THE RELATIONAL ROLE OF PLACE IN THE PRODUCTION OF RACIAL STRATIFICATION

Nick Graetz

A DISSERTATION

in

Demography and Sociology

Presented to the Faculties of the University of Pennsylvania

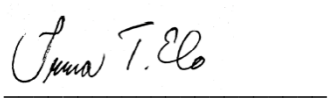
in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

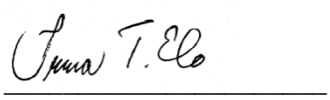
2021

Supervisor of Dissertation



Irma Elo, Professor and Chair of Sociology

Graduate Group Chairperson



Irma Elo, Professor and Chair of Sociology

Dissertation Committee

Courtney Boen, Assistant Professor of Sociology

Xi Song, Associate Professor of Sociology

Tukufu Zuberi, Professor of Sociology, Lasry Family Professor of Race Relations

Daniel Aldana Cohen, Assistant Professor of Sociology

ACKNOWLEDGMENT

Completing this dissertation during all the stress and uncertainty of the pandemic would not have been possible without the unconditional support of so many people in my life.

I've loved taking advantage of so many different scholars at Penn working along so many different intersections of research and policy that I'm interested in. I learned so much during my PhD from my dissertation committee: Irma Elo, Courtney Boen, Daniel Aldana Cohen, Tukufu Zuberi, and Xi Song. I'm immensely grateful for all the ways in which my work has been (and continues to be) influenced by each of you, including feedback on these chapters but also conversations, classes, and your own work in all its forms. I also benefited from many other members of the PSC, including collaborations with Sam Preston and Michel Guillot and my comprehensive exam organized by Chenoa Flippen. Our conversations about theory, methods and measurement helped shaped the way I approach all my research.

Prior to starting my PhD, I benefited so much from the guidance and mentorship of Emmanuela Gakidou. I'm incredibly grateful for all the work you put into cultivating such a great environment for me and other members of the Post-Bachelor Fellowship program to grow and learn together. I leaned on many of these relationships for support during my PhD.

I made a lot of great friendships during my time at Penn, but my writing group was especially supportive during the PhD. Raka, Hannah, Yana, Ellen, and Ashleigh helped to create such a supportive and affirming space that was so necessary to avoid the isolation of the PhD process.

Most of all, I want to thank my family – my mom and dad, and brother, Nathan. I've always felt that I could lean on them during challenging times, and always had a place to go home and decompress. My parents really made me into who I am today, sparking my love of science and always demonstrating compassion, curiosity, and a sense of humor. And last, I want to thank my partner, Steph, the smartest and most deeply caring person I know. I wouldn't have been able to finish this dissertation during the pandemic without your love and support. You constantly show me how to see the joy and hope in the world.

ABSTRACT

THE RELATIONAL ROLE OF PLACE IN THE PRODUCTION OF RACIAL STRATIFICATION

Nick Graetz

Irma Elo

In this dissertation, I examine how we quantify the dynamic, cumulative effects of relational social exposures with longitudinal survey data. In Chapter 1, I demonstrate a new mediation framework for describing what are often conceptualized problematically as "neighborhood effects." Findings from this study clarify the reciprocal, life-course process through which neighborhood is implicated in the early production of social inequality. In Chapter 2, I extend this mediation framework to respond to theoretical critiques of how variables for race are used in common regression frameworks in attempts to study structural racism. I demonstrate an alternative counterfactual approach to explain how multiple racialized systems dynamically shape health over time, examining racial inequities in cardio-metabolic risk. I decompose the observed disparity into three types of effects: a controlled direct effect ("unobserved racism"), proportions attributable to interaction ("racial discrimination"), and pure indirect effects ("emergent discrimination"). I discuss the limitations of counterfactual approaches while highlighting how they can be combined with critical theories to quantify how interlocking systems produce racial health inequities. In Chapter 3, I use this framework to examine the Black-white wealth gap in the United States. Descriptive and qualitative analyses have identified many mechanisms underlying wealth correlations across successive generations, but few studies have quantified the relative contributions of these interconnected and racialized systems of reproduction to the total gap we observe today. I define a wealth gap in 2015-17 between the grandchildren of those racialized as Black and the grandchildren of those racialized as white in 1968-70. I use a fully interacted counterfactual mediation framework to decompose this disparity into the historical, racialized contributions of 1) effects of home values in 1968-70 on home values in successive generations and 2) effects via educational attainment in successive generations. Findings from this study contribute to our understanding of the dynamic, racialized process of multigenerational place-based wealth accumulation and support the importance of historically contingent social policy centered on reparative justice.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

Chapter 1.

PREFACE

Complex exposures and critical quantitative methods

In this dissertation, I examine how we quantify the dynamic, cumulative effects of relational social exposures; specifically, the reciprocal construction of race and place within a system of racism and how to quantify the relations that produce (and reproduce) racialized differences in outcomes related to health and material conditions. The first chapter provides a generalized causal mediation framework for estimating what are often conceptualized problematically as "neighborhood effects." The second chapter extends this mediation framework to respond to theoretical critiques for how the "race" variable collected in surveys is used in regression models to study structural racism in the United States. The third chapter brings these perspectives together to examine widening racial stratification in home values and wealth over three linked generations between 1968 and 2017.

These empirical papers are organized around several broader questions related to causal inference in quantitative sociology and how we define the counterfactual contrasts (estimands) that are most relevant to our relational theories of racism and place. Are these the estimands that we are actually identifying with conventional regression models? How can we better approximate our relational constructs with estimands from causal mediation methods and a life course perspective? Modern counterfactual methods can help to estimate entangled, relational processes over time in ways that are more closely aligned to our theoretical constructs, which are often historically co-constituted. This stands in contrast to counterfactuals offered by conventional regression models, which assume a

much higher degree of conceptual separability between variables (fueling the endless and reductive debates on the relative influence of "race vs. class" in stratification research).

Given relational theories of social processes (e.g., racism) rather than crude models of stratification (e.g., race), what is the best we can do using quantitative methods with individual-level national cohort surveys? When is theoretical nuance lost in the translation between constructivist theories of racism and the restrictive assumptions of conventional regression models? It is not my goal in these three empirical papers to suggest that we can always identify the estimands most relevant to our theories of racism from existing individual- and household-level data. As I echo below, there have been many longstanding calls from critical scholars and activists to shift data collection away from surveilling the oppressed (e.g., surveys of individuals) and towards examining the oppressors (e.g., data on the behaviors of banks, police, landlords, and employers) (Bailey et al. 2017; Ford and Airhihenbuwa 2010; Itzigsohn and Brown 2020; Pattillo 2013). Still, longstanding national surveys such as the Panel Study on Income Dynamics can be invaluable in describing how past relational processes unfolded within particular cohorts aging through particular contexts (Esposito 2019; Prins et al. 2021; Sharkey and Elwert 2011; Wodtke, Harding, and Elwert 2011). Coupled with a strong theoretical framework grounded in the Du Boisian tenets of relationality, contextualization, and historicity (Du Bois 1898, 1899, 1935; Itzigsohn and Brown 2020), these analyses can inform and reflect on existing narratives surrounding the fundamental causes of racialized social stratification. But as I discuss in my three empirical chapters summarized below, this requires grappling with difficult statistical assumptions – especially in elucidating

mediating mechanisms – that I argue are directly implied by the social constructivist perspective (Reskin 2012; Sen and Wasow 2016; Sewell 2016; Zuberi, Patterson, and Stewart 2015).

In all three chapters, I build on a generalized framework for causal mediation analysis to describe the mechanisms through which relational place-based processes produce and reproduce racialized stratification. I summarize below the counterfactual philosophy and quantitative framework underlying all three empirical chapters. This dissertation is primarily a project in empirical translation of existing theory; operationalizing social constructivist, relational theories around racism and place-based exploitation with quantitative estimands, choosing the appropriate estimator, and carefully describing its statistical and the theoretical implications. This requires embracing the notion, long advocated by sociological methodologists such as Christopher Winship, Michael Sobel, Felix Elwert, and many others, that the primary goal of counterfactual models is the causal identification and explanation of social relations.

Counterfactual philosophies

First, we must distinguish between quantitative *prediction* and quantitative explanation (i.e., causal inference) (Gangl 2010; Moffitt 2005; Morgan and Winship 2014). A race variable (e.g., a survey measure asking an individual to self-identify as "Black" or "white") should not be used as a proxy for biological or genetic difference (or even for more complex social relations) in predictive models of future conditions; a

comprehensive discussion is beyond the scope of this project but can be found in Zuberi & Bonilla Silva (2008), Roberts (2011), and Zuberi et al. (2015).

In contrast, it is possible that a variable for race may be used as a proxy in quantitative explanation for past racialization within a specific system of racism, if all estimands are clearly defined relative to a precise research question within a social constructivist theory of racism, racialization, and race that is historically contingent and context specific. There are two distinct interpretations of counterfactual quantities in the broader philosophy of causal inference (Gelman and Imbens 2013; Morgan and Winship 2014; Pearl 2014; Schwartz, Gatto, and Campbell 2016):

1) What would happen in the future if we intervened on an exposure?

2) What would have happened in the past had an exposure been different?

The "race variable" commonly collected in individual- and household-level survey data can be used in some cases to understand and explain racist systems of the past, but this requires the latter philosophy of counterfactuals above combined with a clearly articulated theory of how the system of racism and racialization operated over a specific time and place. Throughout the following examples (and all three empirical chapters), I'm specifically discussing what a "race variable" can represent in quantitative analyses with respect to the populations racialized as Black and the populations racialized as white in the United States over roughly the period 1968-2017.

Consider the use of a self-identified race variable to define a causal contrast using observational data (e.g., data collected from a national survey such as the Panel Study on

Income Dynamics). Specifically, consider the average difference in an outcome between individuals racialized as Black compared to individuals racialized as white as we observed them aging through the system of racism characterizing the United States over a specific period. The fields of statistics and economics (and increasingly, epidemiology) almost always adopt, often implicitly in defining their assumptions, the first philosophical frame above: counterfactual quantities describe *what happens if* we intervene to change the exposure. In terms of using the race variable above to define a counterfactual contrast, this has led to an endless debate in quantitative causal inference about whether "race is manipulable."

> "Properties or attributes of units are not the types of variables that lend themselves to plausible states of counterfactuality. For example, because I am a White person, it would be close to ridiculous to ask what would have happened to me had I been Black. Yet, that is what is often meant when race is interpreted as a causal variable" (Holland 2008, p. 100).
>
> "However, race is not something we can intervene on, and the associated counterfactual queries generally strike researchers as meaningless" (VanderWeele & Robinson 2014, p. 474).
>
> "Making causal inferences usually demands a neatly defined, manipulable treatment variable… research questions for which there are no experimental analogies (even hypothetical ones, in a world with unlimited time and research budgets and

omniscient powers) are fundamentally unidentified questions" (Sen & Wasow 2016, p. 504).

Implicit in all of these quotes is a definition of "race" within the forward-facing frame of "potential outcomes," which imagines all causal contrasts in terms of an idealized randomized-control trial. This further implies that there must be a plausible intervention mechanism – which seemingly strikes these authors as a ridiculous proposition. Kohler-Hausmann (2019) argues that such a counterfactual conceptualization "is wrong because to fit the rigor of the counterfactual model of a clearly defined treatment on otherwise identical units, we must reduce race to only the signs of the category, meaning we must think race is skin color, or phenotype, or other ways we identify group status. And that is a concept mistake if one subscribes to a constructivist, as opposed to a biological or genetic, conception of race. The counterfactual causal model of discrimination is based on a flawed theory of what the category of race references, how it produces effects in the world, and what is meant when we say it is wrong to make decisions of import because of race" (Kohler-Hausmann 2019, p. 1163).

Kohler-Hausmann discusses how such a framework lets biological racial essentialism into the definition required to demonstrate or "prove" racial discrimination within this interventionist approach. It implies that "race" proxies something that is independently manipulable from other social relations. Kohler-Hausmann argues that following this logic, in order to design a well-defined experimental manipulation in a trial carried out today, you'd have to manipulate something phenotypical about the individual to determine whether they are racialized as Black or white within a particular system of

racism and racialization (as interventionists often imply that it is an "implausible" or "vague" manipulation to suggest changing everything else about the system of racism).

This misinterpretation of the potential outcomes counterfactual framework as only relevant to (quasi-)experimental manipulations in the interventionist philosophy has existed in sociology for decades (see the annual review from Winship & Morgan (1999) on the estimation of causal effects from observational data). Gangl (2010) revisited this issue in an updated annual review article on causal inference in sociology and concluded the following:

> "The perception that the counterfactual framework would primarily apply to the effects of policy interventions or other explicitly manipulated (or at least manipulable) treatments is perhaps the single more important impediment to its widespread adoption in sociology. This perception is a major misunderstanding on the part of sociologists (cf. also Heckman 2005, Moffitt 2005, Sobel 1998). (Gangl 2010, p. 38).

In discussing how so many regression analyses attempt to identify effects of social relations via race, gender, and class, Gangl (2010) notes that "the practice of mediation analysis (see MacKinnon 2008) is itself in dire need of being realigned with the potential outcomes framework … all the concerns of causal inference will typically apply at the level of generative mechanisms that constitute the actual causal manipulations behind socially relevant attributes and conditions" (Gangl 2010, p. 40).

The normative consequences of "well-defined interventions"

The interventionist philosophy of counterfactuals, focused on "plausible" forward-facing interventions, is often used to argue for a shift towards "consequentialist" and "pragmatic" social science that can be used for decision-making that is "in the policy sphere" (Galea and Link 2013; Galea, Riddle, and Kaplan 2010; Robinson and Bailey 2020; Schwartz, Prins, et al. 2016). Before discussing how variables for race and place can be used more critically in causal explanation, it is worth noting how this interventionist frame has hampered counterfactual reasoning more generally around complex exposures in quantitative sociology – especially as quantitative sociologists and demographers are increasingly likely to cede this intellectual ground by simply avoiding use of the words "cause" and "causation" in favor of the more vague language of "links" and "drivers" (Broadbent 2019; Hernán 2018; Jackson and Arah 2019).

Causal frames and their associated quantitative methods can premise either 1) technocratic, incremental solutions that rely on the unilateral academic-policymaker relationship or 2) an emancipatory, democratic theory of change in solidarity with existing social movements to affect systems (Ford and Airhihenbuwa 2010; Itzigsohn and Brown 2020; Krieger 2011; Prins and Schwartz 2020; Robinson and Bailey 2020). Itzigsohn & Brown (2020) describe this phenomenon in their manifesto for a Du Boisian sociology:

> "Although we hold 'policy relevance' in high regard, this notion is very narrowly and technocratically conceived and see as worthy only when these contributions directly impact the process of policy making. Furthermore, mainstream sociology

> detaches policy making from the political process that makes policy making possible" (Itzigsohn & Brown, 2020; p. 209)

There are many social and political arrangements in the production of social scientific research that reward the interventionist causal frame and its normative authority in debates of social policy. By limiting our definition of methods for "causal explanation" to only those within the interventionist frame, we risk reifying specific paradigms as "objective" and assigning normative authority to specific technocratic theories of social change (Schwartz, Prins, et al. 2016).

> "Should we primarily address our research to – and identify professionally with – people in positions of power who determine the policy space (Chomsky and Foucault, 2006), or should we direct it toward grassroots social movements that are largely alienated from – and must apply pressure from outside – the policy space? …We suggest that, particularly with regard to social factors, the well-defined intervention view of 'policy relevance' sidesteps these questions, in favor of a narrow consideration of how social policy is developed and how scientific evidence informs this process. The determination of what is relevant and plausible is thus presented as an objective, scientific exercise when in reality it is often also a political calculus." (Schwartz et al. 2016; p. 256).

Consider the theoretical pitfalls of the famous Moving to Opportunity (MTO) experiment, which randomly provided housing vouchers to low-income families in order to study "neighborhood effects" purged of "selection bias" via randomization of the

treatment variable (Sampson 2008b). A research project such as the MTO experiment can yield precise causal evidence for a very particular "plausible" intervention: expanding voucher access in the private housing market. While helpful academically in generating specific causal contrasts that may be less influenced by real-world (non-random) "selection" forces into and out of neighborhoods, the research project is also used to neatly define and maintain a "policy sphere" for the provision of housing that consists of only those programs for which we can generate precise experimental evidence via seemingly airtight manipulations, such as voucher programs. Investment in the MTO experiment as the paradigm for "objective social science" is fundamentally non-threatening to many entities that have a vested interest in preserving the status quo of housing provided via the private market (i.e., speculative investors in real estate). It is unlikely that evidence generated by the MTO experiment will be used to argue for housing as a human right or establishing a more robust public option (e.g., a Social Development Housing Authority) from a fundamental cause framework (Baiocchi et al. 2020; Link and Phelan 1995; Williams and Collins 2001), because the experiment is specifically designed to be decoupled from the real-world forces that currently determine where individuals live, how they are influenced by those places, and how this all influences where they might live in the future (Pattillo 2013; Slater 2013). The high internal validity of a precise manipulation comes at the cost of theoretical generalizability that might speak more directly to the fundamental causes of residential segregation and cyclical "neighborhood effects" in the United States.

The MTO experiment reflects the sociological perception critiqued by Gangl (2010) and others of what constitutes "real causal inference" – and this logic carries into quantitative studies of racism and racial stratification. For example, the interventionist perspective premises audit studies of employers which neatly separate manipulable variables on a resume into "race" (e.g., the name of a jobseeker) and control for variables that are "not race" (e.g., the jobseeker's educational attainment). In cases involving hiring discrimination, the residual racial variation that remains after considering other factors related to hiring, such as educational attainment, is interpreted as *the* effect of racial discrimination (Kohler-Hausmann 2019). Audit studies are often interpreted as premier causal evidence of racial discrimination, as if naming a hypothetical job-seeker Jamal vs. Brendan, for example, while holding everything else constant, is capturing something that is more *truly representative of racism* than the complex, longitudinal systems that produce racialized distributions of everything else that might appear on a resume.

While audit studies identify important dynamics of racist systems, they typically offer no concurrent analysis of how much the *total observed racial disparity in an outcome* such as hiring is due to employer decision-making at the final step before employment versus, for example, racial differences in exposure to penal systems that shape the probability of an employer receiving differently racialized individuals' resumes in the first place. Indeed, by holding all else constant, these studies implicitly suggest that it is possible to separate "racial discrimination" from factors that are "not race," like socioeconomic position. But racial variation in *any* characteristic can only be the result of historical and contemporary projects of racism and racialization (Bonilla-Silva 2009;

Reskin 2012; Roberts 2011; Sewell 2016; Williams, Priest, and Anderson 2016). As with the MTO experiment, audit studies create an artificial world (decoupling the racialized status of the jobseeker with all other racialized variables that affect employment decisions) where it is possible to isolate the effect of a precise manipulation at the cost of losing sight of the broader entangled system of real-world racialization and racial discrimination which results in *total* racialized disparities. While we might find a significant p-value for the experimental average treatment effect, we have no idea how consequential this particular effect pathway is within the broader system of racism producing racialized differences in employment outcomes. Does it plausibly explain 50% of the total racialized difference in hiring outcomes? 1% of the difference? As with the MTO experiment, audit studies implicitly premise a policy agenda focused on antibias training for employers evaluating resumes rather than a broader agenda focused on dismantling the interconnected structural racism in education, labor, and carceral systems (Cogburn 2019; Reskin 2012). It risks primarily focusing on individual acts and litigating the causal evidence required to "prove" racial discrimination rather than interrogating the larger ecology of structural racism itself (Bonilla-Silva 2009; Kohler-Hausmann 2019).

"Realized counterfactuals" of the past

None of the above discussion of the interventionist counterfactual framework is meant to suggest that carefully designed (quasi-)experiments and audit studies are not useful for generating a certain type of causal evidence for social policy, and these studies are of course relevant to sociological theory. Importantly, the motivations for these experimental studies are grounded in concerns related to selection and post-treatment

confounding which are endemic to observational studies of racism and place-based relations, and these must be addressed to construct the closest counterfactuals to our relational theories (all three empirical chapters consider these methodological issues) (Gangl 2010; Morgan and Winship 2014; Oakes et al. 2015; Sen and Wasow 2016; Sharkey and Faber 2014). The response to charges of "vagueness" and "description" leveled at sociological work based in counterfactual regression models should not be to cede the grounds of "real" causal inference to the interventionist frame and only use words such as "links" and "drivers" (Hernán 2018; Winship and Morgan 1999; Winship and Sobel 2004), but rather to sharpen our estimands so that they most accurately reflect theoretically plausible assumptions about time-varying confounding and multiple mediation – *without* sacrificing theoretical generalizability for the sake of internal validity and precision. In other words, embracing that the goal of counterfactual models is most often *causal* understanding, which is more or less justified *given the underlying social theory, study design, and counterfactual assumptions*. As long argued by sociological methodologists such as Christopher Winship, Michael Sobel, and others (Gangl 2010; Morgan and Winship 2014; Winship and Morgan 1999; Winship and Sobel 2004), the use of regression models with many control variables has increased exponentially in applied quantitative sociology, with coefficients of the target independent variable interpreted as demonstrating a "link" to the outcome. These scholars contend that the widening semantic divide between "associational" studies and "causal inference" when applying these models to observational data is used to sidestep critical inferential issues (especially in longitudinal settings) related to study design, confounding, and mediation. Ceding of the "causal inference" space in quantitative

sociology to only (quasi-)experimental designs – and thereby ceding normative authority to fields grounded in positivist rather than relational theoretical frameworks – is a disservice to the many important ways in which holistic, etiologic counterfactual models speak to *causal explanation* in social theory and policy (Elwert and Winship 2014; Itzigsohn and Brown 2020; Mackie 1974; Matthay et al. 2020; Moffitt 2005; Muntaner 2013; Pearl 2014; Schwartz, Prins, et al. 2016; Schwartz, Gatto, and Campbell 2017; Sharkey and Elwert 2011; Winship and Sobel 2004; Wodtke et al. 2011).

The alternative counterfactual philosophy to the interventionist frame might ask: *What would have happened had an individual been racialized differently by the entire time- and place-specific system of racism through which they aged?* Here it is not necessary to articulate a specific intervention mechanism (and indeed, would require the simultaneous manipulation of many entangled social, economic, cultural, and political systems). We cannot observe history again under a radically different set of circumstances, but we can construct a plausible counterfactual comparison, operationalizing our theory using the observed time-varying and reciprocal relationships across variables (Naimi 2016; Schwartz, Gatto, et al. 2016). This is far from a new or novel conceptualization of this specific research question. The question of whether those racialized as white would fare the same as those racialized as Black in the United States if subjected to the same types of injustice has been repeatedly posed since at least the early 1800s (for review, see Krieger 2014). However, as I will argue in my three empirical chapters, this continues to be a question frequently posed but not accurately answered with the conventional methods of applied sociology. I attempt to demonstrate how we can

better operationalize research questions around the retrospective, relational process of structural racism using common individual- and household-level survey data, while paying attention to the fundamental limitations of this data and framework.

In this retrospective conceptualization of the target causal contrast, a race variable is *still a static proxy for the underlying relational social construct*. Whereas the forward-facing interventionist perspective often implies that race is (in part) a proxy for biological or genetic differences and uses this as an argument for "non-manipulability" (Holland 2008; Sen and Wasow 2016; VanderWeele and Robinson 2014), constructing a counterfactual of the past still requires using self-reported race (given a restrictive set of categories) to proxy the ways in which individuals are racialized by many diverse social, economic, and cultural relations. We often assume in measurement and models that these processes are constant over the periods of time and space between observations. For example, say we are interested in the effect pathways through which higher incomes are translated to higher home values. An individual self-identifies as Black in a survey and we measure their income at one time point and their home value at another time point. We can draw the causal arrow between racialized status and the relationship between income and housing that we know exists given a theory of structural racism and a mountain of empirical evidence from qualitative and quantitative study of the racialized private housing market. But this still flattens the underlying process, assuming that a given individual who identifies as Black in our survey is racialized as Black by all intervening processes. We know that racialization is fluid and processes can shift as individuals move through different social, cultural, and physical spaces. For example,

Sewell (2016) expands on the multifaceted and reciprocal "racism-race reification process." At various stages throughout exchange in the housing market, individuals may be racialized in different ways (and subsequently treated in racist ways) based on physical appearance, name, place of current residence, occupation, and more (Sewell 2016).

Still, carefully constructed retrospective counterfactuals of structural racism remain the best that we can do with most publicly available longitudinal data used to describe the dimensions of economic and social stratification, such as the Panel Study of Income Dynamics. While quantitative researchers can do better with the cohort data available, we cannot lose sight of the data needed to produce social justice as advocated by critical scholars and activist-academic collaborations (Watson-Daniels et al. 2020). Self-identified race and flat measures of social stratification (e.g., income, educational attainment) will always be insufficient and efforts should be expanded to measure the political economy of racist actors and institutions (employers, housing, healthcare, banks, agents of state violence) (Sewell 2016; Williams et al. 2016). In this dissertation I only examine how we can better operationalize research questions around the relational process of structural racism using publicly available longitudinal cohort data collected at the individual- and household-level. I acknowledge that until the object of data collection radically shifts from the oppressed (individuals, households) to the relations of oppression, this process will always involve a tradeoff between construct validity (specific relations of racism and racialization) and external validity (nationally representative patterns reflecting structural racism). While we must shift to collecting

data on the practices of racist actors and institutions, we must also practice better critical quantitative methods with individual- and household-level data (Sewell 2016; Williams 2019; Zuberi and Bonilla-Silva 2008). We need to state the limitations of what we can do quantitatively with proxies for racism and racialization to understand *what did happen in the past* – especially when this requires many parametric assumptions and complex exposures, as in studying the entangled systems of structural racism (Reskin 2012).

Counterfactual mediation and post-treatment bias

Both counterfactual philosophies often rely on the same estimators, core assumptions (e.g., consistency, positivity, exchangeability), and definitions of confounding and mediation to describe the identification of counterfactual contrasts (Gangl 2010; Morgan and Winship 2014; Winship and Sobel 2004).

- **Assumption 1**: Consistency (i.e., the exposure is well-defined)
- **Assumption 2**: Positivity (i.e., exposure overlap over all confounders, or "common support").
- **Assumption 3**: Exchangeability (i.e., no unmeasured confounders).
- **Assumption 4**: For a time-varying exposure, correct handling of time-varying confounding. This requires g-methods (inverse-probability-of-treatment weighting, g-formula, etc.).
- **Assumption 5**: For mediation analyses, correct handling of multiple dependent mediators and post-treatment confounders influenced by the exposure. This requires g-methods (inverse-probability-of-treatment weighting, g-formula, etc.).

As demonstrated in the quotes on "race" being a "non-manipulable exposure" above, Assumption 1 is often conflated with the need to articulate a specific intervention to change exposure levels within a given individual (Gangl 2010; Glymour and Glymour 2014; Krieger 2014; Schwartz et al. 2017). Sometimes this is possible in an experimental setting, such as the audit study where the racialization of a given individual by an employer can be manipulated (to a very specific degree) by changing the individual's name on a resume while holding all else constant (Pager 2003). More often with observational data, it is not possible to consider exactly how we might go back in time and precisely change something about the world such that a given individual racialized one way was instead racialized another way by all intervening systems of racism influencing the outcome.

But this doesn't mean we cannot use a race variable to retrospectively identify important causal contrasts resulting from a system of racism. If we clearly define the exposure status of "Black" from a relational, constructivist perspective of "being racialized as Black within a specific period by a specific system of racist actors and institutions," then it is trivial to imagine multiple exposure states for a given individual – all of which may result in different outcomes. Again, this attempted interpretation of a race variable in a conventional regression model is not new. But because this exposure would have affected virtually everything else that we might measure in a survey, all other measured variables are most appropriately considered as mediators, which forces us to reckon with issues related to post-treatment bias (Assumptions 4 and 5) (Jackson and VanderWeele 2019; Naimi 2016; Sen and Wasow 2016). By mishandling these post-

treatment assumptions and simply controlling for all other variables related to racism and racialization (e.g., income), conventional regression results in static counterfactual contrasts that are not theoretically coherent with our stated relational exposure ("being racialized as Black within a specific period by a specific, historically contingent system of racist actors and institutions").

Tremendous progress has been made in overcoming issues related to time-varying confounding (especially in mediation analyses) over the past two decades, largely in statistics and epidemiology (Gangl 2010; Hafeman and Schwartz 2009; Morgan and Winship 2014; VanderWeele and Tchetgen Tchetgen 2017; Wang and Arah 2015). Uptake has been limited in quantitative sociology, in part due to the aforementioned epistemological misinterpretations that makes applied translation difficult (Gangl 2010). In this way, my empirical chapters described below are about *translation*; how we operationalize nuanced relational theories using quantitative counterfactual inference and mediation analysis – and importantly, where this approach is fundamentally limited. What quantitative methods should we use with observational longitudinal data? How should they be interpreted, and what are the limitations? How do methods center different political philosophies of social change? How can modern counterfactual inference, especially rigorous consideration of time-varying confounding dynamics, be incorporated into a broad, holistic theory of etiologic causal inference around racism and place-based effects? Ultimately, we must acknowledge that quantitative methods make implicit theoretical assumptions about how we think a given social process works, and these assumptions must be clearly defined. We can then acknowledge the usefulness (and

limitations) of counterfactual thinking, and how to best triangulate with a diverse set of empirical methods.

Summary of empirical chapters

In Chapter 1, I contextualize several longstanding problems in the quantitative literature on the influence of place-based social exposures on individual outcomes. Until recently (Pais 2017; Sharkey and Faber 2014; Wodtke and Parbst 2017), the traditional "neighborhood effects" literature has largely been preoccupied with identifying an ambient and amorphous *direct effect* of so-called "neighborhood disadvantage." Problematic theories of "social efficacy/cohesion" often implicate a "culture of poverty" causal narrative. But there is a fundamental identification problem in virtually all quantitative studies of neighborhood effects. Examined over time, individual poverty status is related to neighborhood attainment ("selection bias", especially interacted with racial segregation in the U.S.) and neighborhood attainment is related to individual poverty ("cumulative disadvantage" via the place-based relations that distribute socioeconomic resources in the U.S.). These reciprocal dynamics cannot be handled with conventional regression methods. Scholars virtually always control for individual poverty in longitudinal regression models, which eliminates the former path (confounding) but over-controls the second path (mediation). I use a counterfactual approach based on the parametric g-formula that can handle both dynamics over time to decompose a common "total effect of neighborhood disadvantage" to direct and indirect effect pathways via seven hypothesized mediators. In these data, the majority (~73%) of the large *total neighborhood effect* on an indicator of child vocabulary development is due to the

production of household poverty and low-quality schools. Relatively little of the remaining *direct effect* is explained by the often-cited theories of "neighborhood effects" regarding social efficacy, family disruption, or physical environment. I contend that the "neighborhood effect" on inequality in child developmental trajectories is better understood as the totality of ways in which place is used in the United States to organize the racialized production of household poverty and access to high-quality schools, rather than some amorphous direct effect that exists above the individual. I discuss how this empirical framework challenges common conceptions of what is being measured in studies of "neighborhood effects" and the restrictive assumptions of conventional regression models for examining dynamic social processes. The effect pathways identified here are very racially stratified in the United States, and there is increasingly little overlap in the distributions of most measures of neighborhood poverty between those populations racialized as Black compared to those racialized as white. An ideology of neighborhood effects that controls for individual and household socioeconomic position to isolate small marginal pathways via constructs like "social cohesion" is not only frequently misattributed due to time-varying confounding but may also serve to obscure and unjust political economy of place while reinforcing individualistic narratives that pathologize racialized families and spaces (Harvey 2021; Logan and Molotch 1987; Slater 2013; Wacquant 2008).

In Chapter 2 (co-authored with Courtney Boen and Michael Esposito), we discuss persistent issues with the inclusion and interpretation of a self-identified race variable in conventional regression frameworks for studying the life-course emergence of racialized

health disparities. Building on theories of the dynamic social processes of racism, racialization, and race described by Bonilla-Silva, Zuberi, Roberts, Sewell, and Reskin, we demonstrate an alternative counterfactual approach that follows directly from a social-constructivist theory of how the historical and contemporary process of racism dynamically shapes health through observed and unobserved mutually reinforcing systems of structural racism (Bonilla-Silva 1997; Reskin 2012; Roberts 2011; Sewell 2016; Zuberi 2001). We estimate the cumulative contributions of racially stratified mediating exposures (e.g., early-life and adult socioeconomic exposures) to the total racialized disparity in adult cardio-metabolic risk via racial stratification in their distributions (i.e., *emergent discrimination*) and their effects (i.e., *racial discrimination*) – all of which are part of the "race discrimination system" described by Reskin (2012). In contrast to conventional regression, this method accounts for the complex interplay of time-varying confounding and mediation that is required in operationalizing a "race" variable as part of a dynamic social *process* via multiple mediating systems rather than a static, separable characteristic of the individual. We find that 69% of the total racialized disparity in adult cardio-metabolic risk in this cohort can be explained by the cumulative racial and emergent discrimination of racialized individuals via the socioeconomic indicators considered, while the remaining 31% is explained by unobserved mediating systems of racism. These pathways represent the fact that respondents racialized as Black who completed college and who have a mother who completed college still experience racial discrimination in the underlying systems translating those capital to better health outcomes in the US. There is also significant emergent discrimination via stratified exposure to neighborhood poverty; in other words, the distribution of neighborhood

poverty is extremely segregated but ultimately affects health regardless of racialized category. This feature is important in operationalizing the "race discrimination system" because it highlights how certain mediating subsystems produce racialized health disparities, even if the specific connections between that subsystem and health are not racialized. In summary, we use self-identified race to make retrospective inferences regarding the underlying process of racism which racialized individuals in this cohort and acted upon them in racist ways to influence important biomarkers. Modern quantitative causal inference in the study of health is pushing research questions towards "well-defined interventions" seen as more "proximal" to the individual, which reinforces a neoliberal paradigm of social change predicated on marginal interventions that do not fundamentally threaten structural or institutional arrangements. At the same time, conventional regression models used to study mediating pathways of structural racism are subject to time-varying confounding issues, and at worst can reify and essentialize notions of individual race as a separable "risk factor" from all other social relations. But an uncritical application of more complex methods risks losing sight of the fundamental causes (e.g., racism in the housing market) that govern the distribution of more proximal risk factors, and ultimately the total racialized disparity in adult population health. Our analysis offers one framework for leveraging new developments in quantitative causal inference to decompose a population disparity observed at a given time to the historical life-course process through which individuals come to embody racist structural systems – or as described by Quincy Stewart (2008), "swimming upstream."

In Chapter 3, I combine this counterfactual mediation framework with an intergenerational analysis of the role of home values in producing and maintaining the Black-white wealth gap between 1968 and 2017. There is an expanding literature focused on wealth inequality and the long-term pathways connecting parents' wealth to children's outcomes. Recently, there has especially been an increased focus on racialized wealth gaps in the United States. There are two important gaps in this literature: 1) only scant descriptive work on the persistence of wealth over multiple generations at the household-level and 2) an absence of empirical work examining the influence of various historical mechanisms connecting familial wealth over multiple linked generations to contemporary wealth outcomes. Most of the mediators hypothesized to explain the persistence of the racial wealth gap focus on the momentum resulting from extreme historical dispossession of Black wealth and the continued exclusion of Black households from the financial instruments necessary for wealth accumulation. Most notable across these financial instruments is the empirical focus on home ownership. However, not all home ownership is equal, and significant expansion of Black home ownership in the 1950s and 1960s was in extremely substandard housing stock (Taylor 2019). Using home ownership as a proxy for access to an important instrument of household wealth accumulation may mask significant racialized heterogeneity across groups. I examine *home value* as an important mediator of the Black-white wealth gap since 1968. Using linked data for three generations (1968, 1984, 2017), I examine grandparents' 1968 home values as a fundamental intergenerational determinant of contemporary racialized gaps in home value and total net wealth for grandchildren in 2017. I use multiple mediation analysis to decompose the total racial disparity in grandchildren into 1) *indirect effects* operating

through grandparent and parent home values influencing *educational attainment* in the subsequent generation, and 2) *direct effects* through which home values in one generation influences home values in the next generation through other pathways.

