuses attention on the exposure – easier to compare treatment and control groups on a single index to examine imbalance compared to the multivariable setting
- Reduces need for model assumptions – e.g. less dependence on functional form for linking covariates to outcomes

cons

- Unmeasured confounding remains an issue
- Lack of coverage, and need for pruning affects inference to a target population; particularly severe if substantial loss of cases in the treatment group
- Samples size (reduced by matching; but also depending on the required # of matches balance requirement can be large)
- Residual imbalance is not completely avoided

Pro: For a summary of recent debates and methods, see: Guo, S., Fraser, M., & Chen, Q. (2020). Propensity score analysis: recent debate and discussion. *Journal of the Society for Social Work and Research*, *11*(3), 463-482. See

Con: For an extensive critiques of the use of PSM Gary King and Richard Nielsen. 2019. "Why Propensity Scores Should Not Be Used for Matching." Political Analysis, 27, 4, Pp. 435-454.

Propensity Score Matching

Figure 1: Estimated, Unadjusted, Total Health care Expenditures by Nativity Status and Age

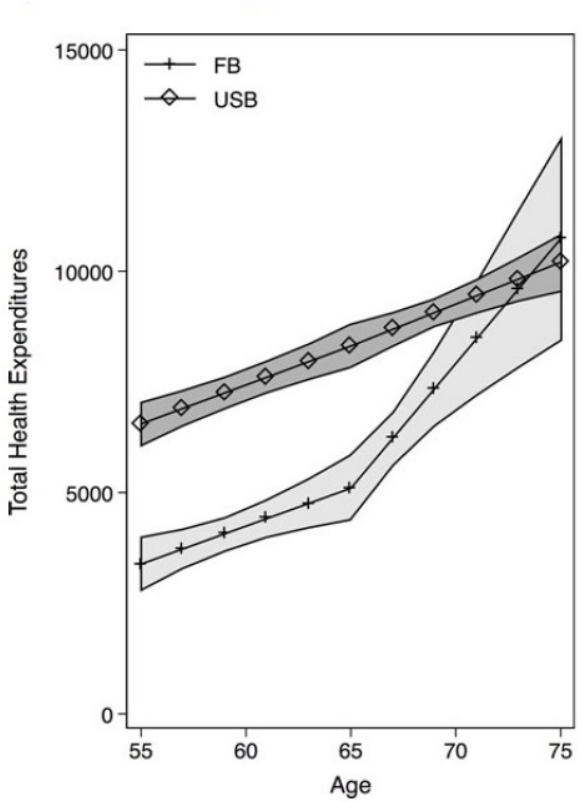
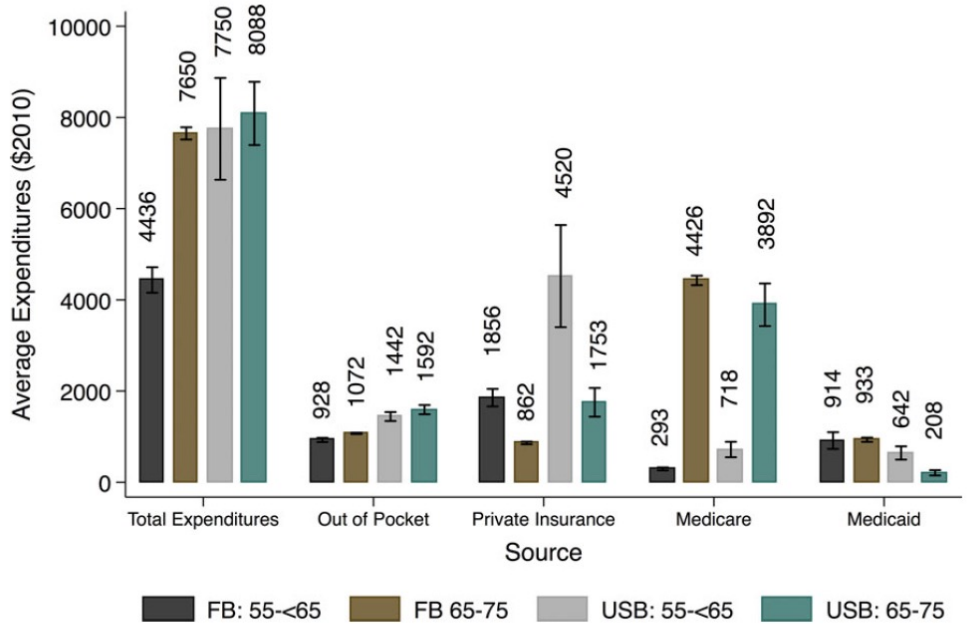


Figure 2: Estimated Total and Payer-Specific Health Care Expenditures by Nativity Status and Age Group



Notes. Estimates are based on demographic and socioeconomic age groups and health needs and health care preferences matched nativity groups. Results are from generalized linear model with a log link using data from respondents ages 55–75 years in the Medical Expenditures Panel Survey (2000–2010). The 95% confidence bounds reflect bootstrapped standard errors based on 500 replicate samples.

	Δ-in-Δ ($)*	*SE†*	*p-value‡*
Overall expenditures	−2,877	626	.000

*Δ-in-Δ = (USB, post-Medicare—FB, post-Medicare)—(USB, pre-Medicare—FB, pre-Medicare). For example, overall spending Δ-in-Δ = $438–$3,314 = −$2,877 and indicates that the average difference in spending (i.e., disparity) between U.S.- and foreign-born respondents decreased by $2,877 in the post-Medicare period.

Tarraf, W., Jensen, G. A., & González, H. M. (2016). *Health Serv Res, 51*(3), 846-871. doi:10.1111/1475-6773.12402

Conclusion

(1) Enhance precision in estimates and allow specific attributions for group differences

(2) Move beyond associations

(3) Exploit available designs; even with secondary observational data

(4) Provide actionable knowledge to inform interventions and policy creation

(5) Move beyond disparities and into equity

General References

 # Resources

Review specific to matching

Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical science: a review journal of the Institute of Mathematical Statistics*, *25*(1), 1.

Survey books on causal inference

Morgan, S., & Winship, C. (2014). *Counterfactuals and Causal Inference: Methods and Principles for Social Research* (2nd ed., Analytical Methods for Social Research). Cambridge: Cambridge University Press. doi:10.1017/CBO9781107587991

Open source full length books on causal inference

Cunningham, Scott. Causal Inference: the Mixtape : https://mixtape.scunning.com/

Bauer, Paul. (2020) Applied Causal Analysis (with R): https://bookdown.org/paul/applied-causal-analysis/

Books on matching

With R code

Leite, W. (2017). *Practical propensity score methods using r*. SAGE Publications, Inc, https://dx.doi.org/10.4135/9781071802854

With Stata code

Guo, Shenyang, and Mark W. Fraser. *Propensity score analysis: Statistical methods and applications*. SAGE publications, 2014.

Thank you!

wassim.tarraf@wayne.edu