CHAPTER 1

Mechanisms connecting neighborhood disadvantage to child development: A life-course mediation analysis

Nick Graetz[1]

[1]Population Studies Center, University of Pennsylvania

Abstract Many observational studies of neighborhood effects find small associations between indicators of neighborhood disadvantage and markers of childhood development. Recent analyses have demonstrated that this is often due to over-controlling individual- and household-level variables that are part of the neighborhood selection process but also mediate the ways in which neighborhood context affects individuals over time. I use the mediational g-formula to estimate a total cumulative effect of neighborhood disadvantage on childhood development (ages 0-9) using the Fragile Families cohort (1998-2013, N=4,898). Addressing estimation issues related to time-varying confounding, I decompose this total effect into direct and indirect effects via six dependent mediators. I find that the majority of the total effect operates indirectly through shaping household poverty and access to high quality schools, with significant racialized heterogeneity in these effects. Findings from this study clarify the reciprocal, life-course process through which neighborhood is implicated in the early production of social inequality.

There has been a rich debate in sociology around the importance of neighborhoods in producing patterns of social stratification (Krysan and Crowder 2017; Massey and Denton 1993; Sampson 2012; Sharkey 2013; Wilson 1987). The focus in quantitative studies almost universally centers on the question of whether or not exposure to some form of neighborhood disadvantage is "important" in explaining an outcome *net* of individual characteristics (Sharkey and Faber 2014). Empirical results have been mixed, with several large studies failing to identify a significant effect of neighborhood disadvantage on indicators of socioeconomic well-being (Ginther, Haveman, and Wolfe

2000; Jencks and Mayer 1990; Sampson, Morenoff, and Gannon-Rowley 2002; Small and Newman 2001). However, Chetty et al. (2016) recently analyzed data from the Moving to Opportunity Experiment (MTO) and found longer-term, positive effects for children moving to lower-poverty neighborhoods at younger ages.

With the exception of experiments such as MTO, the broader "neighborhood effects" literature relies almost entirely on observational cohort or cross-sectional data where the outcome and neighborhood residence are measured simultaneously. It is well-established that selection into neighborhoods is far from random, compelling researchers to control for many other factors in an attempt to overcome "selection bias" in estimating effects (Sampson 2008a). Recently, scholars have raised concerns of using conventional regression techniques for studying an exposure that not only changes over time but is *reciprocally intertwined* with individual characteristics, such as socioeconomic status, that also impact outcomes and future exposure (Kravitz-Wirtz 2016; Sharkey and Elwert 2011; Wodtke et al. 2011).

Insights from the life-course perspective and cumulative (dis)advantage theory posit that there are complex pathways through which exposures such as neighborhood context may accumulate to influence individuals over time (Diprete and Eirich 2006; Kane et al. 2018; Kramer et al. 2017; Krieger 2001a), but this temporal complexity is rarely operationalized in quantitative models of observational data. Research in the field continues to employ conventional regression or matching approaches that fail to capture the time-varying direct and indirect effects of neighborhood on outcomes (Douglas S. Massey et al. 2018; Haskins and McCauley 2018; Sharkey and Faber 2014). While

common individual- or household-level controls (e.g. poverty status) are useful in addressing selection bias, they also over-control the ways in which neighborhood exposure influences outcomes indirectly by patterning poverty status over time. Several recent studies overcome these statistical issues by drawing on new developments in causal inference to identify effects of neighborhood on health, high school graduation rates and development using inverse-probability-of-treatment weighting (Kravitz-Wirtz 2016; Sharkey and Elwert 2011; Wodtke et al. 2011). These frameworks demonstrate that the total neighborhood effect is larger when accounting for indirect accumulation, but they do not specifically tell us *why* it is larger (Sharkey and Faber 2014).

In this study I use a g-computation framework to identify the total effect of neighborhood disadvantage on children's cognitive development (Hernán and Robins 2019; A. Naimi, Cole, and Kennedy 2016; Wang and Arah 2015). Moving beyond an examination of the total effect, I estimate *how* neighborhood exposure influences children's early development both directly and indirectly through patterning other time-varying exposures: household poverty, school quality, unemployment, family structure, physical environment, and collective efficacy (Hernán and Robins 2019; Robins and Hernan 2009; VanderWeele 2016). I also fully stratify this effect decomposition by race to examine racialized heterogeneity in these mediating processes. By reframing point-in-time individual-level characteristics as accumulations of lived experience, this study contributes to scholarship on disparities in early development and moves quantitative life-course methodology closer to theory regarding the cumulative pathways through

which neighborhood is used in the arrangement and production of social inequality (Diprete and Eirich 2006; Kramer et al. 2017; Slater 2013).

The study of neighborhood effects

Many studies of neighborhood effects have focused on outcomes in childhood as a critical window for the intergenerational transmission of place-based disadvantage. Healthy cognitive development in early life has been particularly scrutinized, as this is linked to later socioeconomic status, health, and educational attainment (Christopher Auld and Sidhu 2005; Murnane and Levy 1996; Singh-Manoux et al. 2005). Mayer & Jencks (1989) provided the first comprehensive review of the quantitative literature on neighborhood effects, noting mixed evidence on childhood academic achievement and graduation rates net of individual and family characteristics (Mayer and Jencks 1989). More recently, Burdick-Will et al. (2011) brought together data from three studies in Chicago that leveraged quasi-experimental designs to demonstrate that living in a highly disadvantaged neighborhood was associated with a 0.15-0.25 standard deviation decrease in scores on various tests for healthy early development (Burdick-Will et al. 2011).

These studies largely inform the question of whether or not neighborhood context is important in explaining variation in childhood outcomes, with the problematic implication being that the answer is a dichotomous *yes* or *no*. In their review of the field of quantitative neighborhood effects research, Sharkey & Faber (2014) called for scholars to move beyond this question (what they term the *dichotomous research question of neighborhood importance*) and focus on applying a life-course perspective to understand

why, when, and how neighborhood influences outcomes (Sharkey and Faber 2014). Neighborhood cannot be considered as a static exposure that is separable from individual and family characteristics, but rather an accumulated experience that manifests in characteristics often considered as individual-level controls.

Neighborhood in a life course perspective

In considering the longitudinal analysis of neighborhood effects, it is useful to invoke the life-course perspective. Life-course theory focuses on sequential timing and duration of events, responses to those events, and how events and responses are reciprocally intertwined with each other (Elder 1985, 1998; Elder, Crosnoe, and Kirkpatrick 2003; Elder Jr 1994). In this perspective, neighborhood patterns many other characteristics over time, including factors related to economic wellbeing, family transitions and chronic stressors (Chetty, Hendren, and Katz 2016; Ludwig, Greg J Duncan, et al. 2013; Ludwig, Greg J. Duncan, et al. 2013; MacMillan and Copher 2005; T. and J. 2003; Wen, Browning, and Cagney 2003). In turn, these characteristics influence the outcome *and* selection into different types of neighborhoods. Figure 1 describes these simultaneous issues: 1) "selection bias", whereby neighborhood is not randomized but rather a function of many other characteristics associated with the outcome, and 2) "cumulative (dis)advantage", whereby exposure to neighborhood patterns other characteristics associated with the outcome over time. Time-varying characteristics may act as both *confounders* and *mediators* of the total neighborhood effect across the life-course.

Issues of selection bias compel researchers to control for many of these other time-varying characteristics in attempts to identify an unbiased "neighborhood effect" (Sampson 2012). However, in conventional regression models this simultaneously over-controls indirect pathways through which neighborhood exposure accumulates in other characteristics over time, potentially dramatically underestimating the total neighborhood effect (Kravitz-Wirtz 2016; Sharkey and Elwert 2011; Wodtke et al. 2011). Although the concept of cumulative (dis)advantage has long existed explicitly or implicitly in the quantitative literature, this temporal complexity is rarely operationalized in observational designs in ways that might elucidate the mechanisms of accumulation that reproduce inequalities in childhood development. This has been highlighted by scholars such as Cummins et al. (2007), who call for a more relational approach to estimating the effects of space and place. Slater (2013) argues that the fixation on identifying a "direct effect" of neighborhood disadvantage *net* of individual characteristics is problematic for two reasons. First, such research isolates a marginal effect that may be only a small part of the total process in which place is used to arrange and produce social inequality. Second, the explanation for *why* a direct effect might exist often relies on problematic causal narratives focused on individual behaviors (e.g. the "culture of poverty" hypothesis) (Slater 2013).

Acknowledging the failure of conventional regression models to meet these theoretical needs, Wodtke et al. (2011) and others have used marginal structural models to estimate the total effect of neighborhood disadvantage by accounting for these time-varying confounding pathways using inverse-probability-of-treatment weighting

(Kravitz-Wirtz 2016; Sharkey and Elwert 2011; Wodtke et al. 2011). The primary result is a significant total effect that is larger than the estimate from a conventional regression. The statistical distinction between these studies and those using conventional regression is important from a causal perspective and informs the debate on whether neighborhood context is relevant or not as a mechanism for explaining outcomes. However, as with the conventional analyses, these studies still fall under Sharkey & Faber's (2014) criticism of only informing the *dichotomous research question* of whether neighborhood is "important" or not (Sharkey and Faber 2014).

It is therefore necessary to develop a methodological framework that not only accounts for the important issue of time-varying confounding highlighted by these innovative analyses using marginal structural models, but also extends to estimation of the specific direct and indirect pathways through which neighborhood influences childhood development. As Sampson (2012) argues, the social mechanisms and dynamic processes accounting for neighborhood effects, while thoroughly theorized, have remained largely a "black box" in quantitative analyses (Sampson 2012). Distinguishing the relative magnitude of these pathways will provide theoretical insight into the most important components of this dynamic social process. I examine several broad mechanisms hypothesized to be both differentially distributed across levels of neighborhood disadvantage and to influence language development: family's material resources, access to quality schools, family structure, physical home environment, and collective efficacy.

Mechanisms connecting neighborhood disadvantage to child development

Material resources measured on an individual- or household-level, such as poverty, are perhaps the most commonly cited confounders in studies of neighborhood effects given concerns surrounding selection bias (e.g. increased poverty leads to increased "selection" into highly disadvantaged neighborhoods). However, neighborhood disadvantage also produces differential risk of poverty by structuring exposure to both harmful and beneficial systems (e.g. access to a diverse labor market and the availability of high-wage jobs). Gradients of child development across levels of household poverty have been consistently demonstrated, and are hypothesized to be driven by mechanisms linked to familial stress, spending on childcare, and resources for early learning (Duncan, Brooks-Gunn, and Klebanov 1994; Korenman, Miller, and Sjaastad 1995; Petterson and Albers 2001).

Institutional theories emphasize the lack of quality schools, daycare centers, and recreational areas in disadvantaged neighborhoods, all of which have been demonstrated to be positively associated with childhood development (Brooks-Gunn, Duncan, and Aber 1997; Small and Newman 2001; Wilson 1987). Public school funding in the United States is closely linked to local home values, and the socioeconomic composition of schools is largely determined by the composition of the school catchment area. Under-resourced schools with a large proportion of students from low-income families have been linked to lags in early development via lack of resources for effective teaching and in-class disruptions (Duncan and Murnane 2011; Kahlenberg 2001; Willms 2010).

A large body of research documents the association between family structure and instability and child development (Brown 2006; Carlson and Corcoran 2001; Cavanagh

and Huston 2006; Fomby and Cherlin 2007). Hypothesized mechanisms include familial stress and disruptions in schooling and childcare. As noted by Lee & McLanahan (2015), the question of whether this association is causal is difficult to adjudicate as various family structures differ in many observed and unobserved ways over time. While family structure and instability are often cited as a confounder in studies of neighborhood effects, it is less clear whether neighborhood disadvantage itself influences family instability over time. To the extent that neighborhood disadvantage influences family poverty, it may influence pressures and preferences around family formation (Cherlin et al. 2008; Edin and Kefalas 2011).

Environmental theories emphasize the role of physical conditions in disadvantaged neighborhoods, particularly the substandard quality of housing for families with children (Gielen et al. 2012). Unsafe physical conditions and overcrowding may be related to familial stress, child stress, and child physical health (Coley et al. 2013). Many studies focus on child health, for example the interactions between physical environmental factors and chronic stress in producing susceptibility to asthma (Suglia et al. 2010). Health issues, particularly in early life, may impede school progression and cognitive development.

Lastly, "collective efficacy" is defined as the degree to which a group of individuals feel connected and are confident in the willingness and ability of the group to act on behalf of its members. In terms of neighborhoods, Sampson applied a theory of collective efficacy in *Great American City* to describe cohesion and mutual trust among residents with shared expectations for intervening in support of neighborhood social control

(Sampson 2012). Sampson demonstrated how collective efficacy varies strongly by neighborhood and predicts many outcomes independent of racial composition and economic disadvantage. Collective efficacy has been hypothesized to be connected to health cognitive development in early childhood by influencing various forms of social control, the regulation of deviant behaviors, and fostering a sense of community (Ichikawa, Fujiwara, and Kawachi 2017; Sampson et al. 2002; Smith et al. 2011).

Current study: Identifying direct and indirect pathways of accumulation

The current study extends the literature on quantifying neighborhood effects by explicitly reframing characteristics related to material resources, exposure to institutions, family, environment, and collective efficacy as accumulations of lived experience. Building upon the counterfactual framework of marginal structural models with a life-course methodology, I estimate the direct and indirect effects of neighborhood disadvantage on development through early childhood using the mediational g-formula to address issues of time-varying confounding (Esposito 2019; Hernán and Robins 2019; Keil et al. 2014; Robins and Hernan 2009; Vangen-Lønne et al. 2018). Focusing on the gap between highly advantaged and disadvantaged neighborhoods that emerges during this sensitive stage of the life-course, I estimate the indirect effects of neighborhood that operate through several mediating variables closely related to the hypothesized mechanisms highlighted above. By identifying specific aspects of *why* neighborhoods shape development and the relative importance of these pathways, findings from this study shed new light on how social policies might be better structured to mitigate the emergence and divergence of place-based disparities in child well-being.

Data, Measures, and Analytic Strategy

Data

I use data from the Fragile Families & Child Wellbeing Study (FFCWS), a prospective cohort study that follows children from ages 0 to 15 across large cities within the United States. The sample is designed to be nationally representative of unmarried births in cities with populations greater than 200,000 (Reichman et al. 2001). Sampling is clustered within 75 hospitals across 20 cities at the time of each child's birth (McLanahan 2009). This dataset heavily oversamples poor families and non-Hispanic Black unmarried parents, who comprise 69% of the sample, compared to 32% of the unmarried national population (Reichman et al. 2001). The FFCWS collected data at birth and in the year each child reached ages 1, 3, 5, 9, and 15. I merge restricted-use data for contextual characteristics based on the Census tract of residence at each wave and school characteristics from the National Center for Educational Statistics. I only use data through age 9, as cognitive scores were not measured at age 15. I consider the full non-censored sample across these five waves, resulting in 19,997 person-years of observation (Appendix A provides detailed information on inclusion criteria).

Time-varying outcome

I measure children's early development using scores from the Peabody Picture Vocabulary Test-Revised (PPVT-R) recorded during the in-home assessments at ages 3, 5, and 9 (standardized to have a mean of 0 and standard deviation of 1). This test assesses the number and range of words that children understand on a continuous scale, and has

been used extensively as a comparable index of early cognitive development (Copp et al. 2018; Lee and McLanahan 2015).

Time-varying treatment

Based on the measure developed by Wodtke et al. (2011) and subsequently used in other studies (Douglas S. Massey et al. 2018; Wodtke and Parbst 2017), I calculate an index of neighborhood disadvantage based on a principal component analysis of Census tract-level characteristics: proportion of households living in poverty, proportion of households receiving welfare assistance, proportion of individuals unemployed, proportion of female-headed households, and proportion of individuals above age 25 without a high-school degree (Appendix B). As in Wodtke et al. (2011), I discretize the index into an ordinal variable consisting of five quintiles in order to allow for non-linear exposure-outcome and mediator-outcome effects. The exposure of interest is the duration-weighted quintile (i.e. the average quintile lived in up until the current time), representing cumulative exposure to neighborhood disadvantage (Wodtke et al. 2011).

Time-varying characteristics

I include two potential mediating variables related to material resources. Maternal poverty status (measured as 0/1 below the federal threshold) is collected in each wave and I included this as a duration-weighted exposure (i.e. the total years lived in poverty up until the current time). I also include mother's employment status (employed vs. unemployment).

To create a general measure of school quality, I use restricted data from the FFCWS on school characteristics in Years 5 and 9. I conduct a principal component analysis on the student-to-teacher ratio, percent of students receiving free lunches, and whether the school receives Title I funding. I use the first principal component as a proxy for general school quality (or the material resources available to a given school), where high values characterize schools with a low student-to-teacher ratio, a low percent of students receiving free lunches, and no Title I funding (Appendix B).

As a measure of family structure, I used a binary indicator for whether the child's mother is married/cohabitating vs. single. In sensitivity analyses I tested only using marital status and using the cumulative number of transitions to and from a co-residential (cohabitating or married) union, an operationalization used by Lee & McLanahan (2015) in their analysis of child development. Given little difference in substantive results across family structure variables, I report results using the simple binary indicator of mother married/cohabitating vs. single.

I measured physical housing environment using a binary indicator of positive response on at least one of two items reported by the FFCWS interviewers in Years 5 and 9: "Is environment inside home unsafe for young children? Answer 'Yes' is one or more potentially dangerous health or structural hazards (Examples: frayed electrical wires, mice or rats, broken glass, poisons, falling plaster, broken stairs, peeling paint, cleaning materials left out, flames and heat within reach of young children)"; "Is inside of home crowded? (Examples: many people living in a very small house or apartment, difficult to

find a private place to interview respondent, frequent interruptions and people bumping into each other."

Last, I include an instrument for measuring neighborhood collective efficacy adapted from Sampson et al. (1997) (Sampson, Raudenbush, and Earls 1997). I use the sum of nine identical items (each on a 4-point Likert scale of agreement) from Years 5 and 9. These include agreement on statements such as "This is a close-knit neighborhood" and "People around here are willing to help their neighbors." This total score is included as a continuous variable, with lower scores indicating a higher degree of collective efficacy.

Time-invariant confounders

Characteristics measured at baseline include parents' age (continuous), parents' nativity status (native-born, foreign-born), parents' race-ethnicity (non-Hispanic White, non-Hispanic black, Hispanic, other), parents' education (0/1 college attainment), whether the mother was living with her parents at age 15, child sex, and whether the child was born low birth weight (< 2,500 grams).

G-computation for total effects

G-computation is a method of standardization that allows for the estimation of unconfounded summary effects without requiring the assumption that these effects are constant across levels of confounders (i.e. no time-varying confounding), as is assumed in conventional regression models (Esposito 2019; A. Naimi et al. 2016; Robins and Hernan 2009; VanderWeele 2016; Vanderweele and Vansteelandt 2009; Wang and Arah

2015). Equation 1 expresses the population mean PPVT score, Y, standardized across all values of a stratifying variable, A.

$$E[Y] = \sum_{y}\sum_{x} y \cdot P(Y = y|A = a) \cdot P(A = a) \tag{1}$$

This generalized formula, or "g-formula," for the mean outcome at a given age can be extended over all stratifying variables, $\boldsymbol{V}$, which confound the association between A and Y, as well as variables which mediate the association, $\boldsymbol{M}$.

$$E[Y] = \sum_{y}\sum_{x}\sum_{m}\sum_{v}\begin{Bmatrix} y \cdot P(Y = y|A = a, \boldsymbol{M} = \boldsymbol{m}, \boldsymbol{V} = \boldsymbol{v}) \cdot \\ P(\boldsymbol{M} = \boldsymbol{m}|A = a, \boldsymbol{V} = \boldsymbol{v}) \cdot \\ P(\boldsymbol{V} = \boldsymbol{v}) \cdot \\ P(A = a) \end{Bmatrix} \tag{2}$$

In Equation 2, variables in $\boldsymbol{V}$ confound the exposure-outcome and mediator-outcome relationship. In Equation 3 we consider a specific mediator M which is causally dependent on a vector of previous mediators $\boldsymbol{L}$:

$$E[Y] = \sum_{y}\sum_{x}\sum_{m}\sum_{l}\sum_{v}\begin{Bmatrix} y \cdot P(Y = y|A = a, M = m, \boldsymbol{L} = \boldsymbol{l}, \boldsymbol{V} = \boldsymbol{v}) \cdot \\ P(M = m|A = a, \boldsymbol{L} = \boldsymbol{l}, \boldsymbol{V} = \boldsymbol{v}) \cdot \\ P(\boldsymbol{L} = \boldsymbol{l}|A = a, \boldsymbol{V} = \boldsymbol{v}) \cdot \\ P(\boldsymbol{V} = \boldsymbol{v}) \cdot \\ P(A = a) \end{Bmatrix} \tag{3}$$

As discussed above, variables in $\boldsymbol{L}$ are referred to as "time-varying confounders" when estimating total effects. When considering the effect on Y of changes to A via a specific mediator M, variables in $\boldsymbol{L}$ are referred to as "exposure-induced mediator-outcome confounders" because they are affected by the exposure and confound the relationship

between M and Y. The presence of such confounding means that we cannot estimate the counterfactual associated with a given value of A while holding $\boldsymbol{L}$ constant (as in conventional regression or matching estimators), because such a world would be impossible to observe. However, variables in $\boldsymbol{L}$ also themselves mediate the relationship between A and Y. Figure 2 describes this causal model visually using a directed-acyclic diagram.

This generalization of the entire conditional probability space over age is the critical contribution of g-formula standardization because it has important implications for estimating the population-level change associated with a counterfactual change in treatment. In conventional regression models or demographic decomposition (e.g., Das Gupta decomposition), estimates of counterfactual change associated with a given counterfactual are calculated under the assumption that no other conditional probabilities change as a result of the intervention (i.e., no time-varying confounding). In contrast, the g-formula makes explicit all cascades of conditional probabilities for all variables over age. In this way, we can estimate the change associated with a given counterfactual at a given point in time, conditional on how we expect that counterfactual change to affect all other variables over time (and how those in turn may affect values of the manipulated exposure variable itself over time). This *total effect* (TE) can be estimated using the g-formula in Equation 3, and is analogous to the total effect identified by the marginal structural model in previous studies of total neighborhood effects using observational data (Kravitz-Wirtz 2016; A. Naimi et al. 2016; Sharkey and Elwert 2011; Wodtke et al. 2011):

$$\text{TE} = E(Y_a) - E(Y_{a^*}) \quad (4)$$

G-computation for multiple mediation analysis

In mediation analysis we are often interested in estimating the natural direct effect (NDE) and natural indirect effect (NIE) (Robins and Greenland 1992). The NDE asks: What would have been the difference between the exposed group (a) and unexposed group (a^*) if they had both had the unexposed distribution of the mediator (M_{a^*})? In other words, we are isolating the portion of the total effect that was caused "directly" by the treatment (or unmeasured mediating pathways), rather than through the observed mediator, M. The NIE compares the treated group to what would have happened had the treated group had the same distribution of the mediator as the untreated group. In other words, we are isolating the portion of the total effect operating through the mediator, rather than directly from the treatment itself.

$$\text{NDE} = E\left(Y_{aM_{a^*}}\right) - E\left(Y_{a^*M_{a^*}}\right) \quad (5)$$

$$\text{NIE}^{(M)} = E\left(Y_{aM_a}\right) - E\left(Y_{aM_{a^*}}\right) \quad (6)$$

Identification of these mediating effects is based on four key assumptions: 1) no unobserved exposure-outcome confounding, 2) no unobserved exposure-mediator confounding, 3) no unobserved mediator-outcome confounding, and 4) no unobserved treatment-induced mediator-outcome confounding (Appendix D). As discussed above, this fourth assumption is virtually always violated using observational data to estimate the effect of a neighborhood exposure (Wodtke et al. 2011; Wodtke and Parbst 2017).

However, under sequential ignorability, a randomized interventional analogue can be estimated that only relies on the first three assumptions (Vanderweele, Vansteelandt, and Robins 2014). Let $\bar{G}_{\bar{a}}$ denote a random draw from the distribution of the mediator that would have been observed in the population if the treatment $\bar{A}$ had been set to $\bar{a}$. In the presence of a single time-varying mediator, the interventional direct and indirect effects can be defined using the g-formula or IPT weighting (Vanderweele et al. 2014).

$$rNDE = E(Y_{\bar{a}\bar{G}_{\bar{a}^*}}) \; - \; E(Y_{\bar{a}^*\bar{G}_{\bar{a}^*}}) \tag{7}$$

$$rNIE^{(M)} = E(Y_{\bar{a}\bar{G}_{\bar{a}}}) \; - \; E(Y_{\bar{a}\bar{G}_{\bar{a}^*}}) \tag{8}$$

In the presence of repeated observations of a time-varying exposure, mediators, and outcome (as in the present study), VanderWeele & Tchetgen Tchetgen (2017) define the rNDE and rNIE using what the authors term the "mediational g-formula," extended from the standard g-formula in Equation 3.

$$\begin{aligned} E\left(Y_{\bar{a}\bar{G}_{\bar{a}^*}^{(1)}\ldots\bar{a}\bar{G}_{\bar{a}^*}^{(n)}}\right) = &\sum_{\bar{m}_T}\sum_{\bar{l}_T} E\left(Y_T|\bar{a}_T,\bar{m}_T,\bar{l}_T,v\right)\prod_{t=1}^{T} P\left(\bar{l}_T\middle|\bar{a}_T,\bar{m}_{t-1},\bar{l}_{t-1}v\right)P(v) \\ &\times\sum_{\bar{l}_T^{\dagger}}\prod_{t=1}^{T}\prod_{k=1}^{n-1} P\left(m_t^{(k)}\middle|\bar{a}_t^*,\bar{m}_{t-1},\bar{l}_t^{\dagger},v\right)P\left(\bar{l}_t^{\dagger}\middle|\bar{a}_t^*,\bar{m}_{t-1},\bar{l}_{t-1}^{\dagger},v\right)P(v) \\ &\times\prod_{t=1}^{T} P\left(m_t^{(n)}\middle|\bar{a}_t,\bar{l}_t,\bar{m}_{t-1},m_t^{(1)},\ldots,m_t^{(n-1)},v\right) \end{aligned} \tag{9}$$

I refer to the rNDE and rNIE defined using the mediational g-formula in Equation 9 as the "cumulative rNDE" and "cumulative rNIE" for each mediator. In other words, how

does a time-varying exposure (e.g., neighborhood disadvantage) cumulatively affect an outcome (e.g., PPVT score) by influencing the distribution of a specific time-varying mediator (e.g., household poverty)?

Parametric estimation of the mediational g-formula

VanderWeele and Tchetgen Tchetgen (2017) describe an approach using IPT weighting of marginal structural models to estimate the mediational g-formula, but this approach can perform poorly and inefficiently with continuous exposures and mediators. Alternatively, the parametric g-formula can be extended to estimate Equation 8 using standard regression models, a computational approach based on simulations (Lin SH et al. 2017). I use this approach, which is achieved by specifying a regression model for the time-varying exposure, each mediator, and the outcome and estimating the mediational g-formulas (Equation 8) required for the rNDE and $rNIE^{(M)}$ described in Equations 6 and 7. This stochastic simulation process is conducted in R and described in Appendix C. For a comprehensive review of this approach, see Lin et al. (2017).

For the exposure (NH disadvantage quintile), each mediator, and the outcome (normalized PPVT score), I specify a survey-weighted generalized linear model:

$$y_{i,a} = \beta_0 + \boldsymbol{\beta_1 B}_i + \boldsymbol{\beta_2 X}_{i,a-1} + \varepsilon_{i,a}$$

Models are indexed by individual (i) and age (a), with $\boldsymbol{B}_i$ representing time-invariant confounders measured at birth and $\boldsymbol{X}_{i,a-1}$ representing lagged time-varying characteristics (including the lagged value of the dependent variable itself, meaning this can be considered a "value-added" framework) (Wodtke and Parbst 2017). Poverty status

and the NH disadvantage quintile are modeled directly at each time step, but the duration-weighted values up to each age are used in the model for child's PPVT score (Wodtke et al. 2011). Models parameterize $y_{i,a}$ with appropriate likelihoods (normal, binomial, ordinal) and link functions (identity, logit) given the structure of each dependent variable. All models account for complex survey design by including survey weights of mothers at baseline, and standard errors are clustered by individual (Lumley 2010, 2018).

Adjustment for item non-response and panel attrition

I create 30 multiply imputed datasets using chained equations to account for missing observations on study variables due to item-nonresponse. In addition, it is possible that attrition in the FFCWS is nonrandom and may bias effect estimates. In g-formula estimation, simulations begin with the full sample in the first period and every individual is simulated through all subsequent periods (A. Naimi et al. 2016; Robins and Hernan 2009). In effect calculations, I am then essentially including simulated values of all person-years that are unobserved in the survey sample. This is analogous to the approach used in marginal structural models to correct for panel attrition, where inverse-probability-of-treatment weights are "stabilized" by additionally adjusting for the differential probability of being observed in each wave (Robins, Hernán, and Brumback 2000; Sharkey and Elwert 2011).

Results

Sample summary statistics are presented in Table 1. There is a sharp contrast in children's PPVT scores across levels of neighborhood disadvantage (100.37 in the first

quantile, or NH 1, at Age 9 compared to 87.43 in NH 5). There is extreme neighborhood segregation across several time-invariant characteristics, particularly the proportion of mothers with a college degree (37% in NH 1 compared to 2% in NH 5) and race (59% white in NH1 compared to 3% white in NH 5). High disadvantage neighborhoods are characterized by large proportion of mothers in poverty, lower school quality, higher maternal unemployment, higher proportions of single motherhood, higher risk of a hazardous housing environment, and worse collective efficacy. There is a large jump in poverty from child Age 0 to Age 1 in NH 5, with poverty increasing from 55% to 66% for these new mothers.

In order to ensure that mediating relationships are not based on extrapolation out-of-sample, Figure 3 examines potential issues of common support in the FFCWS sample by considering the joint distribution of neighborhood disadvantage with each mediator. All characteristics have a sufficient number of person-years in the smaller strata. The sample is well-balanced across neighborhood-poverty strata, as this was part of the sampling design for the FFCWS, and across neighborhood-school and neighborhood-efficacy strata. The sparsest area of the neighborhood-mediator joint distributions is in the extreme cells for neighborhood-school strata (less than 1% of the sample falls in the highest quintile of neighborhood disadvantage *and* highest quintile of school quality, and vice versa). However, there are a sufficient number of person-years in all other cells of the joint distribution, and even the person-years in these extreme cells are similar to those in a similar analysis using the Panel Study of Income Dynamics (Wodtke and Parbst 2017).

Table 2 presents coefficients from the models for NH disadvantage, all mediators, and PPVT score. As discussed above, none of these estimates in isolation can be interpreted without assuming no time-varying confounding. Over this nine-year period of observation, neighborhood attainment itself is not strongly associated with any time-varying characteristics outside of past neighborhood attainment and time (i.e. child age). In the mediator models, higher neighborhood disadvantage is associated with increased risk of poverty, lower school quality, higher (worse) score on the collective efficacy index, and higher risk of a hazardous physical home environment. It is only weakly associated with an increased risk for unemployment and single motherhood. The fully adjusted models for PPVT scores at each age suggest that scores are positively associated with mother's education and school quality, and negatively associated with cumulative exposure to poverty and to a lesser extent neighborhood disadvantage.

Figure 4 illustrates the means and confidence intervals for each time-varying mediator estimated by using the coefficients from Table 2 in the natural course g-formula simulation. I overlay the observed survey-weighted means to illustrate concordance between the empirical cohort and the synthetic g-formula cohort at each age. The means and confidence intervals from the g-formula simulation closely match the observed data for all characteristics, indicating that the g-formula is well-calibrated and suggesting valid incorporation of survey weights throughout estimation (Appendix C). We also observe the sharp divergence across the counterfactual g-formula simulations for neighborhood Quintile 1 (low disadvantage) and Quintile 5 (high disadvantage) in terms of the proportion of families expected to be living in poverty, the expected school quality, the

expected physical environment, and the expected collective efficacy. Neighborhood context influences the likelihood of single motherhood and unemployment to a lesser extent.

Table 3 shows the total, direct, and indirect cumulative effects of living in each quintile of neighborhood disadvantage over this period compared to living in the most disadvantaged neighborhood (5^{th} Quintile) on normalized PPVT score at Age 9. This is juxtaposed with the total effect estimated from both the adjusted conventional regression (Table 2) and an unadjusted model only controlling for age and neighborhood disadvantage. As expected, the total effect estimated by the g-formula (0.27 standard deviations for Quintile 1, $p < 0.05$) is smaller than the unadjusted effect (1.23, $p < 0.001$) but larger than the fully-adjusted effect (-0.01, $p > 0.05$).

In the effect decomposition, the direct effect represents the proportion of the total effect that is not explained by the observed mediators. Focusing on the comparison between living in the lowest quintile of neighborhood disadvantage (Quintile 1) vs. the highest quintile (Quintile 5), the direct effect accounts for 19% of the total cumulative effect of neighborhood disadvantage on child's PPVT score by Age 9, while the time-varying mediators explain 81%. There is a large indirect effect via increasing risk of household poverty, 0.15 ($p < 0.001$), which mediates 56% of the total effect. Conceptually, this indirect effect can be interpreted as: "What if those living in Quintile 5 experienced the same household poverty levels across these nine years as would be expected if they lived in Quintile 1, but all other mediators unfolded as expected in Quintile 5?" In other words, we are isolating the proportion of the effect of living in

Quintile 1 compared to Quintile 5 that is solely due to the expected changes in household poverty across these nine years, keeping all other variables at the values they would have had in Quintile 5. This large indirect pathway emerges due to the significant associations between neighborhood and household poverty across the early life course *and* the significant associations between household poverty and PPVT scores. In contrast, though neighborhood strongly influences other mediator distributions such as environment and collective efficacy, these have negligible associations with PPVT scores conditional on all other mediators and time-invariant controls. While poverty represents a significant indirect effect pathway across all neighborhood quintile comparisons, the only other significant indirect pathway is through school quality in comparing the two extreme neighborhood quintiles, 0.05 ($p < 0.05$; 17% of total effect).

Table 1 illustrates the extreme level of racial residential segregation within this cohort, and others such as Wodtke et al. (2011) and Sharkey et al. (2011) have demonstrated racialized heterogeneity in the total effects of neighborhood disadvantage on development and academic outcomes. To explore potential racialized variation in these indirect effect pathways, I run the same analysis (fit models, g-computation, effect decomposition) fully stratified by child race (nonwhite vs. white). Table 4 reports these effect decompositions, which are illustrated in Figure 5. For the children of non-white parents, I find a similar pattern as in the pooled analysis: 67% of the total effect of neighborhood disadvantage on PPVT scores operates via increasing household poverty and decreasing exposure to high quality schools. I find no evidence of indirect effect pathways through any of the measured mediators for the white population. Across all

analyses, I find no significant mediating effects through the hypothesized pathways of unemployment, physical home environment, single motherhood, or collective efficacy.

Discussion

Previous longitudinal analyses have documented the difference between conventional estimates of neighborhood effects using observational data and estimates that account for time-varying confounding (Kravitz-Wirtz 2016; Sharkey and Elwert 2011; Wodtke et al. 2011). The present study extends this line of research by unpacking the "black-box" of the total neighborhood effect, examining the indirect pathways through which neighborhood disadvantage patterns childhood development measured via PPVT scores. These time-varying dynamics are adjusted for within the inverse-probability-of-treatment weighting used in previous studies, but their relative mediating effects are obscured.

I leverage variables representing a diverse set of theoretical mechanisms linking place-based disadvantage to childhood development and parameterize their entangled accumulation over the first nine years of life. Taken together, these findings suggest that the total neighborhood effect on healthy development in early life for this cohort is driven by 1) indirect effects in distributing risk of poverty and access to quality schools and 2) stronger effects for those racialized as non-white. I primarily find support for material resource theories suggesting that neighborhood disadvantage is most important for child development to the extent that it is strongly implicated in the dynamic arrangement and production of household poverty. Mothers are much more likely to fall into poverty if they live in a highly disadvantaged neighborhood, especially in the year immediately

following birth. This is consistent with other studies that link maternal poverty to early life cognitive development (Lee and Jackson 2017; Lee and McLanahan 2015). By Age 9, access to high quality schools also has a significant indirect influence on children's cognitive development, though this represents a relatively small proportion of the total neighborhood effect (17%). This is consistent with Wodtke & Parbst (2017), who use a similar approach to identify a negligible indirect effect of school poverty on academic achievement. Indirect effects emerging so early in childhood are likely to compound exponentially over time, especially as maternal and school resources become more salient in children's lives and may begin to influence other developmental factors in adolescence. On top of being exposed to neighborhood disadvantage at extremely disproportionate levels, I also find suggestive evidence that these indirect effect pathways are most pronounced for non-white populations.

These findings are important for research on neighborhood effects given that household poverty is perhaps the most ubiquitous control variable used to address selection bias in conventional regression models. The present study clarifies that what is often described as *the* neighborhood effect is typically a *direct* effect, where cumulative mediating pathways via individual- or household-level variables such as poverty are conditioned out. As I demonstrate in this study, the g-formula approach can control for the confounding (selection) effects of household poverty while not simultaneously over-controlling its importance as a mediator over time. I use this framework to quantitatively demonstrate how the neighborhood effect is a cumulative combination of both neighborhood-level exposures (collective efficacy, environment, schools) and entangled

household-level exposures (poverty, employment, family). My findings here regarding poverty and school quality support arguments made by Slater (2013) and others: the majority of "the neighborhood effect" is likely best captured by the structural processes segregating poor, racialized populations into disadvantaged spaces which then organize the production of household poverty and access to beneficial public institutions such as high quality schools (Cummins et al. 2007; Reskin 2012; Slater 2013). While centered in many qualitative studies, these dynamics are often missed in quantitative analyses that condition on many individual- and household-level material resource variables. This framework is also in contrast to mediating explanations involving more amorphous neighborhood effects operating through family formation, social cohesion or collective efficacy, which combined with extreme racial segregation often implicate harmful racist stereotypes (Slater 2013).

The present study reflects broadly on how to align quantitative methodology more closely with theory on cumulative (dis)advantage and life-course processes. Complex temporal dynamics are frequently implied and highlighted in the conceptual frameworks of sociological studies, but the quantitative analyses that follow are often limited in their ability to account for such dynamics. In the analysis of social processes, it is very common for both the dependent and independent variables to vary temporally across the lives of individuals. These variables inevitably become entangled with other time-varying characteristics that are correlated with the dependent variable, resulting in dynamic mediation and confounding pathways across the life-course. In identifying effects within a longitudinal perspective, it thus becomes untenable, both empirically and theoretically,

to ignore the risks of time-varying confounding in our quantitative approaches. Sociologists have begun to make this case within subfields of stratification research such as neighborhood, family, socioeconomic, and health effects in early life by employing marginal structural models (Lee and Jackson 2017; Lee and McLanahan 2015; Sharkey and Elwert 2011; Sharkey and Faber 2014). This study acknowledges and contributes to this endeavor by demonstrating a method to overcome the same issues while also decomposing direct and indirect effect pathways through mediating variables.

Limitations

While combining observational data and g-methods can allow for a more formal causal interpretation than conventional regression, they do not replace the need for carefully designed experiments when possible. The total effect estimated in the g-formula is only unbiased and consistent under the usual observational assumption of no unobserved treatment-outcome confounding, but requires the additional assumptions of no unobserved *treatment-mediator* or *mediator-outcome* confounding (VanderWeele 2016; Wodtke and Parbst 2017). These assumptions are all illustrated in Appendix D. Despite these very strong assumptions which caution against a causal interpretation of the direct and indirect effects, I believe there are substantive conceptual benefits for using these methods to describe and decompose total effects estimated using observational data. If we theorize complex effect pathways in a life-course perspective, it is necessary to employ an appropriate methodological framework that is explicit about all assumptions. Applied to observational data, the g-formula at best provides a causal estimator under the above assumptions. At worst, it provides a more nuanced decomposition of associational

life-course processes that can more effectively reflect theory surrounding accumulation and how disparities between groups develop dynamically over time. Regardless of study design, the g-formula mediation highlights the need for tracking time-varying mechanisms that might explain *why* a long-term total effect emerges. This focus can aid future research in avoiding the trappings of Sharkey & Faber's "dichotomous research question" (Sharkey and Faber 2014).

The mediating variables included here are likely not comprehensive of all possible causal mechanisms connecting neighborhood disadvantage to early childhood development. They are also relatively crude in terms of identifying the specific mechanisms underlying theoretical connections. For example, I demonstrate how a significant portion of the total neighborhood effect manifests through indirectly patterning risk of poverty, which in turn impacts development. Theoretically each of these pathways can be further decomposed (i.e. *how* neighborhood produces household poverty, and *how* household poverty influences development), where we can draw on an extensive development literature in psychology, education, and pediatrics (Blair and Raver 2016; Kramer et al. 2017). Qualitative analyses are also invaluable in parsing the nuance of these high-level direct and indirect pathways. The primary goal of this quantitative analysis is to clarify how a total neighborhood effect can be considered as the cumulative product of several high-level mechanisms over time, while avoiding the pitfalls of conventional regression models which assume these mechanisms and neighborhood are not reciprocally related.

Last, there are several important limitations to this analysis regarding construct validity. It is possible that the outcome variable (PPVT score) is an inappropriate operationalization of the target theoretical construct, healthy cognitive development in early childhood. There has been mixed evidence as to whether the PPVT instrument contains racial, class, or gender biases (Halpin, Simpson, and Martin 1990; Pichette, Béland, and Leśniewska 2019). However, this instrument has been used extensively as an outcome measure of language and vocabulary development, particularly in studies of neighborhood effects. The same is true of a basic household poverty indicator, despite the fact that our theoretical construct of resource deprivation often involves a more nuanced connection between urban space and high-wage jobs, tools of wealth accumulation, etc. The primary goal of this study is to apply a rigorous causal mediation analysis using these common variables to 1) demonstrate how the total neighborhood effect is often underestimated when over-controlling for time-varying characteristics, 2) clarify that the majority of the total effect operates through household poverty rather than other commonly cited mediators, and 3) this mediation process is itself highly racialized.

Future directions

Here I only examine accumulation and selection forces that develop during the period of observation, accounting for non-random selection at baseline (i.e. birth of the focal child) and time-varying confounding throughout. However, selection into neighborhood occurs before this baseline and is patterned by a complex array of social, structural, and historical forces (Sharkey 2013; Sharkey and Elwert 2011). Characterizing selection as a social process in *Great American City*, Sampson notes: “My ultimate argument is thus

that selection is not a 'bias' but rather part and parcel of a dynamic social process – another form of neighborhood effect" (Sampson 2012). This perspective will be especially important in studying how *differences in rates of exposure* to neighborhood disadvantage, for example by race and class, manifest as raced and classed disparities in outcomes. Educational attainment and racialized processes are associated with child development and are also substantial drivers of neighborhood exposure via residential segregation (Krysan and Crowder 2017; Lareau 2011; Massey and Denton 1993). Maternal college attainment and Black racial identity are the strongest observed predictors of the distribution of neighborhood disadvantage at baseline (Table 2), and just 10% of the white population live in the worst two quintiles of neighborhood disadvantage compared to 62% of the Black population. While I demonstrate that the effect pathways between neighborhood disadvantage, mediators, and child development are racialized, it is likely that the unequal distribution of exposure to disadvantage plays a much larger role in population disparities. Jackson & VanderWeele (2018) demonstrate how the g-formula can be used to decompose the impact of *exposure differences* at baseline on disparities in later-life outcomes (Jackson and VanderWeele 2018). This is a promising avenue for future research to disentangle the impact of selection itself in producing long-term racialized and classed disparities.

Conclusions

Many variables typically operationalized as separable characteristics of the individual at a given point in time must instead be studied as dynamic, relational, and entangled social processes. An important theoretical advantage of the g-formula is reframing point-

in-time, individual characteristics as accumulations of lived experiences: individual histories of treatment exposure, confounders, and their reciprocal relationships. This methodology is much more closely aligned with life-course theory in many sociological domains, in contrast to conventional regression which treat characteristics as separable within a given cross-section. Controlling for socioeconomic, family, or school characteristics in a conventional model ignores the ways in which these characteristics influence (and are influenced by) neighborhood context over time. While the triangulation of evidence from the current study and many others on the total neighborhood effect supports the need for sustained place-based urban investment, this approach can also pinpoint specific pathways for mitigation of the intergenerational transmission of neighborhood disadvantage (e.g. decoupling neighborhood from the systems that produce household poverty and access to quality schools may be significantly more impactful than intervening on alternative mechanisms).

To the extent that time-varying confounding exists across the life course, conventional models are chronically underestimating the cumulative total and indirect effects of early life disparities. This paper provides an applied example of one method for disentangling the total effect of neighborhood disadvantage on development in very early life, building on the compelling contributions of Sharkey (2011), Wodtke (2011) and others. This flexible and generalizable method can be applied widely for studying the mechanisms through which many early-life disparities are manifested as later-life disparities, an ongoing discourse across many topics in social stratification and inequality.

Tables and Figures

Table 1: Unweighted sample statistics across five waves (at focal child age 0, 1, 3, 5, 9), stratified by the lowest and highest quintiles of neighborhood disadvantage (NH1 and NH5, respectively).

	Age 0		Age 1		Age 3		Age 5		Age 9	
	NH 1	NH 5	NH 1	NH 5	NH 1	NH 5	NH 1	NH 5	NH 1	NH 5
Time-invariant characteristics										
Child male	0.53	0.49	0.55	0.50	0.52	0.53	0.55	0.53	0.54	0.49
Child low birth weight	0.08	0.12	0.08	0.13	0.07	0.12	0.07	0.11	0.06	0.13
Father age	30.14	27.02	30.18	26.69	30.10	26.55	29.80	26.87	29.51	27.22
Father immigrant status	0.16	0.14	0.16	0.14	0.16	0.13	0.16	0.13	0.15	0.12
Mother age	27.79	24.17	27.77	24.01	27.57	23.89	27.35	24.10	27.05	24.80
Mother immigrant status	0.14	0.12	0.15	0.11	0.14	0.11	0.15	0.10	0.14	0.09
Mother college	0.37	0.02	0.36	0.01	0.36	0.01	0.33	0.01	0.29	0.01
Mother lived with parents	0.62	0.30	0.60	0.32	0.59	0.32	0.57	0.31	0.56	0.32
Mother White	0.59	0.03	0.60	0.04	0.59	0.03	0.53	0.03	0.48	0.02
Mother Black	0.16	0.73	0.15	0.71	0.18	0.73	0.21	0.73	0.27	0.75
Mother Hispanic	0.18	0.22	0.17	0.23	0.17	0.22	0.19	0.23	0.19	0.20
Mother other race	0.08	0.02	0.07	0.02	0.06	0.02	0.07	0.01	0.06	0.02

	Age 0		Age 1		Age 3		Age 5		Age 9	
	NH 1	NH 5	NH 1	NH 5	NH 1	NH 5	NH 1	NH 5	NH 1	NH 5
Time-varying mediators										
Mother poverty	0.12	0.55	0.16	0.66	0.15	0.64	0.18	0.63	0.16	0.56
School quality	--	--	--	--	--	--	0.70	-0.50	0.68	-0.64
Mother unemployed	0.39	0.56	0.41	0.57	0.38	0.52	0.34	0.48	0.32	0.45
Mother married/cohab	0.79	0.47	0.75	0.45	0.71	0.37	0.65	0.32	0.60	0.27
Physical environment	--	--	--	--	0.10	0.34	0.12	0.36	0.09	0.28
Collective efficacy	--	--	--	--	--	--	14.95	18.98	14.44	17.70
Time-varying outcome										
PPVT Score	--	--	--	--	95.38	80.69	101.99	87.44	100.37	87.43
N (unique individuals)	4778		4261		3912		3761		3285	
Percent	100		89		82		79		69	

Note: Values represent means for continuous variables and proportions for binary variables. Missing values from item non-response are completed via multiple imputation with chained equations.

Table 2. Coefficient estimates from generalized linear models for the exposure (NH Quintile), outcome (PPVT score), and each time-varying mediator.

	NH Quintile	Poverty	Unemployed	Mother married/ cohab.	Environment	Collective efficacy	School quality	PPVT (Age 3)	PPVT (Age 5)	PPVT (Age 9)
Intercept	--	-2.07***	-0.98***	-1.61***	-1.90***	16.27***	0.52*	0.29	0.61**	0.46*
Child male	-0.04*	0.07	-0.11	0.23*	0.08	-0.30	0.00	-0.14	-0.05	0.00
Child low birth weight	0.10**	0.03	0.16	-0.32	0.42	-0.20	-0.24*	-0.17	-0.06	-0.15
Father age	-0.01**	0.00	0.01	-0.01	0.01	0.02	0.00	0.01	0.01	0.01
Father immigrant status	0.34***	0.38*	0.08	0.20	0.04	0.35	0.32	-0.21	-0.14	0.13
Mother age	0.00	-0.01	-0.02	0.05**	-0.02	-0.07	0.01	0.00	0.00	-0.01
Mother immigrant status	-0.22***	0.07	-0.04	0.46*	-0.03	0.61	-0.33	-0.15	-0.20	0.03
Mother college	-0.88***	-1.58***	-0.12	0.73***	-1.32***	-0.66	0.31	0.52***	0.24	0.39*
Mother living with parents at age 15	-0.03	-0.10	-0.09	0.39***	-0.10	-0.67*	0.10	-0.02	-0.10	-0.06
Mother Black	0.95***	0.42*	-0.24*	-0.70***	0.09	0.21	-0.66***	-0.31**	-0.32*	-0.28
Mother Hispanic	0.73***	0.41*	-0.25	-0.23	-0.17	0.59	-0.67**	-0.29*	-0.31	-0.25
Mother other race	0.45***	-0.07	-0.29	0.12	-0.25	2.27*	-0.18	-0.23	0.09	-0.34
Child age 3	-0.07	-0.24	-0.21	-0.30*	--	--	--	--	--	--
Child age 5	-0.21***	0.07	-0.42**	-0.93***	0.03	--	--	--	--	--
Child age 9	-0.35***	-0.28	-0.42**	-0.81***	-0.05	-0.76**	-0.04	--	--	--

	NH Quintile	Poverty	Unemployed	Mother married/ cohab.	Environment	Collective efficacy	School quality	PPVT (Age 3)	PPVT (Age 5)	PPVT (Age 9)
Poverty	0.19***	1.49***	0.26	-0.55***	--	--	--	-0.33*	-0.43**	-0.47***
School Quality	--	--	--	--	--	--	--	--	-0.01	0.08
Unemployed	0.19***	0.94***	2.49***	0.28*	--	--	--	-0.05	-0.05	0.01
Mother married/cohab	-0.07**	-0.62***	0.12	3.13***	--	--	--	-0.09	0.09	0.05
Physical environment	--	--	--	--	--	--	--	--	-0.07	0.01
Collective efficacy	--	--	--	--	--	--	--	--	--	-0.01
NH duration-weighted	--	0.32***	--	--	--	--	--	-0.10*	-0.05	0.00
NH 2nd	2.35***	--	-0.09	-0.47**	0.49	0.43	-0.25	--	--	--
NH 3rd	3.51***	--	0.01	-0.13	0.68**	1.98***	-0.50**	--	--	--
NH 4th	4.71***	--	-0.02	-0.27	0.84**	1.84***	-0.33	--	--	--
NH 5th	6.68***	--	0.23	-0.21	0.94***	2.83***	-0.55***	--	--	--
PPVT Score	--	--	--	--	--	--	--	--	0.18***	0.26**

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

NH Quintile refers to relative neighborhood disadvantage, where 5 corresponds to the quintile with the highest disadvantage.

Time-varying characteristics are lagged in all models, and all models include lagged values of the outcome as a predictor.

Age is included as a categorical variable in all models. The reference category for the PPVT model is Age 3, as PPVT score was not collected until Age 3. The reference category for all other models, where data collection began at birth, is Age 1.

Table 3. Natural direct and indirect effects contributing to the total effect of each neighborhood quintile compared to the reference quintile (Quintile 5; highest neighborhood disadvantage). All estimates are in terms of standard deviations in PPVT scores at Age 9.

	Quintile 1			**Quintile 2**			**Quintile 3**			**Quintile 4**		
	Coef.	SE	%	Coef.	SE	%	Coef.	SE	%	Coef.	SE	%
Unadjusted effect	1.23	(0.13)***		0.93	(0.10)***		0.62	(0.06)***		0.31	(0.03)***	
Adjusted effect	-0.01	(0.21)		-0.01	(0.16)		-0.01	(0.11)		0.00	(0.05)	
G-formula decomposition												
Total effect (TE)	0.27	(0.19)*	100	0.19	(0.15)	100	0.12	(0.10)*	100	0.08	(0.05)	100
rNDE	0.05	(0.21)	19	0.04	(0.16)	20	0.02	(0.10)	21	0.01	(0.05)	16
rNIE Poverty	0.15	(0.04)***	56	0.12	(0.03)***	63	0.08	(0.02)***	70	0.04	(0.01)***	53
rNIE School quality	0.05	(0.03)*	17	0.02	(0.02)	13	0.00	(0.01)	4	0.02	(0.02)	23
rNIE Unemployed	0.00	(0.00)	0	0.00	(0.00)	0	0.00	(0.00)	0	0.00	(0.00)	0
rNIE Environment	0.00	(0.02)	0	0.00	(0.01)	0	0.00	(0.01)	0	0.00	(0.01)	0
rNIE Married/cohabitating	0.01	(0.01)	3	0.00	(0.01)	-1	0.00	(0.01)	1	0.00	(0.01)	2
rNIE Collective efficacy	0.01	(0.02)	5	0.01	(0.02)	6	0.00	(0.01)	4	0.00	(0.01)	6

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

The unadjusted model includes only age and duration-weighted neighborhood disadvantage.

The adjusted model includes all time variant and invariant characteristics.

Table 4. Natural direct and indirect effects contributing to the total effect of the first neighborhood quintile (lowest disadvantage) compared to the reference quintile (Quintile 5; highest disadvantage) stratified by white and non-white. All estimates are in terms of standard deviations in PPVT scores at Age 9.

	Quintile 1 (non-white)			**Quintile 1 (white)**		
	Coef.	SE	%	Coef.	SE	%
Unadjusted effect	0.78	(0.19)***		0.29	(0.39)	
Adjusted effect	0.00	(0.18)		-0.02	(0.28)	
G-formula decomposition						
Total effect (TE)	0.23	(0.17)*	100	0.13	(0.28)	100
rNDE	0.05	(0.19)	21	0.12	(0.29)	89
rNIE Poverty	0.11	(0.03)*	47	0.00	(0.07)	-1
rNIE School quality	0.05	(0.02)*	20	0.01	(0.03)	4
rNIE Unemployed	0.00	(0.01)	0	0.00	(0.02)	0
rNIE Environment	0.01	(0.02)	3	-0.01	(0.04)	-5
rNIE Married/cohabitating	0.00	(0.01)	2	0.00	(0.01)	3
rNIE Collective efficacy	0.02	(0.02)	9	0.01	(0.04)	11

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

The unadjusted model includes only age and duration-weighted neighborhood disadvantage.

The adjusted model includes all time variant and invariant characteristics.

Figure 1. Model of the temporal, reciprocal relationships between neighborhood context (NH) and time-varying individual/household characteristics (M) in influencing PPVT scores (P) at a specific time. The dotted arrow illustrates how M *confounds* the relationship between NH and P ("selection bias") at one time point. The thick solid arrow illustrates how M *mediates* the relationship between NH and P at a subsequent time point ("cumulative (dis)advantage").

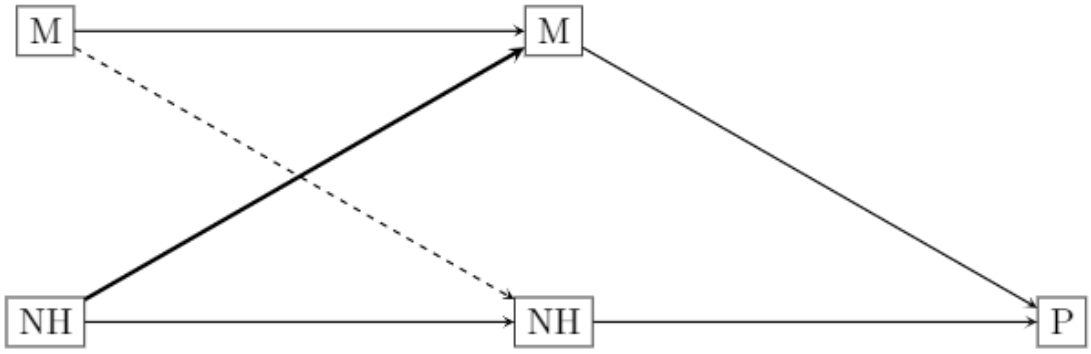


Figure 2: Conceptual life-course diagram representing all relationships in the directed-acyclic graph (DAG) to be estimated via the parametric g-formula. For visual clarity, all time-varying characteristics are suppressed to one node (M) for a given age and all time-invariant characteristics suppressed entirely (but controlled for in all connections). The heavy arrow represents a *direct effect pathway* of neighborhood disadvantage on PPVT score. The dotted arrow represents an *indirect effect pathway* through which neighborhood disadvantage influences the probability of mothers falling into poverty, which then indirectly affects both PPVT scores and neighborhood disadvantage.

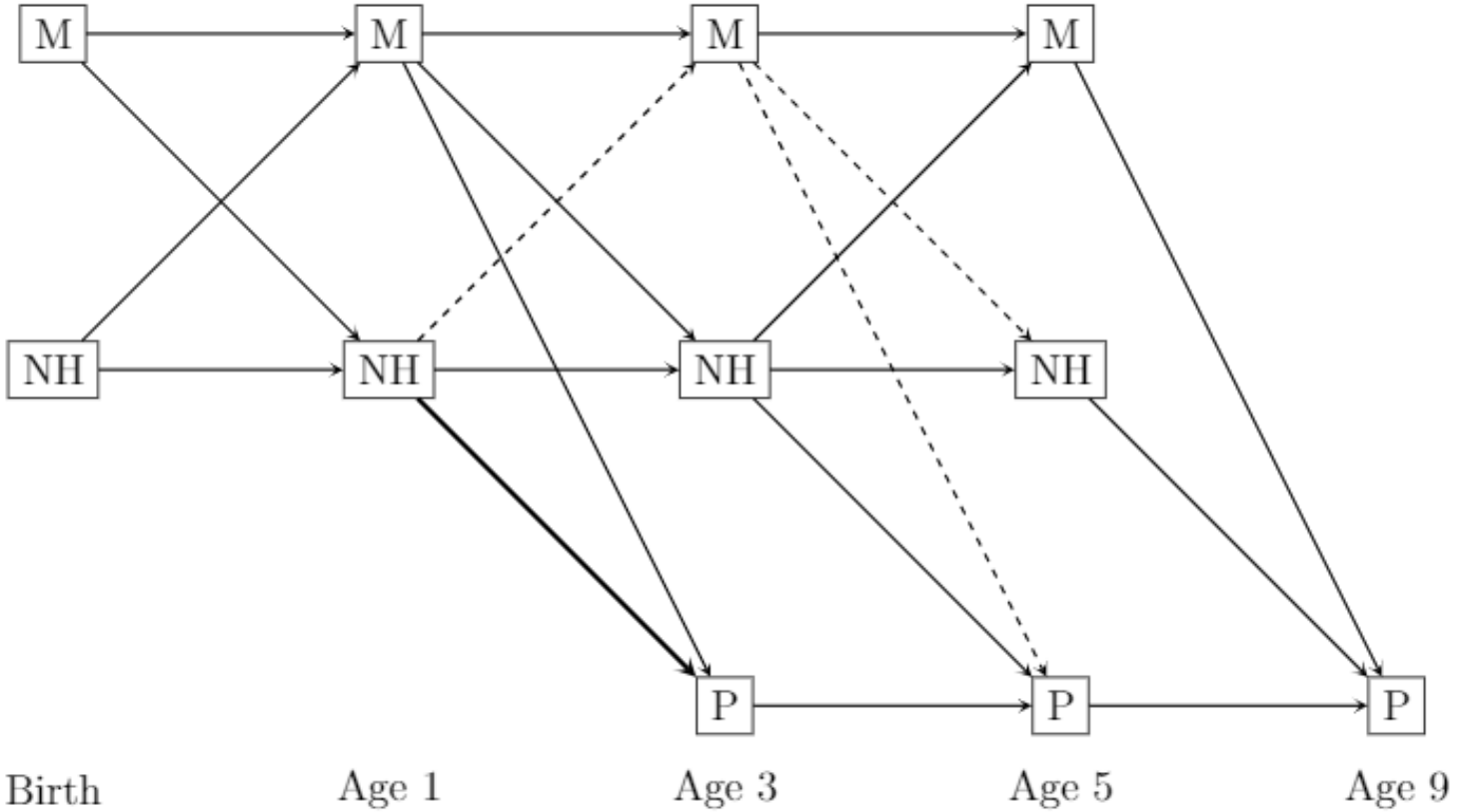


Figure 3. Joint treatment-mediator distributions based on five quintiles of neighborhood disadvantage and all potential mediating characteristics (unweighted data pooled over 30 imputations). Within each cell, the top number indicates the global proportion of person-years in that treatment-mediator combination and the bottom number indicates the total person-years. Continuous school quality scores and collective efficacy scores are categorized using quintiles for this purpose.

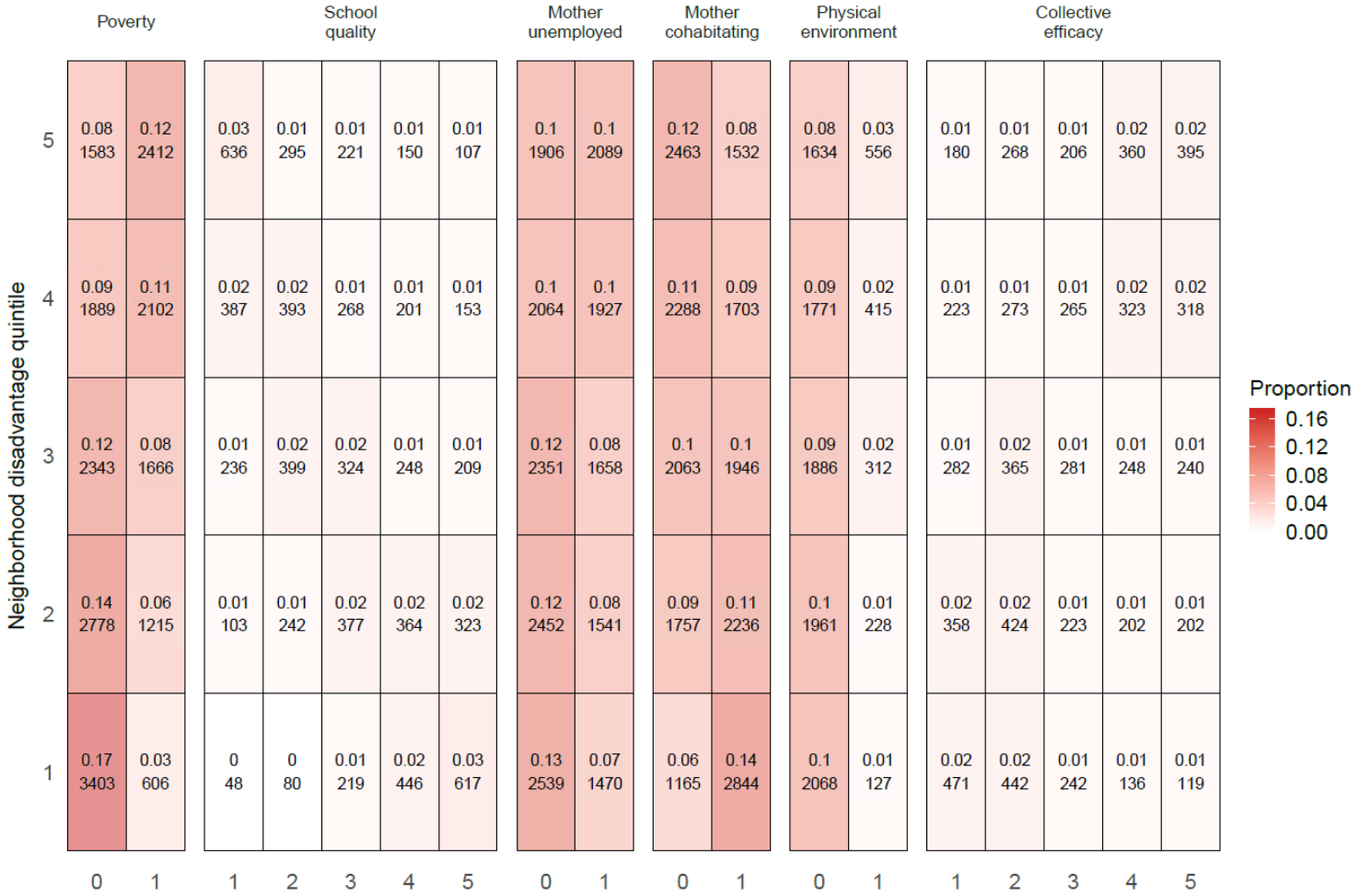


Figure 4. Population averages (proportions or means) and confidence intervals (dashed lines) for all time-varying mediators predicted in the *natural course* g-formula simulation, in addition to the high and low neighborhood disadvantage counterfactual course simulations (confidence intervals on these courses suppressed for visual clarity). The survey-weighted means and confidence intervals of the observed data are plotted at each age.

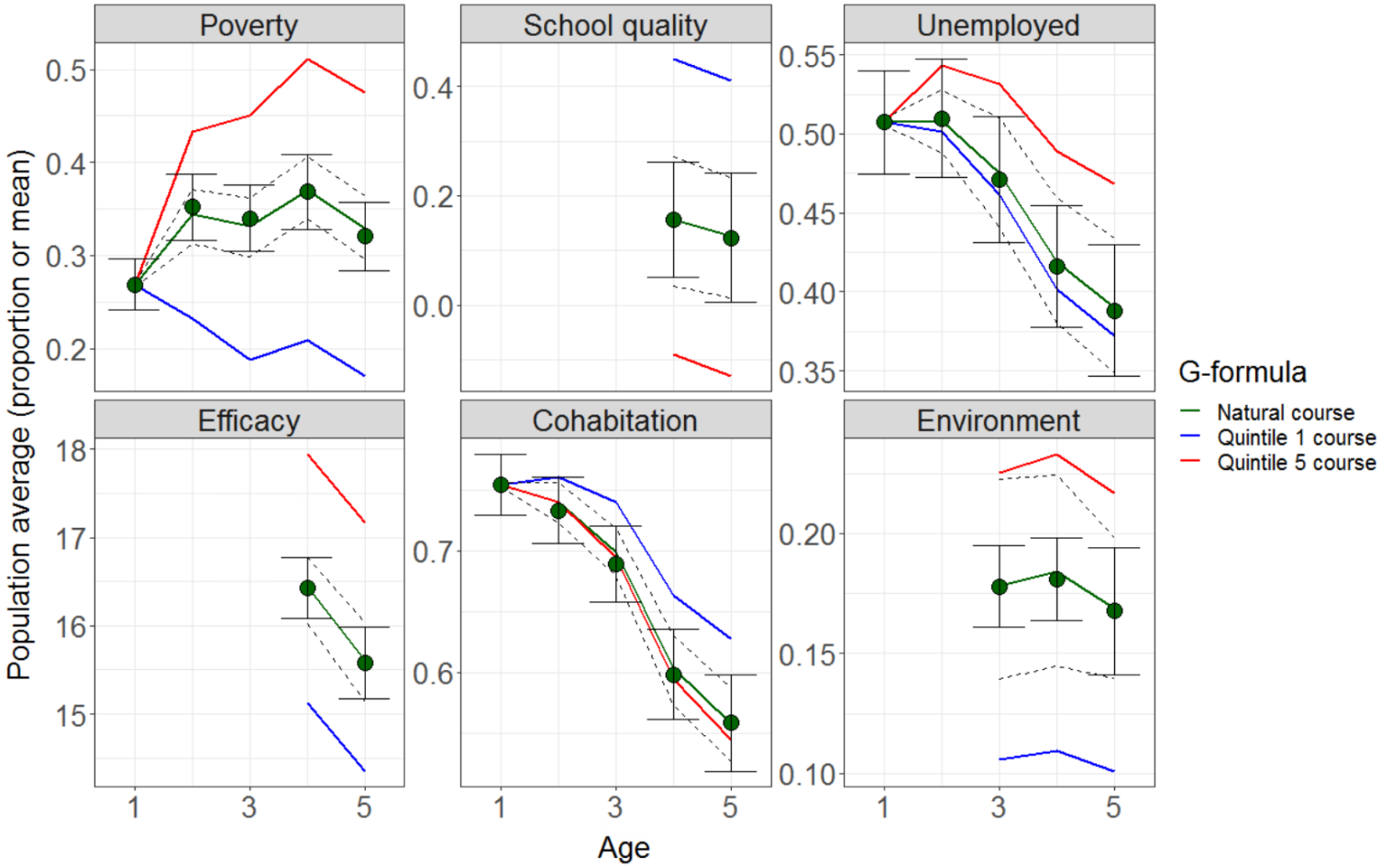
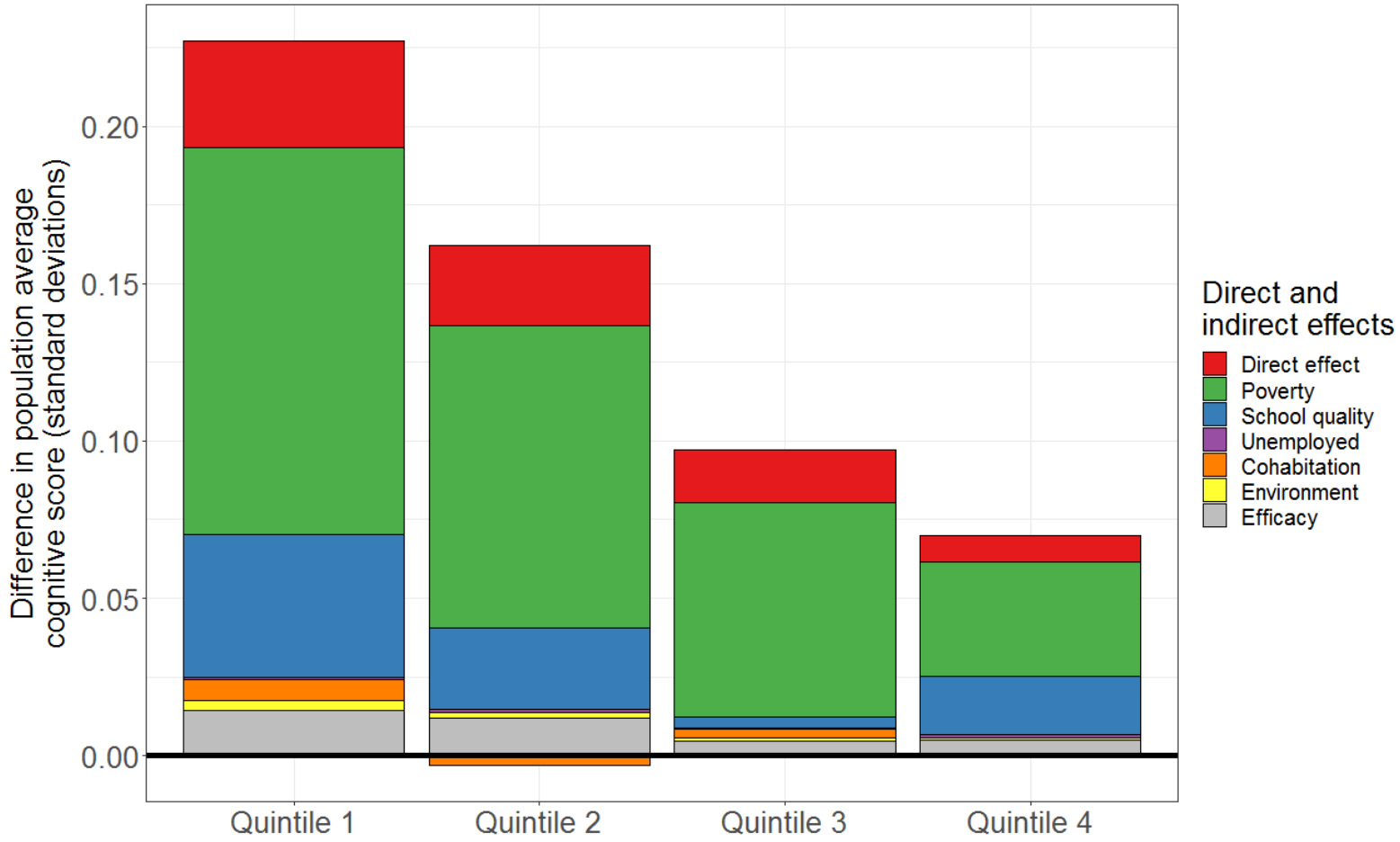
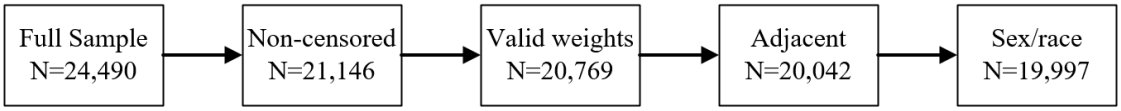


Figure 5. Decomposition of total effect (difference in PPVT score standard deviations between *high disadvantage course*, the 5th quintile of neighborhood disadvantage, and each of the four other quintiles). The "direct effect" is the residual difference not explaining by the observed mediating mechanisms, while all other time-varying characteristics are the indirect effects (i.e. the effect of living in a Quintile 1 neighborhood that operates indirectly on PPVT scores through reducing risk of maternal poverty).



Appendix

Appendix A. Notes on sample construction and multiple imputation



The Fragile Families and Child Wellbeing Study used a stratified random sample of all US cities with 200,000 or more people. The stratification was designed according to policy environments and labor market conditions in the different cities (Reichman et al. 2001). The study recruited an initial sample of 4,898 mothers (for a maximum sample size of 24,490 person-years over the 5 waves from child age 0 to 9), with two sets of baseline weights designed to produce a nationally representative sample and a sample representative of the 77 sample cities. I decided to use the city-level weights, as this produced a larger sample. After excluding those lost to follow-up or censoring and those without a valid city-level survey weight, I also excluded those who were not surveyed in adjacent waves. For example, a mother may be surveyed at child age 1, not surveyed at child age 3, but then surveyed again at child age 5. Because I include the lagged values of all predictor variables, I am not able to use the data on this mother collected at child age 5. Lastly, I exclude those few mothers without valid data on sex or race at baseline. This produces my final analytic sample consisting of 19,997 person-years across all focal children ages 0 to 9.

For missing information due to item non-response, I use a multiple imputation procedure via chained equations (Buuren and Groothuis-Oudshoorn 2011; Robitzsch,

Grund, and Henke 2019). This sort of procedure has been implemented in other studies using these same data (Lee and McLanahan 2015). I find 30 imputations with 30 burn-in iterations to produce sufficient convergence on all variables. These 30 imputed datasets are included in the fitting and pooling of survey weighted longitudinal models using the *survey* R package (Lumley 2018). Details on the inclusion of imputed datasets in g-formula computation are in Appendix C.

Appendix B. Neighborhood disadvantage and school quality indices.

Table B.1. Component weights and correlations from principal component analyses (PCA) of neighborhood and school characteristics.

PCA 1: Wodtke index	Weight	Correlation
Percent poverty	0.480	0.950
Percent unemployed	0.450	0.900
Percent receiving welfare	0.460	0.910
Percent female-headed households	0.430	0.860
Percent high school	0.410	0.800
Component variance		3.920
Proportion total variance explained		0.780
PCA 1: School quality		
Percent free lunch	-0.660	-0.840
Percent Title I	-0.340	-0.830
Pupil to teach ratio	-0.660	-0.430
Component variance		1.590
Proportion total variance explained		0.530

Appendix C. G-formula course simulation.

Stochastic simulation for estimating effects in the parametric mediational g-formulas (Equation 9) involves the following general steps. The g-formula associated with no changes to the observed data is termed the *natural course*, and the counterfactual g-formulas associated with the treatment, low disadvantage (NH Quantile 1) vs. high disadvantage (NH Quantile 5), are termed the *treatment* and *control courses,* respectively.

1. Fit a survey-weighted longitudinal model for each time-varying mediator and the outcome.
2. Subset pooled data to first period.
3. Simulate *natural, treatment,* and *control courses* by predicting data forward using the fitted models.
4. Repeat simulations predicting the outcome variable but drawing all other individual values from either the treatment course or the control intervention course in order to isolate all *direct* and *indirect effect courses*.
5. Calculate population means for the outcome under all simulation courses (expected values in Equations 7 and 8).
6. Calculate all effects in Equations 7 and 8 by differencing the population means across appropriate courses.
    a. Cumulative total effect (TE): *treatment course – control course.*
    b. Cumulative natural direct effect (rNDE): *direct course – control course.*

c. Cumulative natural indirect effect (rNDE): *indirect course – control course*.

Table C.1 illustrates a set of rules for controlled simulations across the different courses given the intervention of high neighborhood disadvantage (NH Quintile 5) compared to low neighborhood disadvantage (NH Quintile 1). This process is repeated 2000 times to achieve convergence. Details on the inclusion of multiple imputation and Monte Carlo simulation to propagate uncertainty from modelled parameters to final population-level estimates is provided in the computation outline in Table C.2.

A comprehensive overview of the parametric g-formula is covered in Hernan & Robins (2019), and a worked example with formal probability notation in Keil et al (2014) (Hernán and Robins 2019; Keil et al. 2014). A worked example of the mediational parametric g-formula is provided by Lin et al (2017). All analyses are conducted using the R statistical learning software (v3.5.1). Multiple imputation is performed with the *mice* package (v3.3.0). Complex survey design is accounted for in all models using *svyglm* from the *survey* package (v3.34) with the city-level weights.

Table C.1. Rules for updating time-varying characteristics in g-formula simulation under an intervention of fixing the neighborhood index at the highest and lowest quintiles.

Time-varying characteristic	Natural course	Treatment course	Control course	Direct effect course	Indirect effect course (poverty)
Cognitive score	Predict	Predict	Predict	Predict	Predict
Neighborhood index	Predict	*Quintile = 1*	*Quintile = 5*	*Treatment course*	Control course
Mother poverty	Predict	Predict	Predict	Control course	*Treatment course*
School quality	Predict	Predict	Predict	Control course	Control course
Mother unemployed	Predict	Predict	Predict	Control course	Control course
Mother married/cohab.	Predict	Predict	Predict	Control course	Control course
Physical environment	Predict	Predict	Predict	Control course	Control course
Collective efficacy	Predict	Predict	Predict	Control course	Control course

Table C.2. Computation outline for stochastic simulation of the parametric g-formula.

Multiple imputation for missing data (30 imputations with 30 burn-in iterations)[a].

Fit a survey-weighted generalized linear model (GLM) with appropriate likelihood for each time-varying variable across the 30 imputed datasets[b].

Combine models to get pooled estimates of 1) the vector of coefficient means and 2) the variance-covariance matrix for each time-varying variable using Rubin's rules[c].

Pool imputed datasets by taking the crude average over imputations.

Draw 2000 random samples from each vector of coefficient means and corresponding variance-covariance matrix.

Estimate g-formula to calculate total, direct, and indirect effects of each time-varying variable.

For each sample in 1:2000 {

- Subset pooled data to first period. Replicate 30 times to remove Monte Carlo error in stochastic prediction of individual-level responses during simulations.
- Simulate *natural, treatment,* and *control courses* by predicting pooled data forward using the sample of coefficients. Individual-level response prediction is stochastic for binomial variables.
- Repeat simulations predicting the outcome variable but drawing all other individual values from either the treatment intervention course or the control intervention course in order to isolate all *direct and indirect effect courses*.
- Calculate population means for the outcome under all simulation courses using survey weights.
- Calculate all controlled effects by differencing the population means across appropriate courses[d]:
    - TE: *treatment course – control course.*
    - rNDE: *direct course – control course.*
    - rNIE: *indirect course – control course.*

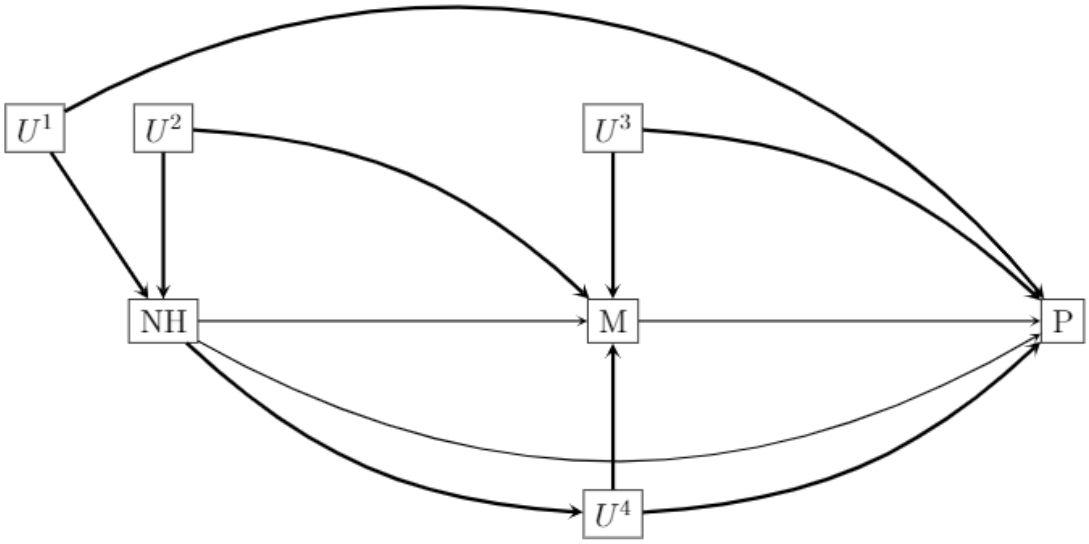
}

Summarize all effects over the 2000 means.

[a] I use the *mice* R package to create 30 datasets via multiple imputation with chained equations.
[b] I use the *svyglm* function in the *survey* R package to account for the complex survey design in estimation of variance.
[c] I use the *pool_mi* function in the *miceadds* R package to average the estimates of the complete data model. This procedure computes the total variance over the repeated analyses and computes the relative increase in variance due to nonresponse and the fraction of missing information.
[d] Given the g-formula estimator is a form of standardization, the direct effect and all indirect effects sum exactly to the total effect.

Appendix D. Causal identification of mediating effects

Figure D.1. Directed acyclic diagram representing threats to causal identification for the indirect effect of neighborhood disadvantage (*NH*) on PPVT score (*P*) that operates through a given mediator (*M*). There may be unobserved exposure-outcome confounding (U^1), unobserved exposure-mediator confounding (U^2), unobserved mediator-outcome confounding (U^3), and/or unobserved treatment-induced mediator-outcome confounding (U^4).



CHAPTER 2

Racism and quantitative causal inference: A life-course mediation framework for decomposing racial health disparities.

Nick Graetz[1], Courtney Boen[1, 2], Michael H. Esposito[3]

[1]Population Studies Center, University of Pennsylvania
[2]Department of Sociology, Population Aging Research Center, University of Pennsylvania
[3]Institute for Social Research, University of Michigan

Abstract Quantitative studies of racial health disparities often use static measures of self-reported race and conventional regression estimators, which critics argue is inconsistent with social constructivist theories of race, racialization, and racism. We demonstrate an alternative counterfactual approach to explain how multiple racialized systems dynamically shape health over time, examining racial inequities in cardio-metabolic risk in the National Longitudinal Study of Adolescent to Adult Health. This framework accounts for the dynamics of time-varying confounding and mediation that is required in operationalizing a "race" variable as part of a social process (*racism*) rather than a separable, individual characteristic. We decompose the observed disparity into three types of effects: a controlled direct effect ("unobserved racism"), proportions attributable to interaction ("racial discrimination"), and pure indirect effects ("emergent discrimination"). We discuss the limitations of counterfactual approaches while highlighting how they can be combined with critical theories to quantify how interlocking systems produce racial health inequities.

As early as Du Bois (1899), critical race scholars have linked racialized disparities in health to the unequal social and economic conditions produced by racism. Over the past century, social scientific research has built on the work of Du Bois, generally adopting and expanding a social constructionist view of race and racism. A large and growing body of theoretical and empirical work describes the connection between macro systems of racial oppression, the construction of racial categories, and

how exposure to racist systems becomes "embodied" to produce racialized health inequities (Bailey et al. 2017; Du Bois 1899; Geronimus et al. 2006; Goosby, Cheadle, and Mitchell 2018; Green and Darity 2010; Phelan and Link 2015; Reskin 2012; Sewell 2016).

Still, despite a strong theoretical foundation for a relational, social constructionist view of race and racism, quantitative analyses of racial health disparities generally fall short of translating theories of racism to empirical models beyond descriptive analyses (Muntaner 2013; Reskin 2012). In recent years, critical race scholars have critiqued how "race" is interpreted in counterfactual frameworks, especially regression models (Kohler-Hausmann 2019; Reskin 2012; Sen and Wasow 2016; Williams 2019; Zuberi and Bonilla-Silva 2008). In addition to treating race as a fixed individual trait, critiques regarding the measurement and modeling of race and the effects of racism also stem from the limited set of counterfactuals offered by conventional regression models. Particularly troubling is that, when applied to longitudinal data, conventional regression assumes an absence of time-varying relationships among variables. These assumptions are largely inconsistent with a sociological interpretation of race as capturing part of a relational social process, rather than a fixed characteristic of the individual. In quantitative studies of health, a race variable is often included in a conventional regression model, with additional variables added to "explain away" racial variation in the outcome. Given the interpretation of the race variable as attempting to capture a relational process, however, this approach warrants critique. Is a static comparison of racial groups in which all other

correlated exposures are held equal the most useful counterfactual we can estimate in quantitative studies? We argue that it is not.

What, then, are the most appropriate counterfactuals to quantify how the system of racism works to produce population health patterns? And importantly, are these counterfactuals estimated by traditional regression models? We argue that conventional analyses to understand mechanisms that produce racialized variation in health require several strong, often untenable assumptions. Particularly concerning is that conventional regression estimators risk reifying race as a static individual trait that can separated from other systems of social stratification such as social class position, rather than as mutually co-constituted with these other systems of inequality over time (Kohler-Hausmann 2019; Zuberi and Bonilla-Silva 2008). By contrast, recent advancements in causal mediation analysis have provided new methods for modeling how dynamic social processes produce population health patterns over time (Bauer and Scheim 2019; Esposito 2019; Wang and Arah 2015). We argue that these approaches address several shortcomings of traditional quantitative approaches to studying racial disparities in health that can be of great utility to sociological research on racial health inequities by better accommodating dynamic and relational theories of race and racialization.

In this study, we combine a novel technique for modelling the relative importance of multiple dependent mediators with a robust social theory of the relational system of structural racism (Bailey et al. 2017; Bonilla-Silva 1997; Reskin 2012) to examine life course processes contributing to racial disparities in health. Given that the racial stratification of socioeconomic resources, risks, and opportunities is a key pathway

linking racism to population health inequities (Boen 2016b; Phelan and Link 2015), we pay particular attention to the mediating and cumulative roles of socioeconomic exposures in the production of racial health disparities. Socioeconomic processes are the only factors contributing to racialized gaps in health, but instead they are an important and frequently studied component in mediation analyses. That socioeconomic exposures also evolve across the life course is an important feature for demonstrating the utility of our approach for handling the complex interplay of time-varying mediators and confounders. Using longitudinal data from the National Longitudinal Study of Adolescent to Adult Health (Add Health), including biomarker data and a variety of life course socioeconomic exposures, we decompose the total Black-white disparity in cardio-metabolic risk to three types of cumulative life course effects that each have intuitive interpretations in our theoretical framework (Jackson and VanderWeele 2019; Wang and Arah 2015): 1) the controlled direct effect (CDE) of racism (i.e. unobserved pathways through which racism becomes embodied that do not operate through the observed mediators, which we call *unobserved racism*; 2) the portion attributable (PAI) to interaction for each observed socioeconomic mediator (i.e. how the effects of a socioeconomic exposure on health varies by racialized category, which we label *racial discrimination*; and 3) the pure indirect effect (PIE) for each observed mediator that we call *emergent discrimination*.

By combining a critical race theoretical approach with modern causal inference methods, our study highlights the limitations of conventional regression models for identifying the life course pathways of structural racism governing the production of

racial health inequities. Importantly, we also offer an alternative mediation framework for examining the social roots of racial health inequities that is more consistent with a social constructivist understanding of racism and race than conventional regression.

Background

The social construction of race

In his essay "The Superior Race," W.E.B. Du Bois (1940) imagines himself in a dialogue with a fictional white character, Roger Van Dieman. In the piece, Du Bois tries to explain to Van Dieman that race is a social, cultural, and historical fact – not a natural or inherent biological phenomenon. Van Dieman presses Du Bois, ultimately asking him how he can tell who is Black if race cannot be objectively measured or ascertained biologically, as was widely practiced at the time. Du Bois closes the essay stating,

> *"I recognize it quite easily and with full legal sanction: The Black man is a person who must ride the 'Jim Crow' in Georgia."*

Du Bois's simple retort carries tremendous meaning by concisely pointing to the dynamic and relational processes that serve to create, reify, and give meaning to race. A man is not Black because of some measurable individual trait or phenotypic characteristic; a man is Black because social, institutional, and legal structures *treat* him as Black. Since Du Bois, sociological scholarship has played a prominent role in expanding this social constructivist view of race, highlighting the essential roles of structures and institutions in shifting and maintaining racial boundaries (Bonilla-Silva 1997; Roberts 2011; Sewell 2016; Zuberi 2001; Zuberi et al. 2015).

Theoretically, recent studies of racial health inequities generally adopt a social-constructionist understanding of racism, racialization, and race, with a large and growing body of work in this area focused on the historical, social, and political roots of these disparities (Bailey et al. 2017; Phelan and Link 2015; Williams 2012). Still, identifying the role of structural racism in governing the distribution of various types of risk poses significant empirical challenges. A common approach in quantitative studies of racial health disparities is to include a time-invariant measure of race in regression models and subsequently control for correlated variables, such as socioeconomic status, to isolate variation that is explained by racial categories. This approach treats constructs like "race" and "socioeconomic position" as separable, rather than co-constituted and historically contingent, constructs. Still, structural racism has been described as "the totality of ways in which societies foster racial discrimination through mutually reinforcing systems of housing, education, employment, earnings, benefits, credit, media, health care, and criminal justice" (Bailey et al. 2017). The logic of conceptual separability implicit in regression estimators therefore carries a significant risk of reifying "race" as an independent construct rather than part of the time-varying reciprocal process of racialization and racism within and across socioeconomic, political, and cultural systems (Sewell 2016; Zuberi and Bonilla-Silva 2008; Zuberi et al. 2015).

Studies focusing on "direct effects" of race

The approach to separating "race" from social exposures that are "not race" is the basic counterfactual logic used in audit studies and legal studies of racial discrimination (Kohler-Hausmann 2019). In cases involving hiring discrimination, for example, the

residual racial variation that remains after considering other factors like educational attainment is interpreted as *the* effect of racial discrimination (Kohler-Hausmann 2019). This same logic is deployed in audit studies, which are appealing for the precision of their experimental manipulations. These studies are often interpreted as premier causal evidence of racial discrimination, as if naming a hypothetical job-seeker Jamal vs. Brendan, for example, while holding everything else constant, is capturing something that is more *truly representative of racism* than the complex, longitudinal systems that produce racialized distributions of everything else that might appear on a resume.

While such studies identify important dynamics of explicit forms of racism, they typically offer no concurrent comparison of how much the total observed racial disparity in an outcome like hiring is due to employer decision-making at the final step before employment versus, for example, racial differences in exposure to penal systems that shape the probability of an employer receiving differently racialized individuals' resumes in the first place. Indeed, by holding all else constant, these studies implicitly suggest that it is possible to separate "racial discrimination" from factors that are "not race," like socioeconomic position. But racial variation in *any* characteristic can only be the result of historical and contemporary projects of racism and racialization (Bonilla-Silva 2009; Roberts 2011; Williams et al. 2016).

Structural racism and the race discrimination system

Variation arising from racial stratification in the mutually reinforcing systems of labor, housing, and education is central to theories of structural racism (Bonilla-Silva

1997, 2009; Reskin 2012). Reskin's (2012) description of the "race discrimination system" is useful for understanding the dynamic and relational processes producing racial health disparities. Building on Bonilla-Silva (1997), Reskin's framework highlights the need for a systems approach to describing the process of racial stratification as a broad causal system consisting of various subsystems (e.g., residential and school segregation, housing and mortgage markets, health services). Importantly, Reskin (2012) contends that in the quantitative literature, scholars often attempt to isolate and manipulate a single subsystem and estimate the direct effect, assuming that doing so would not change other subsystems. In contrast, a systems approach to studying structural racism requires scholars to simultaneously consider that: 1) racialized disparities exist across many exposures, 2) disparities across exposures are mutually reinforcing, 3) the source of disparities is racial discrimination, and 4) there exists discrimination in the effects of exposures (Reskin 2012). Reskin's (2012) use of the term "discrimination" can be confusing, given its inconsistent use across academic fields and its colloquial interpretation. To the extent that there is *any* racial difference in an exposure-outcome association (i.e., an interactive effect that shows differential effects by race), the sub-system governing that exposure-outcome relationship is racializing individuals and "discriminating" based on those racial categories. We will use this language of "*disparities* in exposures" and "*discrimination* in exposure effects" throughout this paper.

This longitudinal framework, which considers the causal dependence of multiple sub-systems of structural racism acting on racialized individuals over time, presents an array of methodological issues for quantitative inference, including concerns about

parameterizing mediation and confounding. Overcoming these issues is the primary purpose of this paper, as we believe they are important to addressing foundational concerns raised by Zuberi & Bonilla-Silva (2008), Kohler-Hausmann (2019), and others with regards to the implicit treatment of variables as independently "race" or "not race" in conventional regression.

Traditional approaches for explaining racial disparities in health

Link & Phelan (2015) describe racism as a fundamental cause of racial health disparities, a broad causal theory that is central to study of structural racism. The fundamental cause model is frequently invoked in quantitative studies of racial health disparities, typically paired with a regression model that includes mediating pathways. Importantly, the fundamental cause model emphasizes the importance of metamechaisms, highlighting that the specific mechanisms through which the fundamental cause is operating at a given time are can be easily substituted. Available data often inhibits directly studying the metamechanisms emphasized by the fundamental cause model, which is itself a considerable problem (see Krieger 2018). Still, existing longitudinal cohort data can be usefully combined with cumulative life course theory and frameworks examining mediating pathways (Boen 2016b; Brown 2018; Brown et al. 2016; Kramer et al. 2017; Williams et al. 2016) to retrospectively study how a fundamental cause like racism operates through a particular space and time to produce racial health disparities.

However, the empirical challenges of considering and accounting for confounding—which have been described by Diez-Roux (2012), Reskin (2012), and

Robinson and Bailey (2019)—are important to address in the context of mediation analyses. The most basic mediation model implied by Link & Phelan (2015) is illustrated with a directed-acyclic diagram (DAG) in Figure 1, which attempts to explain how much of the racial disparity in health (*Y*) is explained by some indicator of material deprivation or socioeconomic position (*M*). For illustrative purposes, consider that *M* is household income. This is a causal mediation question that relies on four key assumptions. As we will discuss below, each assumption is predicated on a specific understanding of how the exposure, typically operationalized as self-identified race, fits into the theoretical causal process of racism and racialization.

- **Assumption 1.** No unobserved confounding of Race → *Y* (***C*** contains all relevant confounding variables of race and health).
- **Assumption 2.** No unobserved confounding of Race → *M* (***C*** contains all relevant confounding variables of race and household income).
- **Assumption 3.** No unobserved confounding of *M* → *Y* (***L*** contains all relevant confounding variables of household income and health).
- **Assumption 4.** No *M* → *Y* confounders (***L***) affected by Race (the dotted arrow between Race and ***L*** does not exist; the "cross-world independence assumption").

This DAG is frequently combined with the following mediation formula for how much of the racial health disparity is "explained" by household income (*M*), potentially including other "confounders" (***L***, ***C***) in both models:

$$Y = \beta_0 + \beta_1(Race) + \boldsymbol{\beta_2}(\boldsymbol{L}) + \boldsymbol{\beta_3}(\boldsymbol{C}) + \varepsilon$$

$$Y = \alpha_0 + \alpha_1(Race) + \boldsymbol{\alpha_2}(\boldsymbol{L}) + \boldsymbol{\alpha_3}(\boldsymbol{C}) + \alpha_4(M) + \delta$$

$$\textit{\% of racial disparity in Y explained by M} = \frac{(\beta_1 - \alpha_1)}{\beta_1}$$

Often referred to as Baron-Kenny (1986) mediation, this approach simply asks: conditional on confounders, how much of the racial disparity in health outcome *Y* remains after controlling for the mediator *M* (income). There are three important considerations to discuss in this mediation framework that are often neglected in quantitative attempts to explain racial disparities.

First, in decomposing the total racial health disparity in health (*Y*), we do not consider any "pre-exposure" variables $\boldsymbol{C}$ (e.g., parental income) to be confounders of Race and health. Doing so implicitly treats Race as an individualized exposure that begins at birth, rather that part of a relational, multigenerational system of racism. Mediation analysis after controlling for many "pre-exposure confounders" means decomposing a marginal racial disparity conditional on other variables treated as "not race," such as parental material resources. While this construction of the causal model may be illuminating in explaining an artificial, hypothetical disparity, it is of rather limited utility in explaining the total, real-world observed disparity. Any racial disparity in health (*Y*) is the result of historical and contemporary racism, and every observed "pre-exposure" variable that differs across racialized categories is more appropriately considered a mediator (*M*) rather than a confounder.

Second, there are three distinct causal pathways through which household income (*M*) might play a role in maintaining the association between self-identified race and

health that are not separately identified in the Baron-Kenny approach (Jackson and VanderWeele 2019; Wang and Arah 2015). Understanding the relative magnitudes of each of these pathways is theoretically important, as each map onto different facets of Reskin's (2012) race discrimination system.

- **The controlled direct effect (CDE)**, Arrow (c). This includes all unobserved mediating pathways through which racism becomes embodied or otherwise influences *Y* that do not operate through the measured mediators – in other words, *unobserved racism*.
- **The proportion attributable to interaction (PAI)**, Arrow (a) + the interactive effect of *Race*M* through Arrow (b). This is a form of structural *racial discrimination*, whereby the underlying system governing the relationship between *M* and *Y* operates differently across racialized categories.
- **The pure indirect effect (PIE)**, Arrow (a) + the main effect of *M* through Arrow (b). This is a form of *emergent discrimination*, whereby racism patterns the distribution of *M* via Arrow (a) but *M* affects *Y* regardless of racialized category.

As discussed above, in many studies the CDE is often the target effect. The implicit logic is that this effect is somehow representing the "true effect of racial discrimination" because we have removed other indirect pathways connecting racial categories to *Y* which are conceptually bracketed as "not race" (e.g., income). Still, the CDE is completely dependent on which observed mediators researchers include in their estimation and is therefore more accurately considered a measure of our ignorance, or the proportion of the total racialized disparity in *Y* that cannot be explained through observed

mediating systems. Instead of focusing exclusively on parameterizing the CDE, researchers of racial health inequities should more broadly consider how a system of structural racism produced racial stratification across a particular cohort aging through a particular time and place via *M*; or, as described by Stewart (2008), "swimming upstream."

Third, the conventional mediation analysis requires that Assumption 4 holds by assuming no variables in ***L*** are influenced by Race (no dotted arrow in Figure 1). This is theoretically untenable, as racism is implicated in patterning the distributions of virtually all other variables that we might consider confounding the relationship between household income (*M*) and health (*Y*) (A. I. Naimi et al. 2016a). In this example of the mediating effect of income, we might consider the neighborhood poverty rate to be one of the confounders (***L***) of the relationship between household income and health, as neighborhood context affects both one's level of household income and health. However, it is well-established that racism produces variation in neighborhood context across racialized categories via residential segregation (Charles 2003; Massey and Denton 1993). In a life course perspective, then, neighborhood context and household income are also reciprocally intertwined over time (Sharkey and Elwert 2011; Wodtke et al. 2011).

What does this do to the interpretation of how much of the racial disparity is "explained by" household income in the conventional mediation analysis above? Researchers are left with a tricky counterfactual: what would the racial disparity have been if *M* had been fixed at its reference category *and all other variables (**L**) had also been fixed at their reference categories*? Studies often adjust for how neighborhood

poverty (L) confounds the effect of household income (M) on health (Y), but in doing so, over-control the mediating pathway through neighborhood poverty, Race $\rightarrow L \rightarrow Y$. In assessing the mediating effect of household income on the relationship between race and health, it is impossible in the conventional regression framework above to both control for and not control for neighborhood poverty (Sharkey and Elwert 2011). This is part of the "all else equal" axiom: how much do those racialized as Black differ in Y compared to those racialized as white, *all else equal*? Regardless of intent and emphasis, this counterfactual frame focuses attention on the residual racial disparity (CDE) (Kohler-Hausmann 2019; Zuberi and Bonilla-Silva 2008). In attempts to control for structural confounding, regression and matching estimators are often describing marginal, artificial worlds that bear little resemblance to the reality of how the dynamic, relational *process* of racism operates over time. The quantitative field of racial health disparities increasingly leans on the nuanced theoretical and conceptual models of race, racialization, and structural racism as discussed above, but still often relies on conventional regression estimators.

New developments in quantitative causal mediation

In Figure 2, we consider three sets of mediators of the racial health disparity ($\boldsymbol{M^{(1)}, M^{(2)}, M^{(3)}}$) observed over three age windows. When decomposing a disparity in Health by self-identified race, we are describing a particular component of the historical causal process that connects the system of racism to population variation in health, following a particular racialized cohort through a particular space and time. Using self-identified race in this way provides an incomplete picture of the historical process of

racialization and the effects of racism, which are malleable over space and time for different cohorts. But as many other authors have described in defending the need to monitor health indicators by race, measuring these disparities can still be useful in identifying causal mechanisms maintaining and reproducing racial stratification that lead to embodied health inequities for particular cohorts (Chowkwanyun and Reed 2020; Krieger 2018).

In this longitudinal framework, we are no longer conceptually separating variables into "race" and "not race" in the same way that might be implied by conventional regression estimators. For example, variation in $\boldsymbol{M^{(2)}}$ is explicitly defined as the historical product of several causal pathways:

1. Race is linked to the distribution of $\boldsymbol{M^{(2)}}$ through unmeasured pathways not captured by $\boldsymbol{M^{(1)}}$ (*unobserved racism*).
2. Race is linked to the distribution of $\boldsymbol{M^{(1)}}$ through unmeasured systems of racism.
    a. $\boldsymbol{M^{(1)}}$ then influences $\boldsymbol{M^{(2)}}$ through an underlying system that racializes individuals and acts upon them differently (*racial discrimination*).
    b. $\boldsymbol{M^{(1)}}$ then influences $\boldsymbol{M^{(2)}}$ regardless of racialized category (*emergent discrimination*).

The variables in $\boldsymbol{M^{(2)}}$ are thus not representing some proportion of variation in Health that is separable from Race, but rather are explicitly part of the process of racism through which racial categories and Health become connected. It is important to note that such a causal decomposition is happening at a "high-level." Each arrow is governed by its own

underlying causal system, which includes specific actors producing and interacting with processes of racialization and racism. For example, if we consider $\boldsymbol{M}^{(1)}$ to include neighborhood poverty rate, Sewell (2016) describes the political economy undergirding the Race $\rightarrow \boldsymbol{M}^{(1)}$ arrow, and Kramer et al. (2017) describe the processes of embodiment underlying the $\boldsymbol{M}^{(1)} \rightarrow Y$ arrow. Goosby et al. (2018) describe how stress-related biological mechanisms related to interpersonal discrimination affect health outcomes (Goosby et al. 2018), experiences which could be implicated in the causal arrow connecting Race $\rightarrow Y$ in Figure 2 and within the systems of racial discrimination governing interactive effects of structural mediators (e.g. racial discrimination influencing the effect of $\boldsymbol{M}^{(1)} \rightarrow Y$).

Our primary goal in this paper is not to specify every pathway by further differentiating each arrow in Figure 2, but instead to characterize the high-level structural system producing racialized disparities in an important indicator of early adult health (cardio-metabolic risk) among this particular cohort, as well as the relative magnitudes of each of these broad mechanisms in a way that reflects our theoretical framework.

The present study

The present study uses modern causal mediation methods that are informed by a critical theory of racism as a fundamental cause of health inequities. The focus of this empirical case study is not on the process of racialization or the proximal mechanisms of interpersonal racism, but on the causally interconnected, reciprocal systems of structural racism (Bailey et al. 2017; Reskin 2012). We consider the process of a cohort aging

through a system of racial stratification as a causal inference problem of multiple mediators, where each intermediate exposure affects all others over time and in ways that vary by how an individual is racialized (Bailey et al. 2017; Esposito 2019). We argue that intentionally or not, the restrictive assumptions of regression estimators for mediation analysis often reify a problematic paradigm of race as a separable individual risk factor, rather than examining racism as a time-varying causal process. Building on the work of Arah & Wang (2015) and VanderWeele & Tchetgen Tchetgen (2017), we apply a g-formula method for multiple causal mediation to decompose the total racial disparity in a direct measure of cardio-metabolic risk in a longitudinal cohort of young adults. This decomposition consists of the following three effect pathways, which all represent different facets of racism: the controlled direct effect, the portion attributable to interaction for each mediator, and the pure indirect effect for each mediator. Under the assumption of no unmeasured confounding, we provide causal evidence for multiple mediating pathways through which structural racism produced embodied cardio-metabolic disparities for this specific cohort. We describe the relative magnitude of mediating effects through these systems, including exposure to neighborhood poverty, educational attainment, and household income.

Data and Methods

Data

We use three waves of data from the National Longitudinal Study of Adolescent to Adult Health (Waves I, III, and IV: 1994-2008). Our key outcome is a continuous

measure of *cardio-metabolic risk* at Wave IV, defined as the first principal component of five biomarker variables indicating: 1) elevated waist circumference, 2) elevated blood pressure, 3) elevated triglycerides, 4) reduced high-density lipoprotein, and 5) pre-diabetic value of glycosylated hemoglobin (HbA1c). Previous research has demonstrated that this measure is a leading risk factor for cardiovascular disease in the United States (Kane 2018). We normalize this index to have a mean of zero and variance of one.

We define each wave of data collection as "Adolescence" (Wave I; ages 12-18), "Transition to adulthood" (Wave III; ages 18-25), and "Young adulthood" (Wave IV; ages 25-32) and control for continuous years of age at the time of interview and self-identified sex in all models. Focusing on the role of socioeconomic factors, we consider the following causally ordered variables to mediate the relationship between self-identified race and cardio-metabolic risk in young adulthood (Figure 2):

- Adolescence ($\boldsymbol{M^{(1)}}$): annual household income (continuous US dollars), parental educational attainment (1 = at least one parent completing college), tract-level poverty rate (continuous).
- Transition to adulthood ($\boldsymbol{M^{(2)}}$): tract-level poverty rate (continuous).
- Young adulthood ($\boldsymbol{M^{(3)}}$): college attainment (1 = respondent completed college), tract-level poverty rate (continuous), annual household income (continuous US dollars).

All mediators except neighborhood poverty rate are interacted with race. Following Wodtke et al. (2011), our measure of neighborhood poverty rate is included in the

outcome model as cumulative over the individual's entire life. We use the non-censored restricted Add Health sample, which includes 10,052 unique individuals self-identified as white (25,888 person-years) and 3,893 unique individuals self-identified as Black (9,741 person-years).

Two types of non-random missing data may result in biased effect estimates: item-nonresponse and censoring of observations. We create 30 multiply imputed datasets using chained equations to account for missing observations due to item-nonresponse (Buuren and Groothuis-Oudshoorn 2011). G-formula simulations begin with the full sample in the first wave and every individual is simulated through all subsequent waves. In effect calculations, we are then including all simulated person-years that were censored in the survey sample. This is analogous to the approach used in marginal structural models to correct for non-random censoring biasing effect calculations (VanderWeele 2009).

Parametric g-computation

The "g-formula" or "g-computation" is a generalization of standardization that allows for the estimation of unconfounded summary effects without relying on the restrictive cross-world independence assumption inherent to conventional regression estimators. In conventional regression models (e.g. Baron-Kenny mediation) or demographic decomposition (e.g. Das Gupta or Kitagawa decomposition), estimates of counterfactual change are calculated under the assumption that no other conditional probabilities change as a result of the exposure changing. In contrast, g-formula

standardization makes explicit the sum of all "cascades" of conditional probabilities for all variables as the cohort ages through that time and space, consistent with Stewart's (2008) "swimming upstream" (Figure 2). We use the g-formula, described in Appendix A, to decompose the observed Black-white disparity in cardio-metabolic risk, using fully interacted models to account for the racialized processes underlying all systems in Figure 2.

We first present descriptive statistics for the outcome and all mediating variables by self-identified race. We then present the effect decomposition for the observed difference in cardio-metabolic risk scores between the population racialized as Black compared to the population racialized as white. We map relevant research questions for each decomposed effect to their conceptual definitions in Table 1, using the specific example of educational attainment to demonstrate main ideas.

Results

Table 2 provides summary statistics by self-identified race for each mediating variable and the outcome variable (N=10,052 and N=3,893 for the cohorts self-identified as white and Black, respectively). The reference category for all binary mediators is the value hypothesized to be associated with higher cardio-metabolic risk (e.g. 0 = no college attainment). We observe a significant disparity in our index of cardio-metabolic risk in young adulthood: 0.35 standard deviations higher in the cohort racialized as Black compared to white. We also observe large disparities across all mediating variables. The average neighborhood poverty rate for Black respondents in adolescence was 26 percent

compared to just 11 percent for white respondents, a disparity which largely remained intact into adulthood. In addition, white adolescents lived in households with an average income of $51,66 compared to $30,496 for Black adolescents. A smaller percentage of Black adolescents lived in a household where a parent completed college compared to white adolescents (16 vs. 24 percent, respectively). In young adulthood, the disparity in college attainment was 22% vs. 33%, suggesting a general increase in college attainment across generations but also an increase in the Black-white gap.

Appendix Table A.1 displays parameter estimates from all mediator and outcome survey-weighted models. These are used to parameterize conditional probabilities across the life course in the g-formula (Appendix A) and should not be given a causal interpretation. Table 3 provides results from the mediation analysis. In this counterfactual decomposition, we divide each decomposed effect by the total effect to determine the proportion of the racial health disparity that would be eliminated if that effect pathway had not operated on this cohort as they aged from adolescence to young adulthood (Table 1). For example, results in Table 3 reveal that if the cohort racialized as Black had instead been racialized as white by the system connecting educational attainment to cardio-metabolic risk, we would expect the total disparity in cardio-metabolic risk by adulthood to be reduced by 0.06 standard deviations (-0.11 to -0.01, $p<0.05$), or 18% of the total observed disparity. Figure 3 illustrates the sampling distributions of each decomposed effect across all simulations and multiply imputed datasets.

Overall, results indicate that roughly 58% of the total racial disparity in the index of adult cardio-metabolic risk can be explained by the observed mediating pathways.

Three of these mediating effect pathways are statistically significant ($p<0.05$): the pure indirect effect (PIE) operating via cumulative exposure to neighborhood poverty across the life course (-0.07; -0.12 to -0.02, 21%), the proportion attributable to interaction (PAI) of whether a parent of the respondent completed college (-0.05; -0.09 to -0.01, 15%), and the proportion attributable to interaction (PAI) of whether the respondent completed college (-0.06; -0.11 to -0.01, 18%). The remaining 42% of the observed racial disparity (CDE) is explained by unobserved mediating pathways operating outside the structural and socioeconomic measures included in this analysis (i.e. at the reference values of the included mediators).

The two interactive effect pathways that are most important in explaining the adult disparity in cardio-metabolic risk involve returns to parental educational attainment and returns to personal educational attainment. These pathways both serve as examples of structural racism via *racial discrimination* in the causal systems governing the returns to education on adult cardio-metabolic health in the United States during this particular period. Our results indicate that, in some way, that system connecting educational attainment to health treated individuals racialized as Black in this cohort differently that those individuals racialized as white.

In contrast, results indicate that cumulative exposure to neighborhood poverty across the life course is the most important pure indirect effect contributing to the observed racial disparity in cardio-metabolic risk, which serves as an example of structural racism via what Reskin (2012) terms *emergent discrimination*. Living in a neighborhood with a high poverty rate impacts health regardless of racialized category.

Still, the racialized distributions of this harmful exposure are highly unequal in ways that contribute to racial health inequality. In this cohort, 16 percent of respondents racialized as white lived in neighborhoods with a poverty rate exceeding 20 percent in adolescence, compared with 61 percent of those racialized as Black. In the g-formula framework for causal mediation, we are able to account for the fact that neighborhood poverty at one age influences all subsequent exposures (as shown in Figure 2). However, our estimate of this particular indirect effect (21 percent) is net of those additional indirect pathways (the arrows directly connecting this neighborhood poverty to cardio-metabolic risk in Figure 2), suggesting the particular salience of neighborhood context across the life course in shaping adult health through unobserved mediating pathways beyond its cumulative impact on adulthood education and income.

Discussion

Structural racism as a relational process

If race, racialization, and racism are dynamic relational processes then our quantitative methods for studying racial inequality must move beyond the limitations of conventional regression models, which largely treat "race" as a static trait of individuals that is separable from racialized constructs like socioeconomic position. In this study, we have provided a worked example of a life course mediation framework for decomposing racial disparities in adult cardio-metabolic function, though this framework can be generalized to the decomposition of any population disparity. We have clarified the relative importance of several structural mediators within the system of racism

experienced by the this particular cohort, such as the importance of *emergent discrimination* across the life course via disproportionate exposure to neighborhood poverty experienced by those racialized as Black and *racial discrimination* via the system governing the relationship between educational attainment and adult health. In doing so, this study serves as one example by which a self-identified race variable can be used to help explain the causal process of racism in shaping outcomes in quantitative studies.

Using nationally representative, longitudinal data from Add Health, we decomposed the Black-white disparity in cardio-metabolic risk by young adulthood, paying particular attention to the roles of life course socioeconomic exposures in the production of the racialized health inequity. Our results revealed a large, racialized disparity in cardio-metabolic risk within this cohort, on the magnitude of 0.35 standard deviations. We further decomposed this disparity using several socioeconomic pathways, implicating cumulative dynamics of structural racism arising from both racial discrimination and emergent discrimination across the life course. Theoretically and empirically, we understand the production of this racialized disparity in health as a relational process in which various structural mediating systems both create differential exposures based on racialized category (e.g., *emergent discrimination*) and treat individuals differently based on racialized category (e.g., *racial discrimination*).

A large proportion of the total disparity in cardio-metabolic risk in young adulthood in this cohort — 58 percent — was explained by the observed socioeconomic mediators. Consistent with previous research (Boen 2016; Brown 2018; Phelan and Link 2015), these results indicate the prominence of socioeconomic factors in the production

of racialized health disparities. Much of this is the result of neoliberal political arrangements in the United States focused on the coupling of health and well-being to individualized "human capital," as well as how this process interacts with contemporary and historical racism (Bailey et al. 2017; Laster Pirtle 2020; Muntaner et al. 2010) to produce racialized socioeconomic distributions. We further found that parental socioeconomic status shapes adult health outcomes, with racial disparities in parental SES playing a key role in the generation of racial health inequities. Parental SES shapes health through many structural mediators observed here across the life course (e.g. children's eventual likelihood of college completion), but we demonstrate that they also have significant indirect effects.

Our results indicated that differential returns to parental educational attainment and returns to personal educational attainment by race also produced racialized health disparities. Reskin (2012) highlights these as examples of structural racism via *racial discrimination* in the causal systems governing the returns to education on adult cardio-metabolic health. This finding, consistent with the "differential returns hypothesis," (Boen 2016; Esposito 2019) could be due to many factors, including how systems of education convey advantage in racialized ways that may eventually impact health outcomes (e.g., occupational status, access to social and professional networks, different modes of social and cultural capital, stress, etc.). A key take-away from this finding is that equalizing access to education would be insufficient in closing Black-white health gaps without parallel efforts to dismantle racism in other domains of social, economic, and political life.

Our results further indicated that exposure to neighborhood poverty contributed to the observed racial disparity in cardio-metabolic risk, which serves as an example of structural racism via what Reskin (2012) terms *emergent discrimination*. High levels of neighborhood poverty during adolescence were harmful for health, and Black individuals were significantly more likely than white individuals to be exposed to high levels of neighborhood poverty in early life. While Figure 2 shows how cumulative neighborhood poverty across the life course influenced all subsequent exposures, we also observed pure indirect effects of neighborhood poverty on adult cardio-metabolic risk. In these ways, results are consistent with research pointing to the prominence of childhood and adolescence as sensitive periods for health and development, when exposure to neighborhood poverty shapes future patterns of health directly and indirectly by shaping adult education, occupation, income, and neighborhood context (Kramer et al. 2017; Kravitz-Wirtz 2016).

Roughly 42 percent of the Black-white gap in cardio-metabolic risk was explained by unobserved mediating pathways, including structural racism via sub-systems not accounted for in our analysis (e.g. punitive policing and court systems, access to healthcare, housing, etc.) and biopsychosocial stress response pathways related to interpersonal racism occurring via discrimination, social exclusion, disenfranchisement, media and state surveillance, and exposure to environmental hazards (Alexander 2012; Boen 2020; Goosby et al. 2018). To the extent these unmeasured systems operate along the causal pathways accounted for in our analysis, they will be included in those effects, but could be further decomposed. For example, contact with

punitive policing and court systems may mediate both the relationships between 1) self-identified racialized category and adult health (i.e. explain a portion of the CDE of 42 percent) and 2) the relationship between neighborhood poverty in adolescence and adult health (i.e. explain a portion of that PIE of 21 percent). In other words, additional causal pathways can be included under different arrows in the DAG described in Figure 2.

Constructing causal models to explain health disparities

It is essential to underscore that our results are ultimately predicated on our choices in how we draw our DAG — choices that can only be justified by our theoretical conceptions of each construct, their causal ordering, and their causal relationships. The empirical goal of the current analysis is to characterize the impacts of multiple causally dependent mediators related to structural racism at a high level in shaping racial health inequities, but in doing so we leave many causal arrows as "black boxes" with insufficiently deep explanation. Extensive empirical work, both quantitative and qualitative, exists for any one of these arrows (for example, see Sewell (2016) on the dynamic relationship between race-racism and residential segregation or Braveman et al. (2011) on the connection between educational attainment and health). Still, the empirical approach employed in this study shows promise for identifying and further examining the mediating pathways whereby structural racism contributes to population health gaps and putting those pathways in a broader context relative to the total real-world disparity.

By combining a race interacted model with an analysis of mediating pathways (Diez Roux 2012), our analysis supports a theoretical framework grounded in the causal

dependence between racism and class relations as a fundamental cause of health in the United States (Laster Pirtle 2020; Phelan and Link 2015; Robinson 2019). Our results speak to the critical roles of the unequal racialized distributions of socioeconomic risks, opportunities, and exposures—as well as racialized effects of these exposures—in the production of racial health disparities. Still, as Williams (2019: 4) notes:

> *"It may be more feasible for scholars to move away from asking 'what factors account for racial inequality' to asking, 'what factors maintain racial inequality.' The former question tends to lead to human capital (i.e., individualistic) explanations whereas the latter question lends itself to racism-based (i.e., structural) explanations."*

The causal arrows and their racialized interactions in Figure 2 do not arise naturally or inevitably, as is often implied by uncritical theories of "social determinants of health" that do not simultaneously consider relational systems of social and economic production. Instead, each causal arrow is maintained by a political economy of specific actors who stand to benefit from particular arrangements rooted in capitalist interests and white supremacy (Cogburn 2019; Laster Pirtle 2020; Roberts 2011; Sewell 2016).

Limitations

There are limitations to our study, which scholars should continue to address in future research. The process of racism is gendered (Crenshaw 1991), which may (in its most simplified form) involve three-way interactions of all observed structural mediators with race and gender (Bauer and Scheim 2019). By not including this additional

interaction, we average over gendered heterogeneity in racialized processes. In addition, gender and age are often considered "demographic controls" in quantitative studies but the extent to which gender and age compositions vary across racialized categories is itself the result of historical racism (Howe and Robinson 2018).

In decomposing a population health disparity along the lines of self-identified race, we are focused on retrospectively examining the forward process by which racialized individuals moved through a system of racism; what Quincy Stewart refers to as "swimming upstream" (Stewart 2008). It is beyond the scope of this analysis to also consider how these processes then serve to reify racial categories; for a comprehensive discussion of the "racism-race reification process" see Sewell (2016).

Last, we believe it is important to qualify the fundamental limitations of quantitative counterfactuals. There are philosophical distinctions between different causal frames, such as those between interventionist causation and etiologic causation (Schwartz, Gatto, and Campbell 2011). These have important implications for counterfactual inferences that rely on unobserved or "potential" outcomes. We consider the counterfactual "What if those racialized as Black had instead been racialized as white within the racist structural systems characterizing the specific time and place through which the Add Health cohort aged?" to be based on a "well-described exposure" (e.g. racialization in a racist system) rather than necessarily a "well-defined intervention." This is a key conceptual distinction between the more holistic, etiologic philosophy of retrospective counterfactual logic (Glymour 2006; Hafeman and Schwartz 2009; Holland 2008; Pearl 2014; Schwartz et al. 2011) and the prospective "interventionist" or

"consequentialist" frame often championed by foundational figures in the potential outcomes methodology of economics and epidemiology (Hernán and Robins 2019). We are not interpreting any of the decomposed effects in isolation as *what would actually happen* if we were to somehow equalize Black-white exposure to neighborhood poverty but not allow any of the other effect pathways to operate. But we would agree with arguments made by Robinson & Bailey (2019) that precise quantitative identification of such hypothetical future causal effects is not a prerequisite for supporting a broad policy agenda aimed at dismantling pathways of structural racism affecting health, an agenda based on retrospective quantitative and qualitative causal triangulation, a robust theory describing the system of racism, and longstanding social movements (Bailey et al. 2017; Darity and Mullen 2020; Ford and Airhihenbuwa 2010; Robinson and Bailey 2019; Taylor 2016).

Conclusions

In this paper we provided a worked example of a life course mediation framework for decomposing racial disparities in health outcomes, but this framework can be generalized to the decomposition of any population disparity. We have clarified the relative importance of several important structural mediators within the system of racism experienced by the Add Health cohort, such as the importance of *emergent discrimination* across the life course via disproportionate exposure to neighborhood poverty experienced by those racialized as Black and *racial discrimination* via the system governing the relationship between educational attainment and adult health. However, substantial assumptions and limitations of this analysis (and, we argue, any quantitative

analysis of racial disparities) underscores the crucial importance of causal triangulation across methods and the notion of "inference to best explanation" (Krieger and Smith 2016a). In one of his many examples in studying the causal process connecting racism and racialization to health and well-being, Du Bois (1898) described the need for triangulation across 1) historical study, 2) statistical investigation, 3) anthropological measurement, and 4) sociological interpretation (Du Bois 1898). Parametric quantitative inference, predicated on estimating unobserved counterfactuals from observed data under strong assumptions of full or partial exchangeability, is only one mode of causal inference; it is not *the definition* of causal inference (Schwartz et al. 2017).

Much of modern quantitative causal inference in the study of health pushes research questions towards "well-defined interventions" seen as more "proximal" to the individual, which reinforces a neoliberal paradigm of social change predicated on marginal interventions that do not fundamentally threaten structural or institutional arrangements (Robinson and Bailey 2019; Schwartz, Prins, et al. 2016). At the same time, conventional regression models used to study mediating pathways of structural racism are subject to time-varying confounding issues, which, at worst, can reify and essentialize notions of individual race as a separable "risk factor" from all other social relations. But an uncritical application of more complex methods risks losing sight of the fundamental causes that govern the distribution of more proximal risk factors, and ultimately the total disparity in population health (Bailey et al. 2017; Jackson and Arah 2019; Krieger 2014; Schwartz, Prins, et al. 2016; Schwartz et al. 2017). It is therefore our hope that this study can offer one framework for leveraging new developments in quantitative causal

inference to decompose a population disparity observed at a given time to the historical life course process through which individuals come to embody racist structural systems, supporting more holistic causal narratives rather than isolating effects of marginal changes (Bailey et al. 2017; Diez Roux 2012; Krieger 2001b; Reskin 2012).

Tables and Figures

Table 1. Decomposed effect estimates and interpretation (adapted from Wang & Arah 2015).

Effect	Research question / counterfactual interpretation
ATE (total disparity)	What if those racialized as Black had been racialized and treated as white by all measured and unmeasured mediating systems?
CDE (disparity without mediators)	What if those racialized as Black had instead been racialized and treated as white in the absence of all measured mediators? In other words, how did *unobserved* mediating pathways of racism produce racialized disparities?
PAI (disparity attributable to mediating interaction, e.g. via college attainment)	What if those racialized as Black had instead been racialized and treated as white by the underlying causal system (for example, the system connecting college attainment to health)? In other words, how did the system connecting educational attainment to health racialize individuals and discriminate based on those racial categories (via the *interaction* effect)?
PIE (disparity attributable to mediating main effect, e.g. via college attainment)	What if those racialized as Black had instead been racialized as white by the system producing the distribution of exposures (for example, college attainment) but still racialized as Black by the system connecting college attainment to health? In other words, how much of the impact of education in mediating the total racial disparity was because of racialized differences in attainment rates and a *main* effect of attainment on health? Reskin (2012) refers to this pathway as *emergent discrimination*.

Table 2. Descriptive statistics for all mediators and outcome by self-identified race.

	White		Black		
	Mean or prop.	SE	Mean or prop.	SE	*p*-value
Adolescence (ages 12-18)					
Age (years)	15.90	(0.02)	16.18	(0.04)	***
Sex (1 = male)	0.49	(0.01)	0.50	(0.01)	***
Parent college attainment (1 = at least one parent completed college)	0.24	(0.01)	0.16	(0.01)	***
Parent household income (US$)	$51,666	(474)	$30,496	(811)	***
Tract-level poverty rate (cont.)	0.11	(0.00)	0.26	(0.00)	***
Transition to adulthood (ages 18-25)					
Age (years)	22.09	(0.02)	22.37	(0.04)	***
Tract-level poverty rate (cont.)	0.13	(0.00)	0.23	(0.00)	***
Young adulthood (ages 25-32)					
Age (age)	28.92	(0.02)	29.22	(0.04)	***
Tract-level poverty rate (cont.)	0.13	(0.00)	0.23	(0.00)	***
College attainment (1 = respondent completed college)	0.33	(0.01)	0.22	(0.01)	***
Household income (US$)	$63,648	(519)	$43,246	(844)	***
Cardio-metabolic risk (cont.)	-0.08	(0.02)	0.27	(0.03)	***
Person-years	25,888		9,741		
Unique individuals	10,052		3,893		

Weighted estimates. *p*-value of two-sided t-test: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Table 3. G-formula effect decomposition.

Effect	Mean	Confidence interval	% of total disparity
ATE (total disparity)	-0.33	(-0.39, -0.26)***	100
CDE (disparity without mediators)	-0.14	(-0.26, 0.04)**	42
PIE (Parent income, Adolescence)	0.00	(-0.05, -0.06)	0
PIE (Parent college attainment, Adolescence)	0.00	(-0.01, 0.02)	0
PIE (Cumulative tract-level poverty)	-0.07	(-0.12, 0.02)***	21
PIE (College attainment, Young adulthood)	-0.01	(-0.03, 0.01)	3
PIE (Household income, Young adulthood)	0.00	(-0.05, 0.05)	0
PAI (Parent income, Adolescence)	-0.01	(-0.02, 0.00)	3
PAI (Parent college attainment, Adolescence)	-0.05	(-0.09, -0.01)*	15
PAI (College attainment, Young adulthood)	-0.06	(-0.11, -0.01)*	18
PAI (Household income, Young adulthood)	0.00	(-0.01, 0.01)	0

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Figure 1. Directed-acyclic graph describing how the connection between Race and Health (*Y*) is mediated by material deprivation, denoted *M*. Includes confounders of $M \rightarrow Y$ denoted *L*, and Arrow (b) includes an interaction between Race and *M* in influencing *Y*.

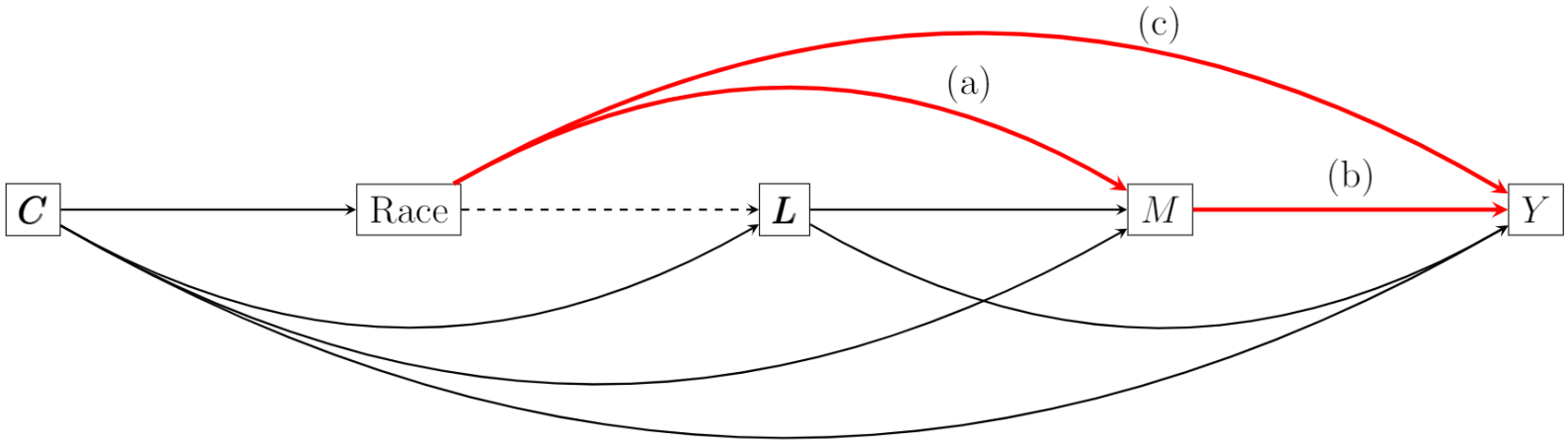
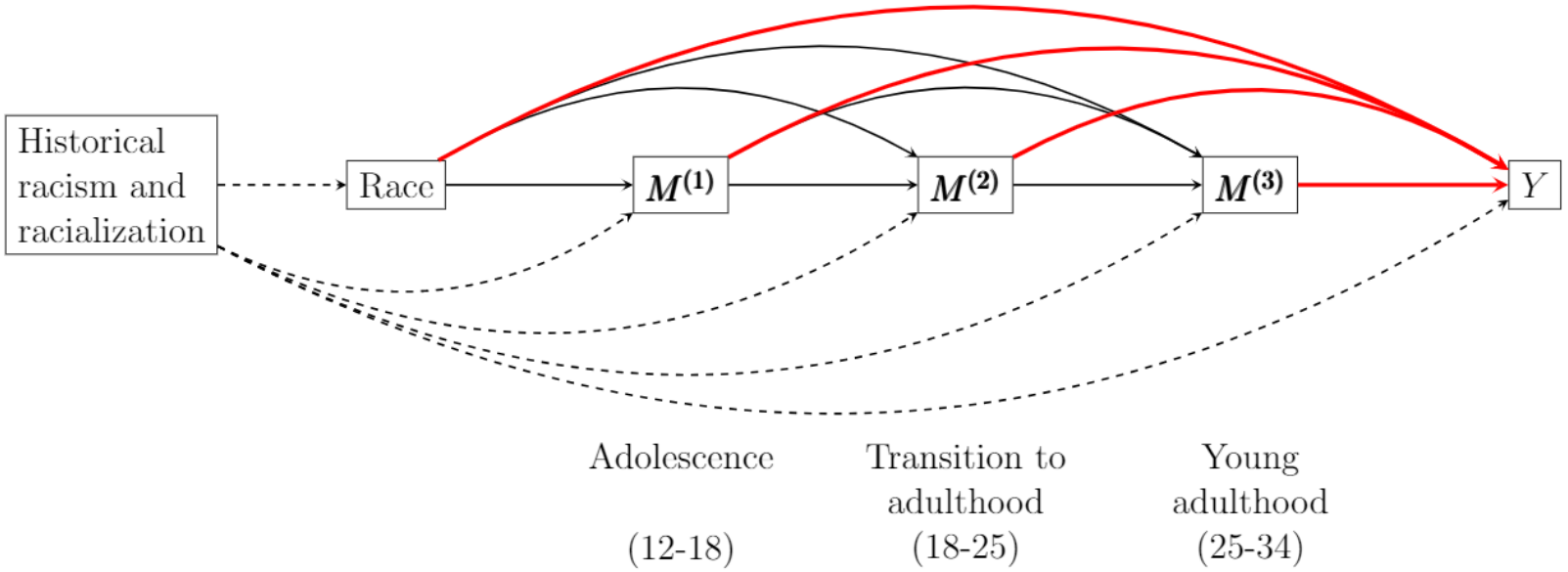
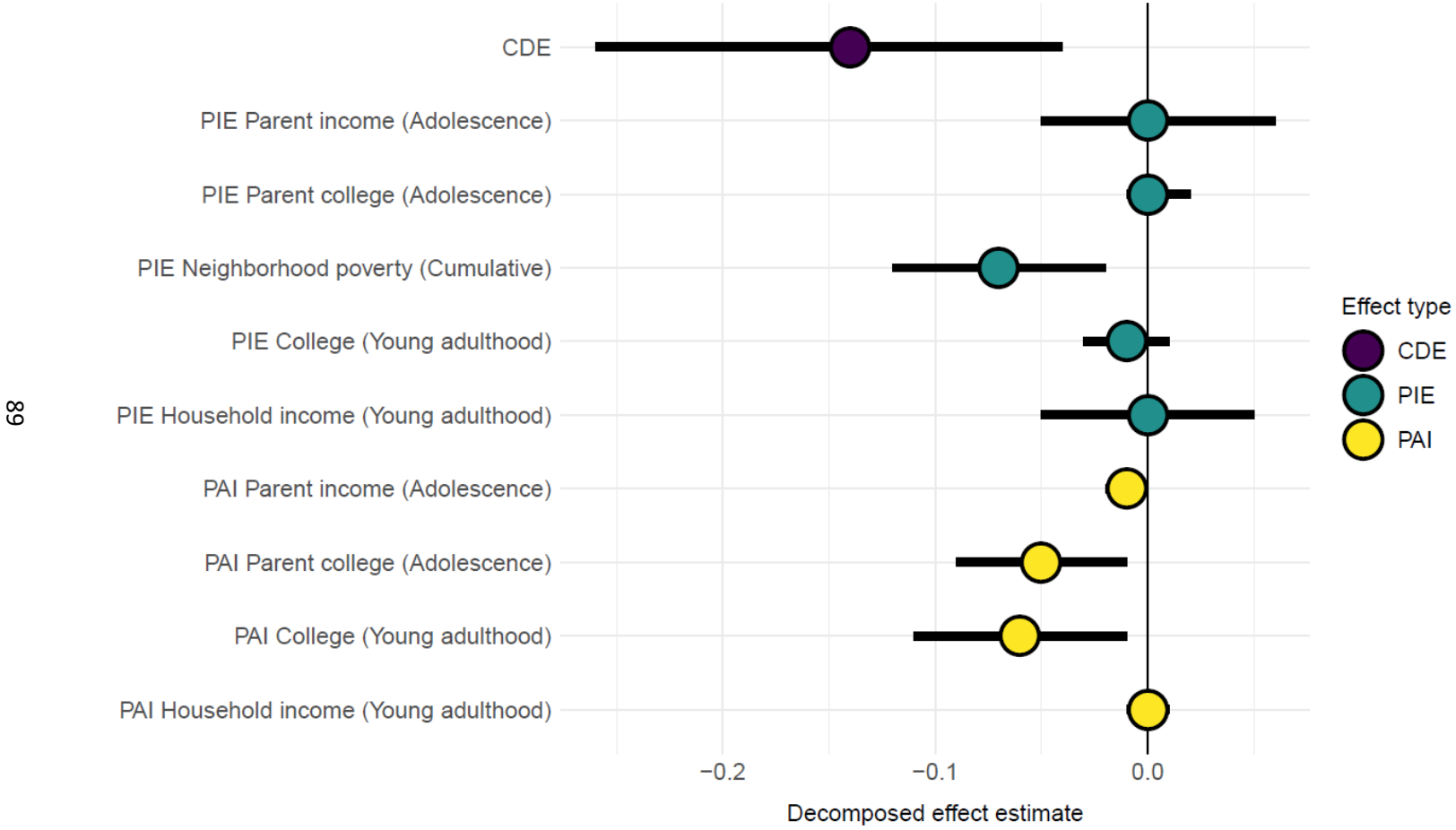


Figure 2. Directed-acyclic graph describing the relationship between race and cardio-metabolic risk by adulthood.



Note: For visual clarity, two pieces of information are suppressed in this DAG compared to Figure 1: 1) Multiple observed mediators measured at the same time point are suppressed to a single node and set of arrows (e.g. $\boldsymbol{M}^{(1)}$) and 2) all arrows except those originating from Race contain an interaction with Race (e.g. the effect of college attainment on household income varies by racialized category).

Figure 3. Decomposition estimates for the total Black-white disparity in cardiometabolic risk



Notes: The x-axis indicates the counterfactual change in the normalized cardio-metabolic risk index (i.e. standard deviations) attributable to that pathway had the population racialized as Black instead been treated as the population racialized as white by all mediating systems.

CHAPTER 3

Historical mechanisms connecting home values over three generations to contemporary Black-white disparities in wealth

Nick Graetz[1]

[1]Department of Sociology, Population Studies Center, University of Pennsylvania

Abstract The Black-white wealth gap in the United States has persisted and widened since the 1960s. Descriptive and qualitative analyses have identified many mechanisms underlying wealth correlations across successive generations, but few studies have quantified the relative contributions of these interconnected and racialized systems of reproduction to the total gap we observe today. Using linked intergenerational data from the PSID (N=3,077), I define a wealth gap in 2015-17 between the grandchildren of those racialized as Black and the grandchildren of those racialized as white in 1968-70. I use a fully interacted counterfactual mediation framework to decompose this disparity into the historical, racialized contributions of 1) effects of home values in 1968-70 on home values in successive generations and 2) effects via educational attainment in successive generations. Findings from this study contribute to our understanding of the dynamic, racialized process of multigenerational place-based wealth accumulation and support the importance of historically contingent social policy centered on reparative justice.

There is an expanding literature focused on the production of wealth inequality within cohorts and the multigenerational pathways connecting wealth across cohorts (Killewald, Pfeffer, and Schachner 2017). Recently there has been an increased focus on racial wealth disparities in the United States, particularly across those populations racialized as Black compared to those racialized as white, and the role of wealth in the reproduction of racial stratification across many social, health, and economic domains (Boen, Keister, and Aronson 2020; Boen and Yang 2016; Darity and Mullen 2020; Hamilton and Neighly 2019; Houle and Addo 2019; Oliver and Shapiro 2006; Pfeffer and Killewald 2018). Most of the mechanisms hypothesized to explain the persistence of the

Black-white wealth gap focus on the momentum of extreme dispossession of wealth and resources from families racialized as Black since slavery and the continued exclusion of Black households from the financial instruments necessary for wealth accumulation (Darity and Mullen 2020; Hamilton and Neighly 2019; Logan and Parman 2017; Zewde 2020). Importantly, these racist arrangements organized through public-private partnerships have played an integral role in reifying and maintaining racial boundaries in the United States (Roberts 2011; Sewell 2016; Sugrue 1996; Taylor 2019; Zuberi 2001).

Many quantitative studies have focused on *homeownership* as a primary mechanism through which wealth gaps are built – but not all homeownership is equally effective for wealth accumulation (Flippen 2004; Howell and Korver-Glenn 2020; Pattillo 2013; Sugrue 1996; Taylor 2019). Since federal housing policy greatly expanded homeownership as an instrument of private wealth accumulation in the New Deal and post-war era, the ability to build home equity has been a critical component of total household wealth across the life course (Killewald et al. 2017). This involves both accessing high value homes and the cumulative appreciation of home values with each successive year of ownership (Killewald and Bryan 2016). Understanding *home values* as socially constructed within a system of racism and white supremacy is critical for quantifying their impacts on intergenerational racial wealth stratification (J. W. Faber 2020; Howell and Korver-Glenn 2018, 2020; Rugh and Massey 2010; Sewell 2016; Zaimi 2020). The racialized system of building (and protecting) property values has often been organized around schools, particularly given the historical and contemporary connections between property values and school funding (Geismer 2015; Jackson 1987; Sugrue 1996; Taylor 2019). In conceptualizing pathways of intergenerational wealth

transmission, treating homeownership as a dichotomous mediator connecting grandparents to grandchildren may mask significant heterogeneity in the distribution of *home values* and the processes of transmission across racialized groups. These processes may have historically operated through direct pathways across generations (e.g., inheritance, gifts, bequests), but also through the indirect ways in which wealth is used to position those in the next generation to access the tools of wealth accumulation – such as the political economy connecting property values to educational attainment in the United States, with higher education translating to future wealth in ways that are also racialized (Emmons and Noeth 2015; Pfeffer and Killewald 2018, 2019).

A relational, intergenerational perspective is also important in describing patterns for those racialized as Black whose families have lived in the United States and been exposed to racist housing systems for *multiple generations*. In measuring the Black-white wealth gap based on repeated cross-sections and monolithic racial categories, many studies of wealth trends and mechanisms obscure how processes of racialization and the composition of the population racialized as Black has shifted over time in the United States. For example, there has been an increase in international migrants racialized as Black – especially over the past two decades – that have very different socioeconomic profiles than their native-born counterparts (Hamilton 2019, 2020). This foreign-born population racialized as Black attains higher homeownership rates and higher home values than the native-born population racialized as Black (Tesfai 2016).

In the present study, I examine *home value* and *educational attainment* as intergenerational mediators connecting the racialized status of grandparents in 1968-70 to

total wealth for their grandchildren in 2015-17. Using linked cohort data across three generations, I use a counterfactual mediation framework to examine grandparents' relative home value rank as a fundamental intergenerational determinant of contemporary racialized wealth gaps for grandchildren. I describe a novel sequential application of the mediational g-formula (Lin et al. 2017; VanderWeele and Tchetgen Tchetgen 2017) to decompose the total Black-white disparity across grandchildren into 1) pathways through which home values in previous generations influenced home values in subsequent generations and 2) pathways through which home values influenced educational attainment in subsequent generations. This decomposition of a contemporary wealth gap is based on fully interacted models by grandparents' racialized category in 1968-70, allowing estimation of effect pathways that are attributable to both *racialized returns* to intergenerational home values and educational attainment as well as their *main effects*, or what Reskin (2012) terms "racial discrimination" and "emergent discrimination" in describing the mutually reinforcing systems of structural racism (Reskin 2012). I quantify how racism in the historical systems distributing home value and education can be structurally embedded in subsequent relations, even if these relations are not themselves explicitly racialized (Bonilla-Silva 2009). I discuss how this distinction allows us to quantify the relative magnitude of different pathways through which structural racism has built the contemporary wealth gap through historical and ongoing racial discrimination in the systems distributing housing and education, as well as through the momentum of past compositional inequalities (Bonilla-Silva 1997, 2009; Reskin 2012).

Background

Housing policy since the New Deal

The median wealth of white households ($171,000) is roughly ten times higher than that of Black households ($17,100) (Darity et al. 2018; Hamilton and Neighly 2019; Thomas et al. 2020). Studies have linked different features of the widening Black-white wealth gap to inequalities in social mobility and the racialized concentration of goods and services, such as high-quality jobs, high-value homes, and well-funded schools (J. W. Faber 2020; Sharkey 2013; Sugrue 1996). Focusing especially on the financialization of the global economy over the past 50 years, scholars have described how Black households have been systematically excluded from the ever-expanding instruments of wealth accumulation, which is an intergenerational process (Hamilton and Neighly 2019; Pfeffer and Killewald 2019; Rugh and Massey 2010).

A critical instrument for private wealth accumulation in the United States is homeownership, which has been widely studied as a binary indicator in the social stratification literature (Killewald et al. 2017). But not all homeownership is equal, and the impact of homeownership on family wealth – especially over generations – is tied to the social process of property valuation and the ability to build home equity (Flippen 2004; Howell and Korver-Glenn 2018, 2020; Korver-Glenn 2018a; Taylor 2019; Zaimi 2020). The history of federal housing policy in the United States has been explicitly racist, particularly around federal subsidies for homeownership in the New Deal and post-war era and the related public-private partnerships arranging the construction and distribution of property values (J. W. Faber 2020; Howell and Korver-Glenn 2018; Rothstein 2017; Sewell 2016; Sugrue 1996; Taylor 2019). The animating ideology of

these arrangements is rooted in the centuries-long history of chattel slavery and dispossession of Black wealth and property in the United States, but a highly salient feature in the most recent century is the "redlining" segregationist logics of the 1930s (J. W. Faber 2020; Rothstein 2017). The Homeowners' Loan Corporation (HOLC) created color-coded maps of majority-Black neighborhoods as undesirable, and the Federal Housing Administration (FHA) refused to insure mortgages in and around those neighborhoods. This arrangement accelerated a deep conflation between race and financial risk in the expanding market of federally subsidized private homeownership (J. W. Faber 2020; Howell and Korver-Glenn 2018; Rugh and Massey 2010). The Housing and Urban Development Act of 1968 coupled with the Fair Housing Act was intended to end this practice and expand Black homeownership over the following decades, but real estate practices of racial exclusion persisted and expanding access to credit continued to support racially tiered housing markets (J. Faber 2020; J. W. Faber 2020; Taylor 2019).

The late 1960s and 1970s were also characterized by a system of "predatory inclusion," whereby individuals racialized as Black continued to not only be systematically *excluded* from certain neighborhoods, but also newly *included* in substandard private housing markets (Taylor 2019). This extended into the development of the dual mortgage market in the United States that remains intact today, a delivery system for residential loans serving lower-income borrowers and minority consumers with a different mix of products than commonly serve higher-income markets (Sewell 2016). Racial minorities, especially those racialized as Black, face a higher risk for being provided credit products that are of poor quality, such as predatory, subprime, and less

regulated loans (Sewell 2016). Indeed, neighborhood racial composition was a stronger determinant of appraised home values in 2015 than it was in 1980 (Howell and Korver-Glenn 2020). Black homeownership nationally has not substantively increased since the Fair Housing Act of 1968, and the overall Black-white wealth gap has only widened (Hamilton and Neighly 2019; Pfeffer and Killewald 2019).

Mechanisms of intergenerational transmission

Recent studies adopt a racialized life-course of wealth accumulation and identify various mechanisms that contribute to the reproduction of the racial wealth gap from early adulthood to retirement (Boen et al. 2020; Houle and Addo 2019; Houle and Warner 2017; Killewald and Bryan 2018; Thomas et al. 2020). Still, the production and reproduction of wealth disparities is inherently multigenerational, extending beyond the correlation of parents' and children's wealth outcomes (Pfeffer and Killewald 2018, 2019; Thomas et al. 2020). This is in part due to the ways in which wealth is directly transferred across generations, but also the ways in which wealth is used to position those in the next generation to access the tools of wealth accumulation.

One important mechanism noted descriptively by Pfeffer & Killewald (2018) is the system through which wealth shapes educational opportunity for children and grandchildren. Educational attainment is in turn tied to occupation, earnings, and homeownership through exchange in racialized labor and housing markets. College attainment has greatly increased within the population racialized as Black since the 1960s, but Seamster & Charron- Chénier (2017) note that this expansion may also

represent a form of "predatory inclusion" similar to Taylor's (2019) description of how the expansion of Black homeownership occurred only through expanding access to substandard credit and property. In many ways, Black Americans have been granted access to post-secondary institutions on exploitative financial terms (Seamster and Charron-Chénier 2017). As profit has been historically maximized in the housing sector via public-private arrangements that burdened Black families with toxic debts, a similar process has unfolded in the burgeoning financialized economy of student debt. Black young adults hold substantially higher levels of student loan debt than similar white young adults (Addo, Houle, and Simon 2016; Houle and Addo 2019; Seamster and Charron-Chénier 2017), which also delays the transition to marriage and homeownership (Addo 2014; Killewald et al. 2017).

Taken together over the past century, the distribution of home values and access to post-secondary institution has developed within a racialized system of *systemic exclusion* and *predatory inclusion*. This intergenerational system results in Black families "swimming upstream" (Stewart 2008) in two ways: 1) being blocked from accessing high-value homes and high-quality post-secondary institutions, and 2) even if gaining access, seeing lower wealth returns from home values and college attainment than those racialized as white (Emmons and Noeth 2015; Flippen 2004). A quantitative intergenerational account of these racialized and entangled effect pathways, particularly their importance relative to one another, can contextualize our empirical understanding of contemporary wealth statistics as the sum of intergenerational histories.

Challenges to estimating intergenerational effects

Common quantitative techniques used to "explain" racialized disparities in outcomes include demographic decomposition (e.g., Kitagawa or Oaxaca-Blinder decomposition) and counterfactual mediation methods based on regression or weighting estimators (e.g., Baron-Kenny mediation) (Jackson and VanderWeele 2018; Lundberg, Johnson, and Stewart 2021; VanderWeele 2016). There are limitations to these techniques, especially when applied in the study of racialized disparities which are necessarily the cumulative result of racism and racialization entangled across many related time-varying exposures (Esposito 2019; A. I. Naimi et al. 2016a; Sen and Wasow 2016; Zuberi and Bonilla-Silva 2008; Zuberi et al. 2015). Williams (2019) recently reviewed the theoretical and empirical problems inherent in "explaining away" portions of the total racial disparity in an outcome over time with conventional decomposition methods (Iceland 2019; Williams 2019), including failing to contextualize the historical social and political construction of race in the United States and the treatment of mediating socioeconomic variables (e.g. home value or education) as "race-neutral."

As demonstrated by Jackson & VanderWeele (2018) and recently extended by others, decomposition of a population disparity can be framed as a counterfactual mediation analysis (Jackson and VanderWeele 2018; Lundberg et al. 2021; A. I. Naimi et al. 2016b; Sudharsanan and Bijlsma 2019). Bauer & Scheim (2019) and Jackson & VanderWeele (2019) discuss how these techniques can account for the complex patterns of time-varying confounding and interaction necessary for studying a dynamic social *process* over time, which is important for addressing the critiques of Zuberi & Bonilla-Silva (2008), Sen & Wasow (2016), and Williams (2019) when it comes to quantifying

dynamic systems of racism and racialization in a counterfactual framework. In terms of describing the production of racialized disparities over time, such a framework can include a self-reported race variable more explicitly as proxying a set of social relations rather than an individual characteristic by addressing issues of "post-treatment bias." This reframing is more than conceptual; it requires carefully describing counterfactual comparisons and their underlying assumptions (Esposito 2019; Kohler-Hausmann 2019; Sen and Wasow 2016).

Describing the racial disparity in terms of a counterfactual mediation analysis also allows for decomposing the total disparity into several theoretically important and distinct intergenerational pathways. Consider the directed acyclic diagram (DAG) in Figure 1 describing connections between parents' racialized status ($R^{(1)}$) and children's wealth by adulthood ($Y^{(2)}$) across two generations. We can consider the theoretical justification for each arrow, supported by previous theoretical and empirical work:

- Arrow (a) exists because those parents racialized as Black are more likely to have lower home values (or more likely to be renters) than those parents racialized as white. As discussed above, this is driven by factors such as systemic discrimination in the housing market and the racist dynamics of retrospective home appraisal systems (Howell and Korver-Glenn 2020; Sewell 2016).
- Arrow (b) exists because parents' home value influences children's educational attainment. There is likely a main effect driven by the coupling of property values to public school funding across the United States. There is likely also an interaction with parents' racialized status, where relative home value may influence the likelihood of

achieving different levels of educational attainment more or less for those racialized as Black compared to white, for example due to racist patterns of place-based public disinvestment.

- Arrow (c) exists because the distribution of educational attainment is racialized via many other unmeasured pathways of interpersonal and structural racism operating outside the system in Arrow (a) distributing relative home value and the system in Arrow (b) distributing the returns to those home values on educational attainment.
- Arrow (d) exists because children's wealth is affected by parents' relative home values via many other unmeasured pathways operating outside educational attainment, likely in racialized ways.
- Arrow (e) exists because children's educational attainment influences children's wealth. There is likely a main effect driven by the ways education is generally tied to higher-wage employment, which in turn is tied to the ability to save, access lending, and more. There is likely also an interaction with racialized status, where all of these relations racialize individuals and operate on them in racist ways. For example, a college degree has a smaller return on income and health for individuals racialized as Black compared to individuals racialized as white (Boen 2016a; Emmons and Noeth 2015; Emmons and Ricketts 2017).
- Arrow (f) exists because parents' history and lived experience in a system of racism affects their children's wealth through many pathways that are not captured by the broad structural systems of racial stratification occurring in housing and education.

Taken together, all arrows in Figure 1 therefore summarize the many *measured* and *unmeasured* pathways through which racism and racialization in a specific time and place can be connected to future outcomes within the life course and across generations, or what Quincy Stewart described as "swimming upstream" for those racialized as Black in the United States (Stewart 2008). The total gap in wealth between the children of those racialized as Black compared to the children of those racialized as white is defined by arrangements that are explicitly racialized, such as discrimination in the underlying system through which educational attainment is translated to increased wealth; the interactive effect through Arrow (e). But this history also includes the fact that home values are connected to educational attainment *at all*, and the distribution of home values is racialized; Arrow (a) plus the main effect of Arrow (b). Both of these connections are part of Reskin's (2012) total "race discrimination system" underlying the historical production of a racialized disparity in a given outcome in a particular time and place (Reskin 2012). These dynamics cannot be operationalized with conventional regression models, which assume that post-treatment variables confounding the relationship between a given mediator and the outcome are not themselves influenced by the treatment or correlated over time (Esposito 2019; Jackson and VanderWeele 2018; A. I. Naimi et al. 2016a).

Translating theory to empirical estimands

For the reasons discussed above, operationalizing this mediation analysis and assumptions therein must be consistent with our theory of historical racism as a relational process. The total racialized disparity in children's wealth ($Y^{(2)}$) in Figure 1 can be

decomposed into the following quantities, each with a distinct interpretation (Vanderweele 2014; VanderWeele and Tchetgen Tchetgen 2017; Wang and Arah 2015):

- The proportion attributable to interaction (PAI) via *H*: The sum of the effect pathways linking race to home values with the *interactive effect* of Arrow (d). In other words, systems of racism and racialization influence the distribution of home values, and home values then influence *Y* in racialized ways.
- The pure indirect effect (PIE) via *H*: The sum of the effect pathways linking race to home values with the *main effect* of Arrow (d). In other words, systems of racism and racialization influence the distribution of home values, and home values then influence $Y^{(2)}$ regardless of racialized status.
- The proportion attributable to interaction (PAI) via *M*: The sum of the effect pathways linking race to educational attainment and the *interactive effect* of Arrow (e). In other words, systems of racism and racialization influence the distribution of educational attainment, and educational attainment then influences $Y^{(2)}$ in racialized ways.
- The pure indirect effect (PIE) via *M*: The sum of the effect pathways linking race to educational attainment and the *main effect* of Arrow (e). In other words, systems of racism and racialization influence the distribution of educational attainment, and educational attainment then influences $Y^{(2)}$ regardless of racialized status.

- The controlled direct effect (CDE): All the unobserved pathways of racism and racialization influencing $Y^{(2)}$ via Arrow (f) that do not operate through the observed mediators above.

Figure 1 is useful in describing the identification challenges for decomposing the total racial disparity into these effects. In Figure 1, $H^{(1)}$ acts as both a *mediator* of $R^{(1)} \rightarrow Y^{(2)}$ and as a *confounder* of $M^{(2)} \rightarrow Y^{(2)}$ that is influenced by $R^{(1)}$. In this situation, $H^{(1)}$ is often termed a "treatment-induced mediator-outcome confounder" with respect to $M^{(2)} \rightarrow Y^{(2)}$ (A. Naimi et al. 2016; Vanderweele et al. 2014; Wodtke and Parbst 2017; Wodtke and Zhou 2020). This creates a difficult counterfactual interpretation if $H^{(1)}$ and $M^{(2)}$ are both included in a conventional regression model predicting $Y^{(2)}$. Analyses using a regression model to assess mediation (often referred to as the Baron-Kenny estimator) – or mediation analyses fitting separate regression models for the both the mediator and outcome – rely on the assumption that all mediator-outcome confounders are *not* associated with the exposure of interest – often termed the "cross-world independence assumption" (A. Naimi et al. 2016; Wang and Arah 2015). In other words, these estimators assume that Arrow (a) does not exist in Figure 1 in order to identify the indirect effects via $M^{(2)}$. Identifying the indirect effects via $M^{(2)}$ requires conditioning on $H^{(2)}$, but in doing so we inadvertently block the mediating pathway through $H^{(2)}$. Using a conventional mediation analysis when this assumption is violated risks dramatically overestimating the mediating impact of $M^{(2)}$ while also corresponding to a counterfactual this is not theoretically coherent (Jackson and VanderWeele 2018; A. I. Naimi et al. 2016a; Naimi 2016; Wang and Arah 2015; Wodtke and Parbst 2017).

Present study

I describe and decompose the historical, intergenerational process of home value attainment in contributing to the gap in total wealth between the grandchildren of those observed in 1968 – the year of the Fair Housing Act and first year of data collection in the Panel Study of Income Dynamics (PSID). The first analysis is descriptive, defining the total racialized disparity in 2015-17 home values by the self-identified race of grandparents in 1968-70 (i.e., the disparity between those who had a grandparent who identified as Black compared to those who had a grandparent who identified as white). I use flow diagrams to examine upward and downward mobility patterns between generations. Here I contextualize the magnitude of these gaps in not only homeownership, but the racialized distribution of home values across owners since 1968-70.

The second analysis focuses on explanation of these gaps by extending Figure 1 across three generations in Figure 2. I use a novel sequential application of the mediational g-formula (VanderWeele and Tchetgen Tchetgen 2017; Wang and Arah 2015) across these three generations to examine the relative importance of mediating pathways through home values and educational attainment, drawing on concepts from Reskin's (2012) "race discrimination system." These pathways include *unobserved mediation*, or the proportion of the total disparity that is not explained by observed intergenerational mediation via home values and educational attainment (i.e., the controlled direct effect; CDE), *racial discrimination via each mediator* (i.e., the

proportion attributable to interaction; PAI), and *emergent discrimination via each mediator* (i.e. the pure indirect effect; PIE) (Reskin 2012).

Data and Methods

I construct a multigenerational dataset beginning with PSID respondents in 1968-70 (G1), their children in 1995-97 (G2), and their grandchildren in 2015-17 (G3). I first describe intergenerational trends in year-specific relative home value ranks between those racialized as Black in 1968-70 and those racialized as white to assess general trends in upward and downward mobility across generations. I then use the mediational g-formula to decompose the total disparity in mean net worth across the grandchildren of Black and white grandparents from 1968-70 into a controlled direct effect and proportions attributable to interaction (PAIs) and pure indirect effects (PIEs) via five broad mediating mechanisms: grandparents' home values, parent's educational attainment, parents' home values, grandchildren's educational attainment, and grandchildren's home values. I apply this mediation analysis *sequentially*, treating each subsequent mediator as the outcome, to examine how previous intergenerational effect pathways become embedded in the distributions of subsequent mediators (Figure 2):

1. How did G3 home values shape G2 educational attainment?
2. How did G2 education shape G2 home values? Are there residual mediating effects from G1 home values?
3. How did G2 home values shape G3 education? Are there residual mediating effects from G1 home values and G2 education?

4. How did G3 education shape G3 home values? Are there residual mediating effects from G1 home values, G2 education, and G2 home values?
5. How did G3 home values shape G3 total wealth? Are there residual mediating effects from G1 home values, G2 education, G2 home values, and G3 education?

Multigenerational Data

Linked multigenerational data comes from the Panel Study of Income Dynamics (PSID). The PSID is the longest running nationally representative, longitudinal survey of individuals and families in the United States. The survey began in 1968 and included two independent samples: a nationally representative sample of families identified by the Survey Research Center at the University of Michigan and an oversample of low-income families drawn from the Survey of Economic Opportunity conducted by the Census. Together, these samples constituted a nationally probability sample of U.S. families. The PSID has interviewed respondents and their family members continuously since 1968, including biennial interviews since 1997.

In addition to describing multigenerational home value trajectories, my primary explanatory goal is to decompose the total 2015-17 disparity in mean net worth between the grandchildren of those racialized as Black in 1968-70 and the grandchildren of those racialized as white. I link children to their biological or adoptive parents and grandparents using the PSID's family identification mapping system (FIMS). I pool grandparents' self-reported home values in 1968-70, parents' self-reported home values in 1995-97, and

grandchildren's home values and total wealth in 2015-17. I subset to ages 25-64 within each generation. I choose these periods to capture 1) those aged 25-64 at the beginning of the PSID, 2) the centered years in which their children are roughly aged 25-64, and 3) the most recent period of observation. Following Pfeffer & Killewald (2018), if only a single grand(parent) is observed, I use their self-reported home value as the sole indicator of (grand)parental home value. For grandparents or parents who did not live in the same household in 1968-70 or 1995-97 (e.g. because they are divorced), I average home values across the separate households (Pfeffer and Killewald 2018). This yields a final sample of 3,077 unique grandchildren in 2015-17 with observed parental and grandparental mediators.

All analyses are weighted using the 1968 PSID family weight and standard errors are clustered by 1968 family using the *survey* R package.

Outcome

The primary outcome is summary net wealth in G3. To capture the full range of variation in negative total wealth values and for ease of interpretation, I use continuous inflation-adjusted wealth truncated at plus/minus $500,000. Sensitivity analyses using percentile ranks and the inverse hyperbolic sine (IHS) transformation produce substantively similar conclusions in mediation analyses.

Exposure

As described in Figure 2, I use grandparents' self-identified race to proxy the ways in which grandchildren's parents and grandparents were racialized by the

intermediate systems described by each arrow in Figure 2. I use this variable to define the total disparity in the decomposition analysis as between the grandchildren of those racialized as Black in 1968 compared to the grandchildren of those racialized as white.

Mediators

There are five ordered dependent mediators between grandparents' racialized category and the total wealth of grandchildren (Figure 2): grandparent home value percentile ($H^{(1)}$), parent educational attainment ($M^{(2)}$), parent home value percentile ($H^{(2)}$), grandchild educational attainment ($M^{(3)}$), and grandchild home value percentile ($H^{(3)}$). Continuous home values are adjusted for inflation and converted to period-specific percentiles (0 = renter). I discretize these into quartiles (0 = renter, 1 – 4 = quartiles) for descriptive tables and figures. Educational attainment is measured using a four-category variable (less than high school, high school, some college, college attainment or more). As described below, all mediator and outcome models also include intermediate controls for the age and sex of respondents in a given generation. In the final analyses, I operationalize educational attainment using a binary indicator for college attainment (0/1), home value using continuous period-specific percentiles (0-100), binary sex (0/1), and continuous age. In sensitivity analyses, I tested every combination of the following: multinomial home value (0 = renter, 1 – 4 = quartiles), multinomial educational attainment (less than high school, high school, some college, college attainment or more), age squared, and various splines for age. Effect estimates and conclusions are substantively similar across all versions of the mediator and outcome

models, so I report the most parsimonious version for ease of interpretation (continuous home value percentiles, binary college attainment).

G-computation with multiple dependent mediators

Conventional regression for mediation analyses, controlling for post-treatment characteristics, assumes that there exists no exposure-induced mediator-outcome confounding; in other words, that the cross-world independence assumption holds (Daniel et al. 2015; Imai et al. 2011; Robins et al. 2000; Wang and Arah 2015). The counterfactual comparison is between two populations that vary on exposure status (being racialized as Black compared to being racialized as white in a specific system of racism), assuming *nothing else* changes across the two populations over time. As described above, theoretical and empirical research has demonstrated this assumption is untenable in the quantitative study of racism as a social process, where racism affects virtually all other observed variables over time (Esposito 2019; A. I. Naimi et al. 2016a; Sen and Wasow 2016; Zuberi and Bonilla-Silva 2008).

The g-formula is a method of standardization that allows for the estimation of unconfounded summary effects without relying on the restrictive cross-world independence assumption inherent to conventional regression estimators (A. Naimi et al. 2016; Robins 1986; Wang and Arah 2015). Equation 1 illustrates the average net worth for grandchildren ($E(Y_3)$) given exposure level $A = a$ (whether grandparent was racialized as Black or white in 1968-70) in terms of confounders unaffected by the exposure (continuous age and binary sex in each generation; v_1, v_2, v_3) and five

dependent mediators (grandparents' home value, parent's education, parents' home value, grandchildren's education, grandchildren's home value; h_1, m_2, h_2, m_3, h_3). Subscripts indicate the generation and for all variables I use x as shorthand for $X = x$.

$$E(Y_3) = \sum_{h_3}\sum_{m_3}\sum_{h_2}\sum_{m_2}\sum_{h_1} E(Y_3|h_3, m_3, h_2, m_2, h_1, a, v_3) \times P(h_3|m_3, h_2, m_2, h_1, a, v_3) \times P(m_3|h_2, m_2, h_1, a, v_3) \times P(h_2|m_2, h_1, a, v_3) \times P(m_2|h_1, a, v_2) \times P(h_1|a, v_1) \times P(v_1)\, P(v_2)\, P(v_3) \quad (1)$$

This equation is analogous to the DAG described in Figure 2. When considering the effect on Y_3 of changes to a via a specific mediator h_2, an intermediate post-treatment variable such as m_2 is often referred to as an "exposure-induced mediator-outcome confounder" because it is *affected by the exposure* and also *confounds the relation* between h_2 and Y_3. The presence of such confounding means that we cannot estimate how the counterfactual Y_3 is influenced via h_2 while holding m_2 constant, because such a world would be impossible to observe. In terms of our theoretical construct of racialization within a system of racism, this reflects on the critiques discussed above from

Zuberi & Bonilla-Silva (2008) and others: what would it *mean* to consider the effect of being racialized one way vs. another *without* any other observed variables changing? At best, we are describing a marginal effect that is difficult to interpret because changing this exposure requires considering changes in everything else that is influenced by and acts on racialized status. At worst, we are reifying the notion that the variable for self-identified race proxies an individualized construct that can be considered (and manipulated) independently from other variables (Sen and Wasow 2016; Zuberi and Bonilla-Silva 2008).

Estimands

Using the g-formula from Equation 1, the total effect (TE) can be estimated using the counterfactual values in Equation 2 (i.e. g-computation) (Wang and Arah 2015):

$$TE = E(Y_{\bar{a}}) - E(Y_{\bar{a}^*}) \quad (2)$$

In handling post-treatment confounding, total effects estimated via g-computation are analogous to those obtained by applying inverse probability-of-treatment weights, or "marginal structural models" (Lee and Jackson 2017; Robins et al. 2000; Wodtke et al. 2011). The g-formula also provides a straightforward decomposition of this total effect or disparity into direct and indirect pathways of accumulation via multiple mediators (Lin SH et al. 2017; VanderWeele and Tchetgen Tchetgen 2017; Wang and Arah 2015). I extend the mediational g-formula introduced by VanderWeele & Tchetgen Tchetgen (2017) to include multiple dependent mediators in a *fully interacted* framework (Equation 3).

$$E(Y_3) = \sum_{h_3}\sum_{m_3}\sum_{h_2}\sum_{m_2}\sum_{h_1} E(Y_3|a \times (h_3, m_3, h_2, m_2, h_1, v_3)) \quad (3)$$
$$\times P(h_3|a \times (m_3, h_2, m_2, h_1, v_3))$$
$$\times P(m_3|a \times (h_2, m_2, h_1, v_3))$$
$$\times P(h_2|a \times (m_2, h_1, v_3))$$
$$\times P(m_2|a \times (h_1, v_2)$$
$$\times P(h_1|a \times (v_1)$$
$$\times P(v_1)\, P(v_2)\, P(v_3)$$

The total effect in Equation 2 (or the total observed racial disparity, as I do not condition on any "pre-treatment" variables) can then be decomposed into the controlled direct effect (CDE; *unobserved mediating pathways*), the proportion attributable to interaction via each mediator (PAI; *racial discrimination in the underlying system connecting the mediator to Y*), and the pure indirect effect via each mediator (PIE; *emergent discrimination*) (Bauer and Scheim 2019; Jackson and VanderWeele 2019; Reskin 2012; VanderWeele and Tchetgen Tchetgen 2017; Wang and Arah 2015). In the estimands defined below, $\bar{G}_{\bar{a}}^{(h_3)}$ corresponds to a series of random draws from the underlying distribution m_3 that would have been observed in the population with $\bar{A} = \overline{a^*}$ (for further discussion of this decomposition, see VanderWeele & Tchetgen Tchetgen 2017).

$$CDE = E(Y_{\bar{a}\bar{h}_3^*\bar{m}_3^*\bar{h}_2^*\bar{m}_2^*\bar{h}_1^*}) - E(Y_{\bar{a}^*\bar{h}_3^*\bar{m}_3^*\bar{h}_2^*\bar{m}_2^*\bar{h}_1^*}) \quad (4)$$

$$PIE^{(h_3)} = E(Y_{\bar{a}^*\bar{G}_{\bar{a}}^{(h_3)}\bar{G}_{\bar{a}^*}^{(m_3)}\bar{G}_{\bar{a}^*}^{(h_2)}\bar{G}_{\bar{a}^*}^{(m_2)}\bar{G}_{\bar{a}^*}^{(h_1)}}) - E(Y_{\bar{a}^*\bar{G}_{\bar{a}^*}^{(h_3)}\bar{G}_{\bar{a}^*}^{(m_3)}\bar{G}_{\bar{a}^*}^{(h_2)}\bar{G}_{\bar{a}^*}^{(m_2)}\bar{G}_{\bar{a}^*}^{(h_1)}}) \quad (5)$$

$$PAI^{(h_3)} = E(Y_{\bar{a}\bar{G}_{\bar{a}}^{(h_3)}\bar{G}_{\bar{a}^*}^{(m_3)}\bar{G}_{\bar{a}^*}^{(h_2)}\bar{G}_{\bar{a}^*}^{(m_2)}\bar{G}_{\bar{a}^*}^{(h_1)}}) - E(Y_{\bar{a}^*\bar{G}_{\bar{a}}^{(h_3)}\bar{G}_{\bar{a}^*}^{(m_3)}\bar{G}_{\bar{a}^*}^{(h_2)}\bar{G}_{\bar{a}^*}^{(m_2)}\bar{G}_{\bar{a}^*}^{(h_1)}}) \quad (6)$$

$$- E(Y_{\bar{a}\bar{G}_{\bar{a}^*}^{(h_3)}\bar{G}_{\bar{a}^*}^{(m_3)}\bar{G}_{\bar{a}^*}^{(h_2)}\bar{G}_{\bar{a}^*}^{(m_2)}\bar{G}_{\bar{a}^*}^{(h_1)}}) + E(Y_{\bar{a}^*\bar{G}_{\bar{a}^*}^{(h_3)}\bar{G}_{\bar{a}^*}^{(m_3)}\bar{G}_{\bar{a}^*}^{(h_2)}\bar{G}_{\bar{a}^*}^{(m_2)}\bar{G}_{\bar{a}^*}^{(h_1)}})$$

Equations 5 and 6 correspond to the pure indirect effect (PIE) and proportion attributable to interaction (PAI) via a specific mediator – respondent home value (h_3) – but are estimated analogously for each of the other four mediators: grandparent's home value (h_1), parent's education (m_2), parent's home value (h_2), and respondent's education (m_3).

Parametric estimation

I adapt the estimation algorithm for g-computation described by Lin et al. (2017) to include survey-weighted models incorporating the grandparents' 1968 family weight:

7. Fit survey-weighted generalized linear models for $Y_3, h_3, m_3, h_2, m_2, h_1$. Summary net worth ($Y$) and home value quartiles (h_3, h_2, h_1) are estimated with a linear regression model, while educational attainment (m_3, m_2) are estimated with a logistic regression model.

8. Draw a random multivariate-normal sample of parameters from all models using fitted coefficients and variance-covariance matrices (e.g., parametric bootstrap).
9. Create 30 replicates of the data to remove Monte Carlo error arising from stochastic individual-level response prediction. Use parameters from Step 2 to predict all conditional values in the g-formula (Equation 3) under two exposures: $A = a$ (individuals racialized as white) and $A = a^*$ (individuals racialized as Black).
10. Impute all counterfactuals in Equations 4-6 by drawing the necessary values from the simulations in Step 2, randomly permuting the distributions of h_3, m_3, h_2, m_2, h_1 to obtain $\bar{G}^{(h_3)}, \bar{G}^{(m_3)}, \bar{G}^{(h_2)}, \bar{G}^{(m_2)}, \bar{G}^{(h_1)}$. Calculate all effects using the formulas from Equations 4-6.
11. Repeat Steps 2-4 as many times as necessary (I create 1,000 draws of all effects) and then calculate the mean of these draws, as well as 2.5% and 97.5% quantiles to generate confidence intervals for all effects .

Identification assumptions

While avoiding the common assumption of no post-treatment bias or time-varying confounding via the *observed* variables, the effect estimates described in Equations 4-6 are unbiased only under the strong assumptions of positivity, no unobserved confounding in all mediator and outcome models, and no model misspecification. This analysis might therefore be interpreted more conservatively as a dynamic demographic decomposition of a total population difference (i.e., the total wealth gap between the grandchildren of those racialized as Black compared to those racialized as white in 1968) rather than a forward-

facing *causal* mediation analysis. This frame is aligned to the philosophy of "realized counterfactuals" and "reverse causal questions" – in other words, the goal of identifying possible causes of the effect (e.g., the total wealth difference between the grandchildren of those racialized as Black vs. white) rather than an effect of a cause (e.g., grandparents' racialized status) (Gelman and Imbens 2013; Krieger and Smith 2016b; Schwartz, Gatto, et al. 2016). The primary advantage of this counterfactual approach over other decomposition or mediation methods is the ability to correctly account for multiple dependent mediators and the related patterns of post-treatment time-varying confounding (including all interactions with racialized status) that can lead to misleading and biased results. In this way, the present analysis is a more dynamic and generalized version of Kitagawa or Oaxaca-Blinder decomposition (Jackson and VanderWeele 2018; Sudharsanan and Bijlsma 2019). Proofs for how this mediation analysis can be reduced to Kitagawa or Oaxaca-Blinder decomposition in the *absence* of post-treatment confounding or dependent mediators are provided in Jackson & VanderWeele (2018).

I first present descriptive statistics and figures illustrating the flow of home values across linked generations. I then present effect estimates from the sequential mediation analysis (Figure 2), first estimating how grandparent home value ($H^{(1)}$) mediates the relation between grandparent racialized status ($R^{(1)}$) and parent college attainment ($M^{(2)}$) and continuing by treating each subsequent mediator as the *outcome* until reaching grandchildren's total wealth in 2015-17. All analyses were conducted using R version 4.0.2.

Results

Table 1 presents weighted summary statistics by generation and self-reported race of grandparents. I collapse year-specific home value percentiles to quartiles. Black homeownership among those aged 25-64 in 1968-1970 was 38%, but the vast majority of this ownership was in low value housing stock. Only 7% of Black families owned homes valued in the upper half of all homes in those years. In contrast, 58% percent of white families owned homes in the upper half, including 33% owning homes in the top quartile. This gap remained largely intact in the second generation and seems to be reproduced again in the third generation, though this generation is younger at the time of observation (average age at observation for grandchildren was roughly six years younger than for parents and grandparents). Over the period 1968-2017 there was a large expansion of college attainment. In 1968-70, 7% of Black grandparents and 12% of white grandparents had completed college. These rates both improved with each successive generation, but the exponential increase for the children and grandchildren of white grandparents drove a widening racial disparity (23% vs. 52% in 2015-17). By 2015-17, the median summary net worth of grandchildren was four times higher for those with white grandparents ($41,440 vs. $10,022).

Figure 3 visualizes the flows of home value attainment from one generation to the next. In the first generation racialized as Black, the vast majority of homeownership existed in the bottom half of home values – particularly the bottom quartile. There was virtually no Black homeownership among the upper half of home values; in contrast, almost three quarters of white homeownership was in the upper half, including roughly 40% in the upper quartile. While the children of those renting were upwardly mobile in

the second generation for both Black and white families, there are large racial disparities in attainment. For example, roughly half of the children of Black renting families in the first generation were themselves renters by mid adulthood, compared to only one fifth of the children of white renting families. These processes were largely reproduced between the second and third generations. White families attained and have held high home values across all three generations, while Black families have only made marginal gains in value attainment. Any gains Black families made in homeownership by the second generation have been reversed by the third generation (2015-17), where only 34% own their homes at the time of observation. While Figure 3 summarizes intergenerational flows between grandparents-parents and parents-grandchildren, Figure 4 visualizes every specific intergenerational trajectory across all three generations. While some Black families have been upwardly mobile across generations, the most common intergenerational trajectory is renting in every generation. This group is much smaller for white families, where we can observe the general reproduction of home value attainment across generations at the upper end of the distribution.

Table 2 presents the results of the five sequential mediation analyses (all effect estimates are visualized in Figure 5). In the first mediation analysis (Mediation 1) predicting college attainment in the second generation, there is a 17.95 (95% confidence interval: 13.55 – 21.79, $p<0.001$) percentage point difference in attainment rate between the children of those racialized as white in 1968 compared to those racialized as Black. Of this total difference, 61% was predicted by the PIE via home value percentile in the first generation. In other words, the expected college attainment rate for the Black sample

in G2 would be 10.99 (2.97 – 20.30, $p<0.01$) percentage points higher if they had the same distribution of home values in G1 as the white sample. The PAI is not significant, implying that there is not sufficient evidence to determine whether the white sample experienced differential returns to G1 home values as the Black sample.

In the second mediation analysis (Mediation 2) predicting home value percentile in the second generation, there is a difference in mean percentile of 29.78 (27.33 – 32.57, $p<0.001$). All four potential mediating effect pathways are significant. The PAI via G1 home value percentile (7.10; 0.07 – 14.33, $p<0.05$) is larger than the corresponding PIE (4.90; 0.12 – 9.41, $p<0.05$), indicating that the white sample experienced significantly higher returns to G2 home value from G1 home value than the Black sample. The same is true of the PAI (2.10; -0.12 – 4.43, $p<0.05$) and PIE (1.68; 0.71 – 2.81) via G2 college attainment. Taken together, these four pathways predict 53% of the total difference in mean G2 home value percentile. The CDE (G1 home value reference value = renter, G2 college attainment reference value = no college) is significant and larger than any given mediating effects via G1 home value or G2 college attainment, indicating that the average difference between those without a college degree and born to white renting grandparents was still 14.00 (10.84 – 17.67, $p<0.001$) higher than those without a college degree born to Black renting grandparents.

In the third mediation analysis (Mediation 3) predicting college attainment in the third generation, there is a 26.93 (22.14 – 31.34, $p<0.001$) percentage point difference in attainment rate between the children of those racialized as white in 1968 compared to those racialized as Black. The PIE and PAI via parent college attainment are significant

and predict 25% ($p<0.01$) and 9% ($p<0.001$) of this difference, respectively. The PIE via parent home value is the largest mediating effect, predicting 47% (12.76; 6.85 – 18.17, $p<0.001$) of the total difference.

In the fourth mediation analysis (Mediation 4) predicting home value percentile in the third generation, there is a difference in mean percentile of 14.95 (11.95 – 18.04, $p<0.001$). As with the mediation of home values in the second generation, the CDE is again an important factor, predicting 61% (9.41; 5.20 – 12.91, $p<0.001$) of the total difference (here the CDE corresponds to having grandparents and parents without a college degree and renting). The PIEs via parent home value and grandchild college attainment are both relatively important, predicting a difference of 4.77 (1.80 – 7.52, $p<0.001$) and 4.88 (2.91 – 7.18, $p<0.001$) percentile points, respectively.

Last, I report results of the mediation analysis for the racialized gap in summary net worth of grandchildren including all mediators (Mediation 5). We see that the intergenerational history of grandparent and parent home value and educational attainment ultimately operate through grandchildren's home values in influencing their net wealth. This manifests as a large PIE via grandchildren's home values, where the expected home value for the grandchildren of white grandparents results in 32.38 thousand dollars (24.31 – 41.83, $p<0.001$) higher net wealth than the expected home value for the grandchildren of Black grandparents. Again, the PIE indicates that the grandchildren of those racialized as white attain higher home values than the grandchildren of those racialized as Black, and home value has an average effect on net wealth regardless of racialized status. However, the PAI of this increased home value is

also large and the same relative magnitude as the PIE. For the same home value, the grandchildren of those racialized as white see an average return of 26.84 thousand dollars (12.47 – 40.96) more to net wealth compared to the grandchildren of those racialized as Black. There also remains a smaller but significant influence of parent's education net of downstream mediators (5.73; 1.05 – 11.79, $p<0.01$).

Discussion

In this study, I used a fully interacted sequential mediation framework to deconstruct the intergenerational legacy of wealth accumulation for those racialized as Black compared to white in 1968. Using a broad structural model for the dependent systems distributing home values and educational attainment over three generations, I illustrate how the racialized intergenerational transmission of home values is persistent and continues to be a primary driver of the total Black-white wealth gap in 2015-2017 for families that have lived in the United States since 1968-1970. While applied here to broad racialized systems of intergenerational transmission, this framework can provide the scaffolding for a more specific decomposition of any given connection (Figure 2) without losing sight of the relative importance of each subsystem in the larger ecosystem of historical and contemporary structural racism underlying observed wealth disparities (Bonilla-Silva 1997; Reskin 2012; Zuberi and Bonilla-Silva 2008). By combining fully interacted models with a counterfactual mediation analysis of intergenerational pathways, the present study supports a theoretical framework grounded in the causal dependence between racism and class relations as a fundamental cause of wealth in the United States (Itzigsohn and Brown 2020; Phelan and Link 2015; Robinson 2019; Taylor 2019). The

causal arrows in Figure 2 do not arise naturally or inevitably, as is often implied by individualistic theories of "social determinants" that do not simultaneously consider relational systems of social and economic production (Krieger 2011; Phelan and Link 2015; Portes 1998). These causal arrows are maintained by specific actors who stand to benefit from particular arrangements, often rooted in capitalist interests and white supremacy (Cogburn 2019; Laster Pirtle 2020; Pattillo 2013; Phelan and Link 2015; Ray 2019; Sewell 2016; Sugrue 1996; Taylor 2019; Zaimi 2020).

Results from the sequential mediation analysis reveal how the racialized distribution of home values and the racialized returns to those home values became *embedded* in subsequent racial stratification across generations since 1968. First examining stratification in college attainment among the children of grandparents, the PIE via grandparent home value illustrates this was largely a compositional transition. Through a wide variety of public-private partnerships and arrangements in the racialized political economy of housing characterizing the post-war era (Sugrue 1996; Taylor 2019), Black families were effectively locked out of the high home values attained by white families, which had a large main effect on college attainment in the subsequent generation. In predicting home values in the second generation, the proportion attributable to interaction (PAIs) were larger than the corresponding pure indirect effects (PIEs), indicating that differential *returns* to grandparent home values and college were more important than the compositional differences. Particularly, the proportion attributable to interaction (PAI) via grandparent home value had the largest influence, indicating that even if the distribution of home values in the white sample would have

been the same as in the Black sample, there were large positive returns only experienced by the children of white grandparents. In predicting grandchild college attainment, the majority of the total difference is attributable to the pure indirect effect (PIE) via home value in second generation. This finding is an important empirical illustration of Reskin's (2012) concept of *emergent discrimination*. The racialized distribution of home values in the second generation, which has a large main effect on the education distribution in third generation, was itself built on a system of *racial discrimination* as seen in the previous mediation analyses. The distribution of home values in the latest generation (2015-2017) remains highly racialized and is historically contingent on grandparent and parent characteristics, resulting in a pure indirect effect on the wealth of grandchildren of white grandparents of $32,380. Racialized differences in home values for this generation are the combination of racist systems of market valuation of homes (Howell and Korver-Glenn 2020) and value appreciation (Flippen 2004; Markley et al. 2020). These systems have continued over generations and shape the abilities of grandparents and parents to influence grandchildren's access to high value housing.

While emergent discrimination in total wealth is predicated on this racialized distribution of home values built over generations, *racial discrimination* via the proportion attributable to interaction (PAI) connecting home values to total wealth in the current generation ($26,840) remains a primary component of the wealth disparity across grandchildren ($44,790). There is also a relatively large proportion attributable to interaction (PAI) operating through parents' home value outside of its influence on grandchildren's home value ($18,660), though this is only marginally significant

(p<0.10). These are the result of the racialized production of home equity and its importance in total net wealth, which again involves value appreciation over time but also mortgage terms and assessment of property taxes. Homes in majority-Black neighborhoods continue to be systematically undervalued over time (Howell and Korver-Glenn 2018, 2020; Markley et al. 2020) and regressive tax structures continue to disproportionately burden Black homeowners with low-valued property (Berry 2021). For similarly valued homes, Black homeowners are much more likely to have subprime loans and high-risk mortgage terms which prevent building equity (Markley et al. 2020; Rugh and Massey 2010; Steil et al. 2018).

In summary, my results speak to both the long-run positioning of private homeownership as an integral feature of racialized wealth accumulation via this pure indirect effect (PIE) and the fact that closing the Black-white gap in homeownership is unlikely to be a panacea for closing total wealth gaps without simultaneously addressing the systems of racial discrimination underlying this proportion attributable to interaction (PAI) (Darity et al. 2018; Darity and Mullen 2020; Markley et al. 2020). Structural racism in housing is historically contingent, but actively reproduced in each generation by a political economy of public and private actors (Korver-Glenn 2018b; Sewell 2016; Taylor 2019).

Limitations and future research

Here I focused on important features of the *intergenerational* process of wealth accumulation via the racialized production and distribution of home values and racist

returns to those values, both directly and indirectly via educational attainment. Examining linked generations is theoretically important for quantifying long-run processes of historical and contemporary structural racism, given that each cohort moves through specific systems of racism and racialization that are contingent on time and space. Still, *intragenerational* wealth accumulation is also integral to our understanding of the production and reproduction of wealth gaps between those populations racialized as Black compared to those racialized as white. Killewald & Bryan (2018) illustrate how the Black-white gap in summary net worth is reproduced particularly across ages 25-40, and Thomas et al. (2020) note the importance of age effects in relation to period or cohort explanations for predicting racialized differences in wealth. Future research can help to further examine the specific racialized instruments and pathways of wealth accumulation in early adulthood, including in the housing market, and how this is built in part on long-run intergenerational histories of discrimination.

Conclusions

Vastly expanded during the New Deal and post-war era in the United States, homeownership remains a hugely influential instrument for private wealth accumulation across the life course and a historically important tool for building and protecting the pipelines to well-resourced K-12 schools and college attainment (Geismer 2015; Jackson 1987; Sugrue 1996; Taylor 2019). Still, the ability to effectively *access* and *use* high home values to build wealth remains blocked for the children and grandchildren of those racialized as Black in 1968. As I demonstrate using a fully interacted mediation, this is a multifaceted intergenerational system of racial and emergent discrimination (Reskin

2012) – dynamics which can be obscured by conventional mediation or decomposition frameworks. Closing the gap in home value attainment and its role in wealth accumulation has never been about simply expanding Black homeownership (Darity et al. 2018; Markley et al. 2020), though such expansion has been used historically for the sake of profit and exploitation via the predatory inclusion of Black Americans into substandard housing markets (Taylor 2019). The ability to translate home value to equity remains stratified through entangled systems of racism, including racial discrimination at various points in shopping for homes, initial valuation at the time of purchase, mortgage and lending terms, value appreciation over time, and assessment of property taxes (Howell and Korver-Glenn 2020; Korver-Glenn 2018a, 2018b; Sewell 2016; Taylor 2019).

Housing is an important part of the ecosystem of structural racism in the United States (Pattillo 2013; Reskin 2012; Sewell 2016). But these structures are maintained and reproduced by racist individual actions, cultural norms, and collective decisions made throughout the private market and related bureaucracy of local, state, and federal policymaking (Korver-Glenn 2018b; Ray 2019; Sewell 2016; Taylor 2019). Importantly, these dynamics reify and maintain racial boundaries in the United States (Sewell 2016). There are evidence-based policies that can be implemented now, such as standardizing the property valuation process, increased lending oversight, and policies supporting historically marginalized homeowners and renters that are focused on reparations and restorative justice (Darity and Mullen 2020; Howell and Korver-Glenn 2020). Still, as argued by Pattillo (2013), Taylor (2019), and others, removing housing from the

reproduction of systems of racial domination in the United States will likely require some amount of decommodification and removal from speculative financial markets, as well as establishing the provision of housing as a human right (Baiocchi et al. 2020; Pattillo 2013; Stein 2019; Taylor 2019). A historical, relational, and intergenerational perspective is necessary to avoid further subsidizing a public-private housing system that has never equitably provided adequate housing to those racialized as Black but has consistently and effectively redistributed property value to those racialized as white. As Taylor (2019) writes: "Homeownership is a central cog in the U.S. economy. Its pivotal role as an economic barometer and motor means that there are endless attempts to make it more accessible to ever-wider groups of people. While these are certainly statements of fact, they should not be seen as statements on the advisability of suturing economic well-being to a privately owned asset in a society where the value of that asset will be weighed by the race or ethnicity of whoever possesses it … The promotion of homeownership by the state is not only an acceptance of these market dynamics; it is also an abdication of responsibility for the equitable provision of resources that attend to the racial deficit created by the inequality embedded in homeownership" (Taylor 2019; 258, 261).

Tables and Figures

Table 1. Weighted descriptive statistics for the intergenerational sample.

	Generation 1 (G1)		Generation 2 (G2)		Generation 3 (G3)	
	1968-1973		1995-2000		2015-2017	
	Black	White	Black	White	Black	White
Age (years)	38.57	40.24	38.36	39.44	33.04	32.07
Male (%)	0.32	0.48	0.23	0.44	0.51	0.49
Renter (%)	0.62	0.20	0.50	0.15	0.66	0.41
Owner (% bottom quartile)	0.18	0.09	0.23	0.13	0.15	0.14
Owner (% second quartile)	0.14	0.14	0.14	0.18	0.10	0.21
Owner (% third quartile)	0.06	0.25	0.12	0.24	0.04	0.15
Owner (% top quartile)	0.01	0.33	0.01	0.30	0.04	0.10
High school or less (%)	0.83	0.57	0.52	0.35	0.36	0.22
Some college (%)	0.10	0.30	0.29	0.27	0.40	0.26
College or more (%)	0.07	0.12	0.19	0.38	0.23	0.52
Summary net worth (median)					10,022	41,440
Summary net worth (mean)					40,697	99,765
N					1,430	1,647

Mean summary net worth is calculated after truncating at -$500,000 and $500,000.

Table 2. All effect estimates across five sequential mediation analyses (effects in Mediation 5 – grandchild total wealth – reported in thousands).

	Mediation 1		**Mediation 2**		**Mediation 3**		**Mediation 4**		**Mediation 5**	
	Mean	%	Mean	%	Mean	%	Mean	%	Mean	%
ATE	17.95*** (13.66, 21.79)	100	29.78*** (27.33, 32.57)	100	26.93*** (22.14, 31.34)	100	14.95*** (11.95, 18.04)	100	57.36*** (44.79, 70.32)	100
CDE	2.48 (-2.36, 7)	14	14.00*** (10.84, 17.67)	47	5.88* (0.76, 10.76)	22	9.14*** (5.20, 12.91)	61	-12.5 (-27.59, 2.00)	-22
G1 House value percentile (PAI)	4.51 (-6.04, 14.8)	25	7.10* (0.07, 14.33)	24	-5.62 (-13.59, 1.57)	-21	-1.97 (-7.84, 3.83)	-13	-5.91 (-27.79, 17.88)	-10
G1 House value percentile (PIE)	10.99** (2.97, 20.3)	61	4.90* (0.12, 9.41)	16	4.34 (-0.42, 10.83)	16	1.81 (-2.29, 5.86)	12	-0.47 (-17.01, 14.61)	-1
G2 Edu attainment (PAI)			2.10* (-0.12, 4.43)	7	6.75** (2.5, 10.92)	25	-2.54* (-5.10, 0.01)	-17	-3.42 (-15.17, 8.20)	-6
G2 Edu attainment (PIE)			1.68*** (0.71, 2.81)	6	2.47*** (0.83, 4.56)	9	0.39 (-0.64, 1.50)	3	5.73** (1.05, 11.79)	10
G2 House value percentile (PAI)					1.21 (-6.84, 10.48)	4	2.78 (-2.11, 7.80)	19	18.66 (-6.56, 39.94)	33
G2 House value percentile (PIE)					12.76*** (6.85, 18.17)	47	4.77*** (1.80, 7.52)	32	4.71 (-7.91, 17.68)	8
G3 Edu attainment (PAI)							-4.31* (-8.33, -0.25)	-29	-7.40 (-21.46, 8.40)	-13
G3 Edu attainment (PIE)							4.88*** (2.91, 7.18)	33	-1.25 (-9.32, 6.21)	-2
G3 House value percentile (PAI)									26.84*** (12.47, 40.96)	47
G3 House value percentile (PIE)									32.38*** (24.31, 41.83)	56

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Figure 1. Directed acyclic diagram (DAG) describing the connection between racialized status of parents ($R^{(1)}$) to the wealth of children ($Y^{(2)}$) via observed pathways through multiple dependent mediators (parents' home values, $H^{(1)}$; children's educational attainment, $M^{(2)}$) and other unobserved pathways.

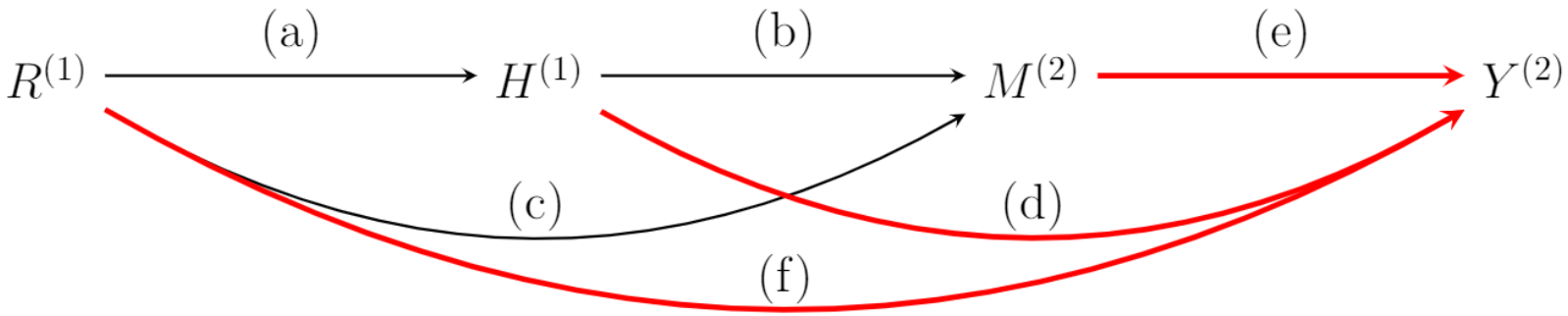


Figure 2. Directed acyclic diagrams (DAGs) describing five sequential mediation analyses connecting the racialized status of grandparents ($R^{(1)}$) and their home value percentile ($H^{(1)}$) to the following generational outcomes: parent educational attainment ($M^{(2)}$), parent home value percentile ($H^{(2)}$), grandchild educational attainment ($M^{(3)}$), grandchild home value percentile ($H^{(3)}$), and grandchild net wealth ($Y^{(3)}$), including generational confounders unaffected by $R^{(1)}$ ($\boldsymbol{V}$; age, age squared, sex). Every intermediate arrow (e.g., $H^{(2)} \rightarrow M^{(3)}$, $H^{(3)} \rightarrow Y^{(3)}$) contains an interaction with $R^{(1)}$. Target outcomes and effect pathways are highlighted in red.

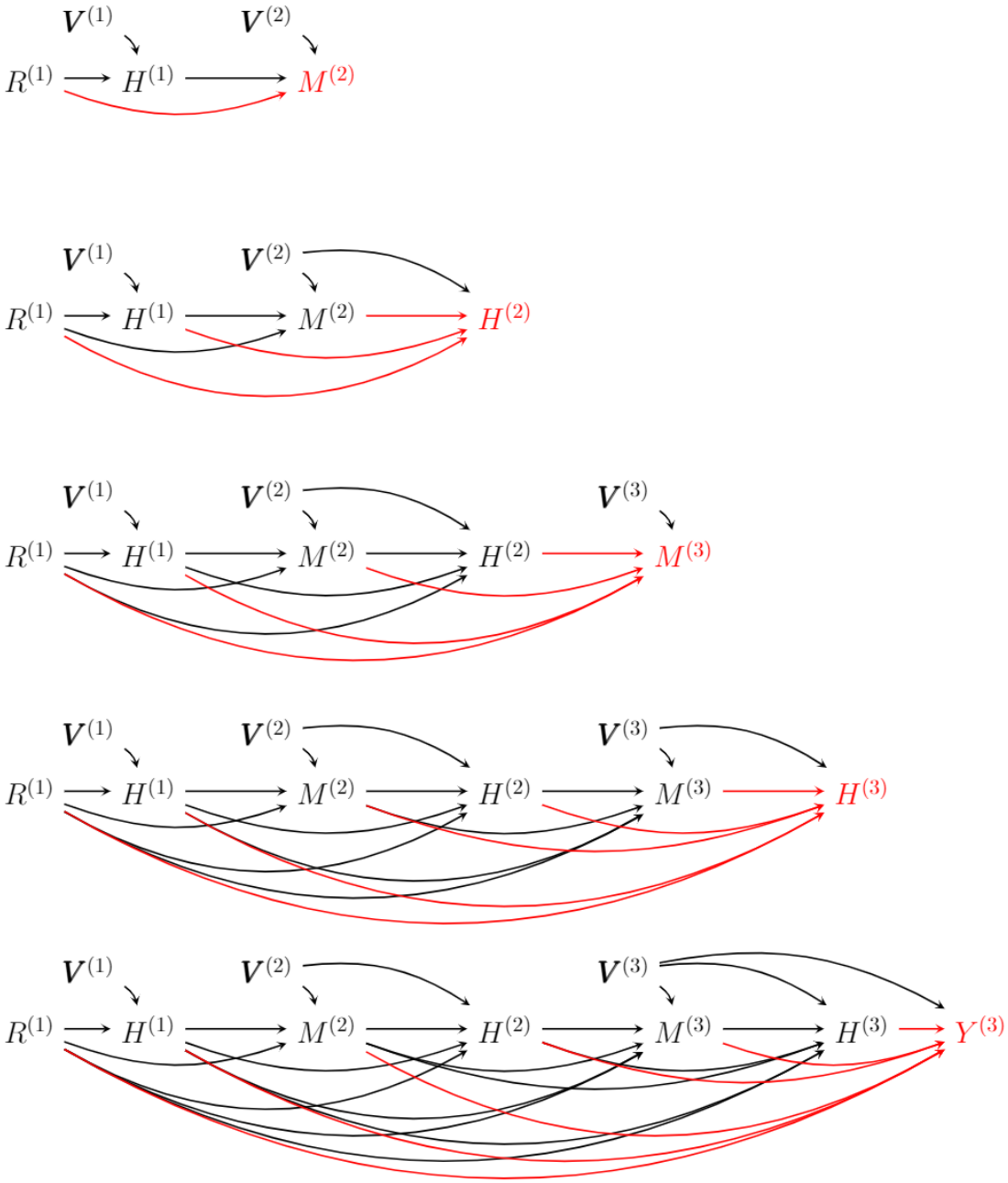


Figure 3. Proportions in each home value quartile across each pair of generations (grandparents to parents: G3-G2, parents to grandchildren: G2-G1) by grandparent (G1) self-reported race. Home values are discretized into 0 = renter, 1 = lowest period-specific quartile … 4 = highest period-specific quartile.

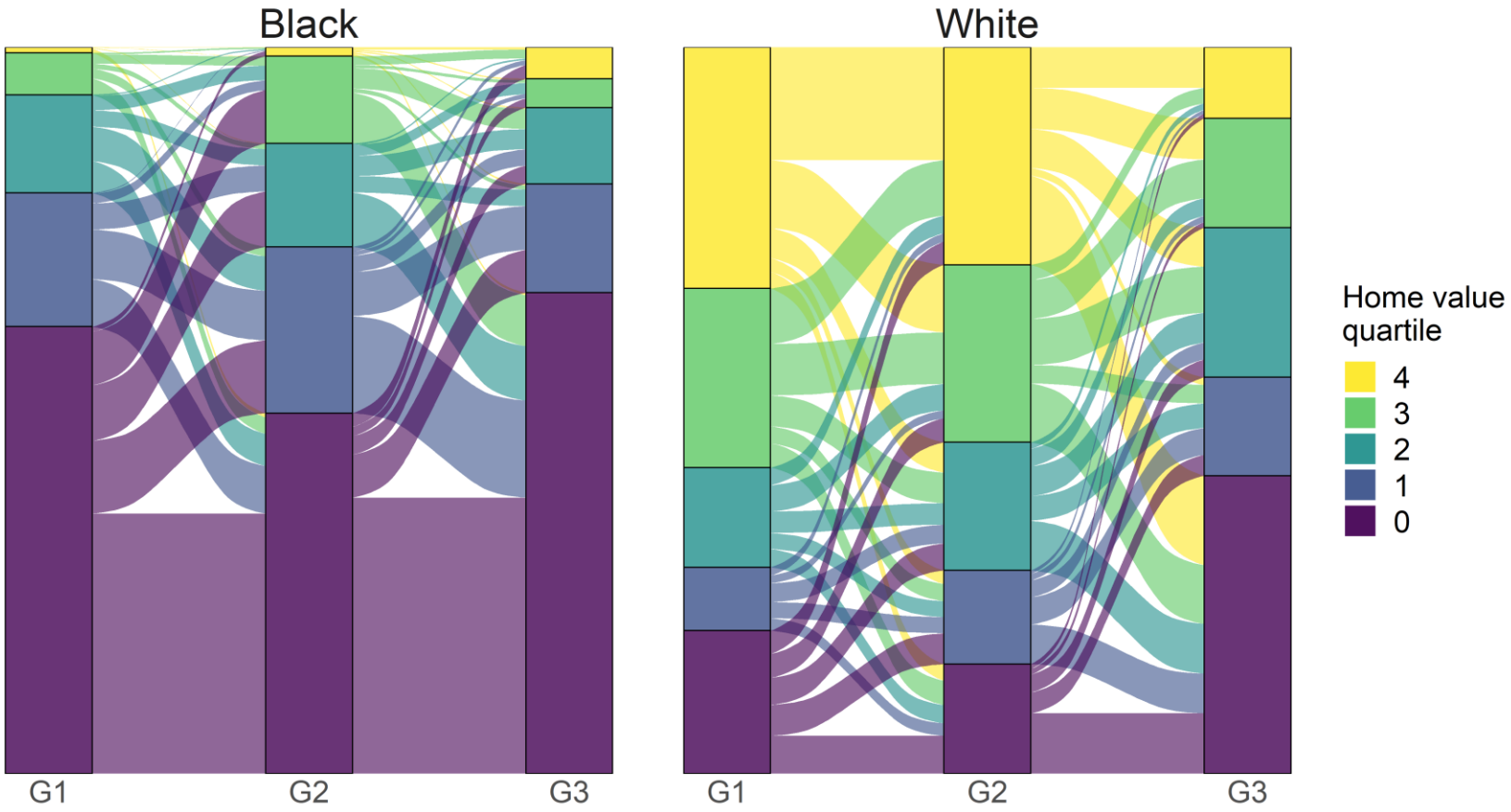


Figure 4. Proportions in each home value quartile across generations by grandparent (G1) self-reported race, following each unique combination (e.g., 0-0-0) over all three generations. Home values are discretized into 0 = renter, 1 = lowest period-specific quartile … 4 = highest period-specific quartile.

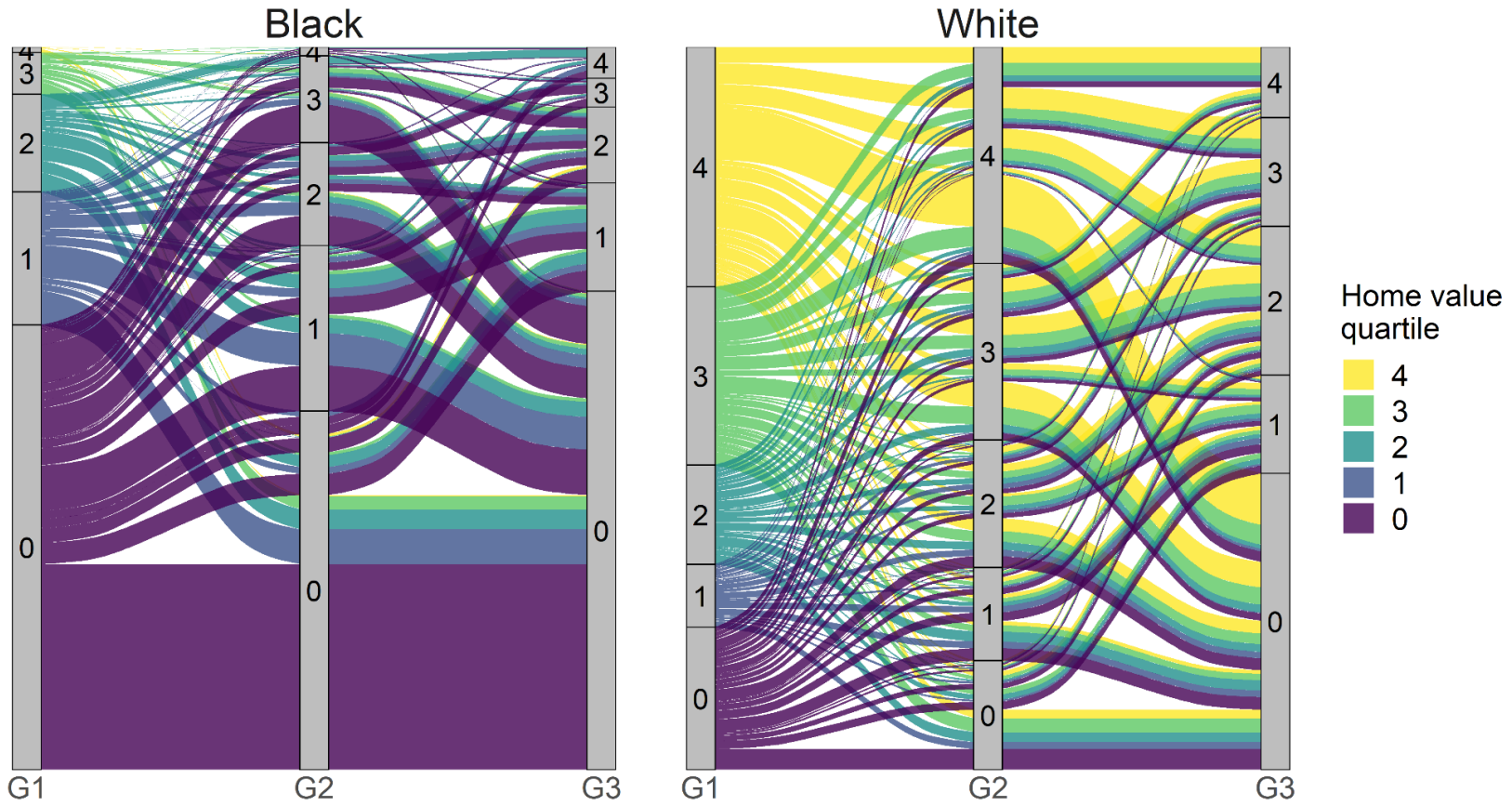
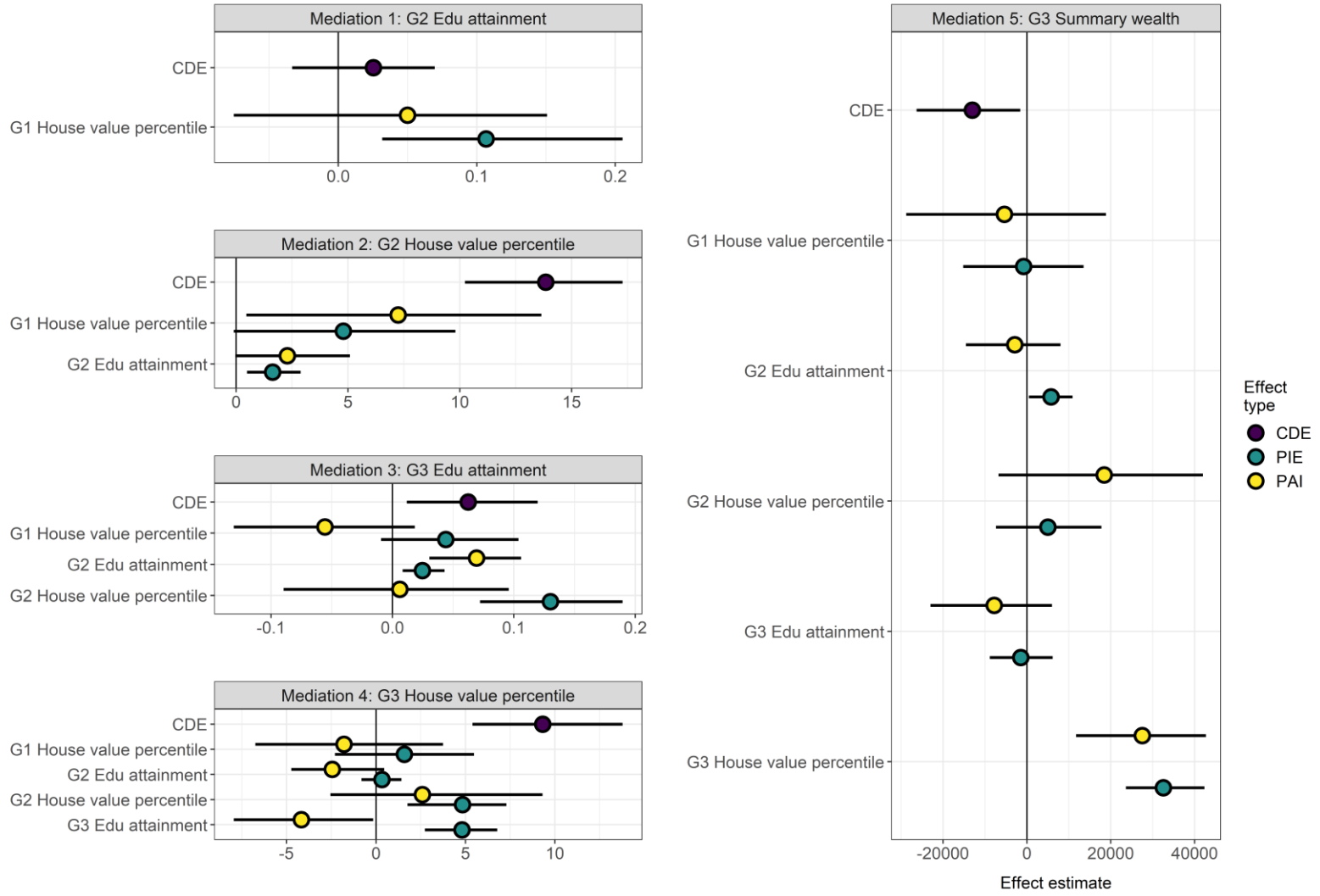


Figure 5. All effect estimates across five sequential mediation analyses.



References

Addo, Fenaba R. 2014. “Debt, Cohabitation, and Marriage in Young Adulthood.” *Demography* 51(5):1677–1701.

Addo, Fenaba R., Jason N. Houle, and Daniel Simon. 2016. “Young, Black, and (Still) in the Red: Parental Wealth, Race, and Student Loan Debt.” *Race and Social Problems* 8(1):64–76.

Alexander, Michelle. 2012. *The New Jim Crow: Mass Incarceration in the Age of Colorblindness*. The New Press.

Bailey, Zinzi D., Nancy Krieger, Madina Agénor, Jasmine Graves, Natalia Linos, and Mary T. Bassett. 2017. “Structural Racism and Health Inequities in the USA: Evidence and Interventions.” *The Lancet* 389(10077):1453–63.

Baiocchi, Gianpaolo, H. Jacob Carlson, Marnie Brady, Ned Crowley, and Sara Duvisac. 2020. “The Case for a Social Housing Development Authority.” (November).

Bauer, Greta R., and Ayden I. Scheim. 2019. “Methods for Analytic Intercategorical Intersectionality in Quantitative Research: Discrimination as a Mediator of Health Inequalities.” *Social Science and Medicine* 226:236–45.

Berry, Christopher. 2021. “Reassessing the Property Tax Christopher.” *SSRN*.

Blair, Clancy, and Cybele Raver. 2016. “Poverty, Stress, and Brain Development:

New Directions for Prevention and Intervention." *Academic Pediatrics* 16(3).

Boen, Courtney. 2016a. "The Role of Socioeconomic Factors in Black-White Health Inequities across the Life Course: Point-in-Time Measures, Long-Term Exposures, and Differential Health Returns." *Social Science & Medicine (1982)* 170:63–76.

Boen, Courtney. 2016b. "The Role of Socioeconomic Factors in Black-White Health Inequities across the Life Course: Point-in-Time Measures, Long-Term Exposures, and Differential Health Returns." *Social Science and Medicine* 170:63–76.

Boen, Courtney. 2020. "Death by a Thousand Cuts: Stress Exposure and Black-White Disparities in Physiological Functioning in Late Life." *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences* 75(9):1937–50.

Boen, Courtney, Lisa Keister, and Brian Aronson. 2020. "Beyond Net Worth: Racial Differences in Wealth Portfolios and Black–White Health Inequality across the Life Course." *Journal of Health and Social Behavior* 61(2):153–69.

Boen, Courtney, and Y. Claire Yang. 2016. "The Physiological Impacts of Wealth Shocks in Late Life: Evidence from the Great Recession." *Social Science and Medicine* 150:221–30.

Du Bois, W. E. B. 1898. “The Study of the Negro Problems.” *Annals of the American Academy of Political and Social Science*.

Du Bois, W. E. B. 1899. *The Philadelphia Negro: A Social Study*. The University of Pennsylvania Press.

Du Bois, W. E. B. 1935. *Black Reconstruction in America, 1860-1880*.

Bonilla-Silva, Eduardo. 1997. “Rethinking Racism: Toward a Structural Interpretation.” *American Sociological Review* 62(3):465–80.

Bonilla-Silva, Eduardo. 2009. *Racism without Racists: Color-Blind Racism and the Persistence of Racial Inequality in America*. Rowman & Littlefield Publishers.

Broadbent, Alex. 2019. “The C-Word, the P-Word, and Realism in Epidemiology.” *Synthese*.

Brooks-Gunn, Jeanne., Greg J. Duncan, and J. Lawrence. Aber. 1997. *Neighborhood Poverty*. Russell Sage Foundation.

Brown, Susan. 2006. “Family Structure Transitions and Adolescent Wellbeing.” *Journal of Marriage and Family* 43(3):447–61.

Brown, Tyson H. 2018. “Racial Stratification, Immigration, and Health Inequality: A Life Course-Intersectional Approach.” *Social Forces* 96(4):1507–40.

Brown, Tyson H., Liana J. Richardson, Taylor W. Hargrove, and Courtney S. Thomas. 2016. “Using Multiple-Hierarchy Stratification and Life Course Approaches to Understand Health Inequalities: The Intersecting Consequences of Race, Gender, SES, and Age.” *Journal of Health and Social Behavior* 57(2):200–222.

Burdick-Will, J., J. Ludwig, SW Raudenbush, Robert Sampson, L. Sanbonmatsu, and Patrick Sharkey. 2011. “Converging Evidence for Neighborhood Effects on Children’s Test Scores: An Experimental, Quasi-Experimental, and Observational Comparison.” Pp. 255–76 in *Whither Opportunity?: Rising Inequality, Schools, and Children’s Life Chances (Copublished with the Spencer Foundation)*.

Buuren, Stef van, and Karin Groothuis-Oudshoorn. 2011. “Mice: Multivariate Imputation by Chained Equations in *R*.” *Journal of Statistical Software* 45(3):1–67.

Carlson, Marcia J., and Mary E. Corcoran. 2001. *Family Structure and Children’s Behavioral and Cognitive Outcomes*. Vol. 63.

Cavanagh, Shannon, and Aletha Huston. 2006. “Family Instability and Children’s Early Problem Behavior.” *Social Forces* 85(1):575–605.

Charles, Camille Zubrinsky. 2003. “The Dynamics of Racial Residential Segregation.” *Annual Review of Sociology* 29(1):167–207.

Cherlin, Andrew, Caitlin Cross-Barnet, Linda M. Burton, and Raymond Garrett-Peters. 2008. “Promises They Can Keep: Low-Income Women’s Attitudes toward Motherhood, Marriage, and Divorce.” *Journal of Marriage and Family* 70(4):919–33.

Chetty, Raj, Nathaniel Hendren, and Lawrence F. Katz. 2016. “The Effects of Exposure to Better Neighborhoods on Children : New Evidence from the Moving to Opportunity Experiment †.” 106(4):855–902.

Chowkwanyun, Merlin, and Adolph L. Reed. 2020. “Racial Health Disparities and Covid-19 — Caution and Context.” *New England Journal of Medicine* 383(3):201–3.

Christopher Auld, M., and Nirmal Sidhu. 2005. “Schooling, Cognitive Ability and Health.” *Health Economics* 14(10):1019–34.

Cogburn, Courtney D. 2019. “Culture, Race, and Health: Implications for Racial Inequities and Population Health.” *The Milbank Quarterly* 97(3):736–61.

Coley, Rebekah, Tama Leventhal, Alicia Lynch, and Melissa Kull. 2013. “Relations between Housing Characteristics and the Well-Being of Low-Income Children and Adolescents.” *Developmental Psychology* 23(1):1–7.

Copp, Jennifer E., Peggy C. Giordano, Wendy D. Manning, and Monica A. Longmore. 2018. “Parental Incarceration and Child Well-Being: Conceptual and Practical Concerns Regarding the Use of Propensity Scores.” *Socius:*

Sociological Research for a Dynamic World 4:237802311877930.

Crenshaw, Kimberle. 1991. “Mapping the Margins: Intersectionality, Identity Politics, and Violence Against Women of Color.” *Stanford Law Review* 43(6):1241–99.

Cummins, Steven, Sarah Curtis, Ana V. Diez-Roux, and Sally Macintyre. 2007. “Understanding and Representing ‘place’ in Health Research: A Relational Approach.” *Social Science and Medicine* 65(9):1825–38.

Daniel, R. M., B. L. De Stavola, S. N. Cousens, and S. Vansteelandt. 2015. “Causal Mediation Analysis with Multiple Mediators.” *Biometrics* 71(1):1–14.

Darity, William A. Jr., and A. Kirsten Mullen. 2020. *From Here to Equality*. The University of North Carolina Press.

Darity, William, Darrick Hamilton, Mark Paul, Alan Aja, Anne Price, Antonio Moore, and Caterina Chiopris. 2018. “What We Get Wrong About Closing the Racial Wealth Gap.” *Samuel DuBois Cook Center on Social Equity* (April):1–67.

Diez Roux, Ana V. 2012. “Conceptual Approaches to the Study of Health Disparities.” *Annual Review of Public Health* 33(1):41–58.

Diprete, Thomas A., and Gregory M. Eirich. 2006. “Cumulative Advantage as a Mechanism for Inequality: A Review of Theoretical and Empirical

Developments.” *Annu. Rev. Sociol* 32:271–97.

Douglas S. Massey, Brandon Wagner, Louis Donnelly, Sara McLanahan, Jeanne Brooks-Gunn, Irwin Garfinkel, Colter Mitchell, and Daniel A. Notterman. 2018. “Neighborhood Disadvantage and Telomere Length: Results from the Fragile Families Study.” *RSF: The Russell Sage Foundation Journal of the Social Sciences* 4(4):28.

Duncan, Greg J., Jeanne Brooks-Gunn, and Pamela Kato Klebanov. 1994. “Economic Deprivation and Early Childhood Development Published by : Blackwell Publishing on Behalf of the Society for Research in Child Development Stable URL : Http://Www.Jstor.Org/Stable/1131385 Economic Deprivation and Early Childhood Development.” *Child Development* 65(2):296–318.

Duncan, Greg J., and Richard Murnane. 2011. *Whither Opportunity? Rising Inequality, Schools, and Children’s Life Chances*.

Edin, Kathryn, and Maria Kefalas. 2011. *Promises I Can Keep: Why Poor Women Put Motherhood before Marriage*. University of California Press.

Elder, Glen H. 1985. “Life Course Dynamics: Trajectories and Transitions 1968-1980.”

Elder, Glen H. 1998. “The Life Course as Developmental Theory.” *Child Development*.

Elder, Glen H., Robert Crosnoe, and Monica Kirkpatrick. 2003. “The Emergence and Development of Life Course Theory.” in *Handbook of the Life Course,* edited by J. Mortimer and M. Shanahan.

Elder Jr, Glen H. 1994. “Time, Human Agency, and Social Change: Perspectives on the Life Course*.” *Social Psychology Quarterly* 57(1):4–15.

Elwert, Felix, and Christopher Winship. 2014. “Endogenous Selection Bias: The Problem of Conditioning on a Collider Variable.” *Annual Review of Sociology* 40:31–53.

Emmons, William R., and Bryan J. Noeth. 2015. “Why Didn’t Higher Education Protect Hispanic and Black Wealth?” *In the Balance* (12):1–3.

Emmons, William R., and Lowell R. Ricketts. 2017. “College_is_not_enough_Higher_e.PDF.” *Economic Research/Federal Reserve Bank OfSt. Louis Review* 99(1):7–39.

Esposito, Michael H. 2019. “Inequality in Process: Income and Heterogeneous Educational Health Gradients Among Blacks and Whites in the USA.” *Race and Social Problems* (0123456789).

Faber, Jacob. 2020. “Contemporary Echoes of Segregationist Policy: Spatial Marking and the Persistence of Inequality.” *Urban Studies* 58(5):1067–86.

Faber, Jacob W. 2020. “We Built This: Consequences of New Deal Era Intervention in America’s Racial Geography.” *American Sociological*

Review 85(5):739–75.

Flippen, Chenoa A. 2004. “Unequal Returns to Housing Investments? A Study of Real Housing Appreciation among Black, White, and Hispanic Households.” *Social Forces* 82:1523–51.

Fomby, Paula, and Andrew Cherlin. 2007. “Family Instability and Child Well-Being.” *American Sociological Review* 72(2):181–204.

Ford, Chandra L., and Collins O. Airhihenbuwa. 2010. “The Public Health Critical Race Methodology: Praxis for Antiracism Research.” *Social Science and Medicine* 71(8):1390–98.

Galea, S., M. Riddle, and G. A. Kaplan. 2010. “Causal Thinking and Complex System Approaches in Epidemiology.” *International Journal of Epidemiology* 39(1):97–106.

Galea, Sandro, and Bruce G. Link. 2013. “Six Paths for the Future of Social Epidemiology.” *American Journal of Epidemiology* 178(6):843–49.

Gangl, Markus. 2010. “Causal Inference in Sociological Research.” *Annual Review of Sociology* 36:21–47.

Geismer, Lily. 2015. *Don’t Blame Us: Suburban Liberals and the Transformation of the Democratic Party*. Princeton University Press.

Gelman, Andrew, and Guido Imbens. 2013. “Why Ask Why? Forward Causal Inference and Reverse Causal Questions.” *Inconnu* 1–7.

Geronimus, Arline T., Margaret Hicken, Danya Keene, and John Bound. 2006. "'Weathering' and Age Patterns of Allostatic Load Scores among Blacks and Whites in the United States." *American Journal of Public Health* 96(5):826–33.

Gielen, Andrea C., Wendy Shields, Eileen McDonald, Shannon Frattaroli, David Bishai, and Xia Ma. 2012. "Home Safety and Low-Income Urban Housing Quality." *Pediatrics* 130(6):1053–59.

Ginther, Donna, Robert Haveman, and Barbara Wolfe. 2000. "Neighborhood Attributes as Determinants of Children's Outcomes: How Robust Are the Relationships?" *The Journal of Human Resources* 35(4):603.

Glymour, Clark, and Madelyn R. Glymour. 2014. "Commentary: Race and Sex Are Causes." *Epidemiology* 25(4):488–90.

Glymour, M. Maria. 2006. "Using Causal Diagrams to Understand Common Problems in Social Epidemiology." *Methods in Social Epidemiology* 387–422.

Goosby, Bridget J., Jacob E. Cheadle, and Colter Mitchell. 2018. "Stress-Related Biosocial Mechanisms of Discrimination and African American Health Inequities." *Annual Review of Sociology* 44(1):319–40.

Green, Tiffany L., and William A. Darity. 2010. "Under the Skin: Using Theories from Biology and the Social Sciences to Explore the Mechanisms behind the

Black-White Health Gap." *American Journal of Public Health* 100(SUPPL. 1):S36.

Hafeman, Danella M., and Sharon Schwartz. 2009. "Opening the Black Box: A Motivation for the Assessment of Mediation." *International Journal of Epidemiology* 38(3):838–45.

Halpin, Gerald, Robert G. Simpson, and Sheila L. Martin. 1990. "An Investigation of Racial Bias in the Peabody Picture Vocabulary Test-Revised." *Educational and Psychological Measurement* 50(1):183–89.

Hamilton, Darrick, and Madeline Neighly. 2019. *The Racial Rules of Corporate Power: How Extractive Corporate Power Harms Black and Brown Communities and How Race-Conscious Solutions Can Create an Inclusive Economy*.

Hamilton, Tod G. 2019. *Immigration and the Remaking of Black America*. Russell Sage Foundation.

Hamilton, Tod G. 2020. "Black Immigrants and the Changing Portrait of Black America." *Annual Review of Sociology* 46:295–313.

Harvey, Michael. 2021. "The Political Economy of Health: Revisiting Its Marxian Origins to Address 21st-Century Health Inequalities." *American Journal of Public Health* 111(2):293–300.

Haskins, Anna R., and Erin J. McCauley. 2018. "Casualties of Context? Risk of

Cognitive, Behavioral and Physical Health Difficulties among Children Living in High-Incarceration Neighborhoods." *Journal of Public Health* 1–9.

Hernán, Miguel. 2018. "The C-Word: The More We Discuss It, the Less Dirty It Sounds." *American Journal of Public Health* 108(5):625–26.

Hernán, Miguel A., and James M. Robins. 2019. *Causal Inference*.

Holland, P. 2008. "Causation and Race." Pp. 93–109 in *White Logic, White Methods*, edited by T. Zuberi and E. Bonilla-Silva. Rowman & Littlefield Publishers.

Houle, Jason N., and Fenaba R. Addo. 2019. "Racial Disparities in Student Debt and the Reproduction of the Fragile Black Middle Class." *Sociology of Race and Ethnicity* 5(4):562–77.

Houle, Jason N., and Cody Warner. 2017. "Into the Red and Back to the Nest? Student Debt, College Completion, and Returning to the Parental Home among Young Adults." *Sociology of Education* 90(1):89–108.

Howe, Chanelle J., and Whitney R. Robinson. 2018. "Survival-Related Selection Bias in Studies of Racial Health Disparities." *Epidemiology* 29(4):521–24.

Howell, Junia, and Elizabeth Korver-Glenn. 2018. "Neighborhoods, Race, and the Twenty-First-Century Housing Appraisal Industry." *Sociology of Race and Ethnicity* 4(4):473–90.

Howell, Junia, and Elizabeth Korver-Glenn. 2020. "The Increasing Effect of

Neighborhood Racial Composition on Housing Values, 1980–2015." *Social Problems* 1–21.

Iceland, John. 2019. *Racial and Ethnic Inequality in Poverty and Affluence, 1959–2015*. Vol. 38. Springer Netherlands.

Ichikawa, Kayoko, Takeo Fujiwara, and Ichiro Kawachi. 2017. "It Takes a Village: Fixed-Effects Analysis of Neighborhood Collective Efficacy and Children's Development." *Journal of Epidemiology* 27(8):368–72.

Imai, Kosuke, Luke Keele, Dustin Tingley, and Teppei Yamamoto. 2011. "Unpacking the Black Box of Causality: Learning about Causal Mechanisms from Experimental and Observational Studies." *American Political Science Review* 105(4):765–89.

Itzigsohn, Jose, and Karida L. Brown. 2020. "The Sociology of W. E. B. Du Bois: Racialized Modernity and the Global Color Line." *NYU Press*. Retrieved March 19, 2021 (https://nyupress.org/9781479856770/the-sociology-of-w-e-b-du-bois/).

Jackson, John W., and Onyebuchi A. Arah. 2019. "Invited Commentary: Making Causal Inference More Social and (Social) Epidemiology More Causal." *American Journal of Epidemiology* 189(3):179–82.

Jackson, John W., and Tyler J. VanderWeele. 2018. "Decomposition Analysis to Identify Intervention Targets for Reducing Disparities." *Epidemiology*

(Cambridge, Mass.) 29(6):825–35.

Jackson, John W., and Tyler J. VanderWeele. 2019. "Intersectional Decomposition Analysis with Differential Exposure, Effects, and Construct." *Social Science and Medicine* 226(January):254–59.

Jackson, Kenneth T. 1987. *Crabgrass Frontier: The Suburbanization of the United States*. Oxford University Press.

Jencks, Christopher, and Susan E. Mayer. 1990. "The Social Consequences of Growing up in a Poor Neighborhood." Pp. 111–86 in *Inner-City Poverty in the United States*, edited by L. E. Lynn and M. G. H. McGreary. Washington, D.C.: National Academies Press (US).

Kahlenberg, Richard D. 2001. *All Together Now : Creating Middle-Class Schools through Public School Choice*. Washington, D.C. : Brookings Institution Press.

Kane, Jennifer B., Kathleen Mullan Harris, S. Philip Morgan, and David K. Guilkey. 2018. "Pathways of Health and Human Capital from Adolescence into Young Adulthood." *Social Forces* 96(3):949–76.

Keil, Alexander P., Jessie K. Edwards, David B. Richardson, Ashley I. Naimi, and Stephen R. Cole. 2014. "The Parametric G-Formula for Time-to-Event Data: Intuition and a Worked Example." *Epidemiology (Cambridge, Mass.)* 25(6):889–97.

Killewald, Alexandra, and Brielle Bryan. 2016. “Does Your Home Make You Wealthy?” *Rsf* 2(6):110–28.

Killewald, Alexandra, and Brielle Bryan. 2018. “Falling Behind: The Role of Inter- and Intragenerational Processes in Widening Racial and Ethnic Wealth Gaps through Early and Middle Adulthood.” *Social Forces* 97(December):705–40.

Killewald, Alexandra, Fabian T. Pfeffer, and Jared N. Schachner. 2017. “Wealth Inequality and Accumulation.” *Annual Review of Sociology* 43(1):379–404.

Kohler-Hausmann, Issa. 2019. “Eddie Murphy and the Dangers of Counterfactual Causal Thinking about Detecting Racial Discrimination.” *Northwestern University Law Review* 113(5):1163–1228.

Korenman, Sanders, Jane E. Miller, and John E. Sjaastad. 1995. “Long-Term Poverty and Child Development in the United States: Results from the NLSY.” *Children and Youth Services Review* 17(1–2):127–55.

Korver-Glenn, Elizabeth. 2018a. “Brokering Ties and Inequality: How White Real Estate Agents Recreate Advantage and Exclusion in Urban Housing Markets.” *Social Currents* 5(4):350–68.

Korver-Glenn, Elizabeth. 2018b. “Compounding Inequalities: How Racial Stereotypes and Discrimination Accumulate across the Stages of Housing Exchange.” *American Sociological Review* 83(4):627–56.

Kramer, Michael R., Eric B. Schneider, Jennifer B. Kane, Claire Margerison-Zilko, Jessica Jones-Smith, Katherine King, Pamela Davis-Kean, and Joseph G. Grzywacz. 2017. "Getting Under the Skin: Children's Health Disparities as Embodiment of Social Class." *Population Research and Policy Review* 36(5):671–97.

Kravitz-Wirtz, Nicole. 2016. "Cumulative Effects of Growing Up in Separate and Unequal Neighborhoods on Racial Disparities in Self-Rated Health in Early Adulthood." *Journal of Health and Social Behavior* 57(4):453–70.

Krieger, Nancy. 2001a. "Theories for Social Epidemiology in the 21st Century: An Ecosocial Perspective." *International Journal of Epidemiology* 30(4):668–77.

Krieger, Nancy. 2001b. "Theories for Social Epidemiology in the 21st Century: An Ecosocial Perspective." *International Journal of Epidemiology* 30(4):668–77.

Krieger, Nancy. 2011. *Epidemiology and the People's Health*. Oxford University Press.

Krieger, Nancy. 2014. "On the Causal Interpretation of Race." *Epidemiology* 25(6):937.

Krieger, Nancy. 2018. "The Science and Epidemiology of Racism and Health: Racial/Ethnic Categories, Biological Expressions of Racism, and the

Embodiment of Inequality — an Ecosocial Perspective." *What's the Use of Race?* (2010).

Krieger, Nancy, and George Davey Smith. 2016a. "Response: FACEing Reality: Productive Tensions between Our Epidemiological Questions, Methods and Mission." *International Journal of Epidemiology* 45(6):1852–65.

Krieger, Nancy, and George Davey Smith. 2016b. "The Tale Wagged by the DAG: Broadening the Scope of Causal Inference and Explanation for Epidemiology." *International Journal of Epidemiology* 45(6):1787–1808.

Krysan, Maria, and Kyle Crowder. 2017. *Cycle of Segregation : Social Processes and Residential Stratification*.

Lareau, Annette. 2011. *Unequal Childhoods: Class, Race and Family Life*.

Laster Pirtle, Whitney N. 2020. "Racial Capitalism: A Fundamental Cause of Novel Coronavirus (COVID-19) Pandemic Inequities in the United States." *Health Education and Behavior* 47(4).

Lee, Dohoon, and Margot Jackson. 2017. "The Simultaneous Effects of Socioeconomic Disadvantage and Child Health on Children's Cognitive Development." *Demography* 54(5):1845–71.

Lee, Dohoon, and Sara McLanahan. 2015. "Family Structure Transitions and Child Development: Instability, Selection, and Population Heterogeneity." *American Sociological Review* 80(4):738–63.

Lin SH, Young J, Logan R, Tchetgen JTT, and Vanderweele TJ. 2017. "Parametric Mediational G-Formula Approach to Mediation Analysis with Time-Varying Exposures, Mediators, and Confounders." *Epidemiology* 28(2):266–74.

Lin, Sheng Hsuan, Jessica Young, Roger Logan, Eric J. Tchetgen Tchetgen, and Tyler J. Vanderweele. 2017. "Parametric Mediational g -Formula Approach to Mediation Analysis with Time-Varying Exposures, Mediators, and Confounders." *Epidemiology* 28(2):266–74.

Link, B. G., and J. Phelan. 1995. "Social Conditions as Fundamental Causes of Disease." *Journal of Health and Social Behavior* Spec No:80–94.

Logan, John R., and Harvey L. Molotch. 1987. *Urban Fortunes: The Political Economy of Place*. 20th anniv. Berkeley, CA : University of California Press,.

Logan, Trevon D., and John M. Parman. 2017. "Segregation and Homeownership in the Early Twentieth Century †." *American Economic Review* (5):410–14.

Ludwig, Jens, Greg J Duncan, Lisa A. Gennetian, Lawrence F. Katz, Ronald C. Kessler, Jeffrey R. Kling, and Lisa Sanbonmatsu. 2013. "Long-Term Neighborhood Effects on Low-Income Families: Evidence from Moving to Opportunity." *American Economic Review* 103(3):226–31.

Ludwig, Jens, Greg J. Duncan, Lisa A. Gennetian, Lawrence F. Katz, Ronald C.

Kessler, Jeffrey R. Kling, and Lisa Sanbonmatsu. 2013. “Neighborhood Effects on the Low-Income Adults.” *Science* 1505(2012):1505–11.

Lumley, Thomas. 2010. “Simple and Stratified Sampling.” Pp. 17–37 in *Complex Surveys*. John Wiley & Sons, Inc.

Lumley, Thomas. 2018. “Survey: Analysis of Complex Survey Samples.”

Lundberg, Ian, Rebecca Johnson, and Brandon Stewart. 2021. “What Is Your Estimand? Defining the Target Quantity Connects Statistical Evidence to Theory.” *American Sociological Review* Forthcomin.

Mackie, John L. 1974. *The Cement of the Universe*. Oxford University Press.

MacMillan, Ross, and Ronda Copher. 2005. “Families in the Life Course: Interdependency of Roles, Role Configurations, and Pathways.” *Journal of Marriage and Family* 67(4):858–79.

Markley, Scott N., Taylor J. Hafley, Coleman A. Allums, Steven R. Holloway, and Hee Cheol Chung. 2020. “The Limits of Homeownership: Racial Capitalism, Black Wealth, and the Appreciation Gap in Atlanta.” *International Journal of Urban and Regional Research* 44(2):310–28.

Massey, Douglas S., and Nancy A. Denton. 1993. *American Apartheid: Segregation and the Making of the Underclass*. Harvard University Press.

Matthay, Ellicott C., Erin Hagan, Laura M. Gottlieb, May Lynn Tan, David Vlahov, Nancy E. Adler, and M. Maria Glymour. 2020. “Alternative Causal

Inference Methods in Population Health Research: Evaluating Tradeoffs and Triangulating Evidence." *SSM - Population Health* 10:100526.

Mayer, Susan E., and Christopher Jencks. 1989. *Growing up in Poor Neighborhoods: How Much Does It Matter?* Vol. 243.

McLanahan, Sara. 2009. "Fragile Families and the Reproduction of Poverty." *The ANNALS of the American Academy of Political and Social Science* 621(1):111–31.

Moffitt, Robert. 2005. "Remarks on the Analysis of Causal Relationships in Population Research." *Demography* 1(1):91.

Morgan, Stephen L., and Christopher Winship. 2014. *Counterfactuals and Causal Inference: Methods and Principles for Social Research*. Cambridge University Press.

Muntaner, Carles. 2013. "Invited Commentary: On the Future of Social Epidemiology--a Case for Scientific Realism." *American Journal of Epidemiology* 178(6):852–57.

Muntaner, Carles, Carme Borrell, Christophe Vanroelen, Haejoo Chung, Joan Benach, Il Ho Kim, and Edwin Ng. 2010. "Employment Relations, Social Class and Health: A Review and Analysis of Conceptual and Measurement Alternatives." *Social Science and Medicine* 71(12):2130–40.

Murnane, Richard J., and Frank Levy. 1996. *Teaching the New Basic Skills :*

Principles for Educating Children to Thrive in a Changing Economy. Free Press.

Naimi, Ashley. 2016. "The Counterfactual Implications of Fundamental Cause Theory." *Current Epidemiology Reports* 3(1):92–97.

Naimi, Ashley, Stephen Cole, and Edward Kennedy. 2016. "An Introduction to G Methods." *International Journal of Epidemiology* 46(2):dyw323.

Naimi, Ashley I., Mireille E. Schnitzer, Erica E. M. Moodie, and Lisa M. Bodnar. 2016a. "Mediation Analysis for Health Disparities Research." *American Journal of Epidemiology* 184(4):315–24.

Naimi, Ashley I., Mireille E. Schnitzer, Erica E. M. Moodie, and Lisa M. Bodnar. 2016b. "Mediation Analysis for Health Disparities Research." *American Journal of Epidemiology* 184(4):315–24.

Oakes, J. Michael, Kate E. Andrade, Ifrah M. Biyoow, and Logan T. Cowan. 2015. "Twenty Years of Neighborhood Effect Research: An Assessment." *Current Epidemiology Reports* 2(1):80–87.

Oliver, Melvin L., and Thomas Shapiro. 2006. *Black Wealth, White Wealth: A New Perspective on Racial Inequality*. New York: Routledge.

Pager, Devah. 2003. "The Mark of a Criminal Record." *American Journal of Sociology* 108(5):937–75.

Pais, Jeremy. 2017. "Intergenerational Neighborhood Attainment and the Legacy

of Racial Residential Segregation: A Causal Mediation Analysis." *Demography* 54(4):1221–50.

Pattillo, Mary. 2013. "Housing: Commodity versus Right." *Annual Review of Sociology* 39:509–31.

Pearl, Judea. 2014. "Interpretation and Identification of Causal Mediation." *Psychological Methods* 19(4):459–81.

Petterson, Stephen M., and Alison Burke Albers. 2001. "Effects of Poverty and Maternal Depression on Early Child Development." *Child Development* 72(6):1794–1813.

Pfeffer, Fabian T., and Alexandra Killewald. 2018. "Generations of Advantage. Multigenerational Correlations in Family Wealth." *Social Forces* 96(4):1411–42.

Pfeffer, Fabian T., and Alexandra Killewald. 2019. "Intergenerational Wealth Mobility and Racial Inequality." *Socius: Sociological Research for a Dynamic World* 5:237802311983179.

Phelan, Jo C., and Bruce G. Link. 2015. "Is Racism a Fundamental Cause of Inequalities in Health?" *Annual Review of Sociology* 41(1):311–30.

Pichette, François, Sébastien Béland, and Justyna Leśniewska. 2019. "Detection of Gender-Biased Items in the Peabody Picture Vocabulary Test." *Languages* 4(2):27.

Portes, Alejandro. 1998. “Social Capital: Its Origins and Applications in Modern Sociology.” *Annual Review of Sociology* 24:1–24.

Prins, Seth J., Sarah McKetta, Jonathan Platt, Carles Muntaner, Katherine M. Keyes, and Lisa M. Bates. 2021. “The Serpent of Their Agonies.” *Epidemiology* 32(2):303–9.

Prins, Seth J., and Sharon Schwartz. 2020. “Toward a Dialectical Social Epidemiology Seth J. Prins and Sharon Schwartz Department of Epidemiology, Columbia University.” *In Progress*.

Ray, Victor. 2019. “A Theory of Racialized Organizations.” *American Sociological Review* 84(1):26–53.

Reichman, Nancy E., Julien O. Teitler, Irwin Garfinkel, and Sara S. Mclanahan. 2001. *Fragile Families: Sample and Design*. Vol. 23.

Reskin, Barbara. 2012. “The Race Discrimination System.” *Annual Review of Sociology* 38(1):17–35.

Roberts, Dorothy E. 2011. *Fatal Invention: How Science, Politics, and Big Business Re-Create Race in the Twenty-First Century*. New Press.

Robins, J. M., M. A. Hernán, and B. Brumback. 2000. “Marginal Structural Models and Causal Inference in Epidemiology.” *Epidemiology (Cambridge, Mass.)* 11(5):550–60.

Robins, James. 1986. “A New Approach to Causal Inference in Mortality Studies

with a Sustained Exposure Period-Application to Control of the Healthy Worker Survivor Effect." *Mathematical Modelling* 7(9–12):1393–1512.

Robins, James M., and Sander Greenland. 1992. "Identifiability and Exchangeability for Direct and Indirect Effects." *Epidemiology* 3:143–55.

Robins, James M., and Miguel A. Hernan. 2009. "Estimation of the Causal Effects of Time-Varying Exposures." Pp. 553–99 in *Advances in Longitudinal Data Analysis*. Chapman & Hall.

Robinson, Cedric J. 2019. *Cedric J. Robinson: On Racial Capitalism, Black Internationalism, and Cultures of Resistance*.

Robinson, Whitney R., and Zinzi D. Bailey. 2019. "Invited Commentary: What Social Epidemiology Brings to the Table-Reconciling Social Epidemiology and Causal Inference." *American Journal of Epidemiology* 189(3):171–74.

Robinson, Whitney R., and Zinzi D. Bailey. 2020. "Invited Commentary: What Social Epidemiology Brings to the Table - Reconciling Social Epidemiology and Causal Inference." *American Journal of Epidemiology* 189(3):171–74.

Robitzsch, A., S. Grund, and T. Henke. 2019. "Miceadds: Some Additional Multiple Imputation Functions, Especially for Mice."

Rothstein, Richard. 2017. *The Color of Law: A Forgotten History of How Our Government Segregated America*. Liverright.

Rugh, Jacob S., and Douglas S. Massey. 2010. "Racial Segregation and the

American Foreclosure Crisis." *American Sociological Review* 75(5):629–51.

Sampson, Robert J. 2008a. *Moving to Inequality: Neighborhood Effects and Experiments Meet Social Structure 1*. Vol. 114.

Sampson, Robert J. 2008b. "Moving to Inequality: Neighborhood Effects and Experiments Meet Social Structure Moving to Inequality: Neighborhood Effects and Experiments Meet Social Structure 1." *Source: American Journal of Sociology* 114(1):189–231.

Sampson, Robert J. 2012. *Great American City: Chicago and the Enduring Neighborhood Effect*. The University of Chicago Press.

Sampson, Robert J., Jeffrey D. Morenoff, and Thomas Gannon-Rowley. 2002. "Assessing 'Neighborhood Effects': Social Processes and New Directions in Research." *Annual Review of Sociology* 28(1):443–78.

Sampson, Robert J., Stephen W. Raudenbush, and Felton Earls. 1997. "Neighborhoods and Violent Crime: A Multilevel Study of Collective Efficacy." *Science* 277(5328):918–24.

Schwartz, Sharon, Nicolle M. Gatto, and Ulka B. Campbell. 2011. "What Would Have Been Is Not What Would Be: Counterfactuals of the Past and Potential Outcomes of the Future." *Causality and Psychopathology: Finding the Determinants of Disorders and Their Cures* 25–46.

Schwartz, Sharon, Nicolle M. Gatto, and Ulka B. Campbell. 2016. "Causal

Identification: A Charge of Epidemiology in Danger of Marginalization." *Annals of Epidemiology* 26(10):669–73.

Schwartz, Sharon, Nicolle M. Gatto, and Ulka B. Campbell. 2017. "Heeding the Call for Less Casual Causal Inferences: The Utility of Realized (Quantitative) Causal Effects." *Annals of Epidemiology* 27(6):402–5.

Schwartz, Sharon, Seth Prins, Ulka Campbell, and Nicolle Gatto. 2016. "Is the 'Well-Defined Intervention Assumption' Politically Conservative?" *Social Science & Medicine* 166:254–57.

Seamster, Louise, and Raphaël Charron-Chénier. 2017. "Predatory Inclusion and Education Debt: Rethinking the Racial Wealth Gap." *Social Currents* 4(3):199–207.

Sen, Maya, and Omar Wasow. 2016. "Race as a Bundle of Sticks: Designs That Estimate Effects of Seemingly Immutable Characteristics." *Annual Review of Political Science* 19:499–522.

Sewell, Abigail A. 2016. "The Racism-Race Reification Process." *Sociology of Race and Ethnicity* 2(4):402–32.

Sharkey, Patrick. 2013. *Stuck in Place: Urban Neighborhoods and the End of Progress toward Racial Equality*. The University of Chicago Press.

Sharkey, Patrick, and Felix Elwert. 2011. "The Legacy of Disadvantage: Multigenerational Neighborhood Effects on Cognitive Ability." *American*

Journal of Sociology 116(6):1934–81.

Sharkey, Patrick, and Jacob Faber. 2014. "Where, When, Why, and For Whom Do Residential Contexts Matter? Moving Away from the Dichotomous Understanding of Neighborhood Effects." *Annual Review of Sociology* 40(1):559–79.

Singh-Manoux, Archana, Jane E. Ferrie, John W. Lynch, and Michael Marmot. 2005. "The Role of Cognitive Ability (Intelligence) in Explaining the Association between Socioeconomic Position and Health: Evidence from the Whitehall II Prospective Cohort Study." *American Journal of Epidemiology* 161(9):831–39.

Slater, Tom. 2013. "Your Life Chances Affect Where You Live: A Critique of the 'Cottage Industry' of Neighbourhood Effects Research." *International Journal of Urban and Regional Research* 37(2):367–87.

Small, Mario Luis, and Katherine Newman. 2001. "Urban Poverty after *The Truly Disadvantaged* : The Rediscovery of the Family, the Neighborhood, and Culture." *Annual Review of Sociology* 27(1):23–45.

Smith, Emilie, D. Osgood, Linda Caldwell, Kathryn Hynes, and Daniel Perkins. 2011. "Measuring Collective Efficacy Among Children in Communitybased Afterschool Programs: Exploring Pathways toward Prevention and Positive Youth Development." *Bone* 23(1):1–7.

Steil, Justin P., Len Albright, Jacob S. Rugh, and Douglas S. Massey. 2018. "The Social Structure of Mortgage Discrimination." *Housing Studies* 33(5):759–76.

Stein, Samuel. 2019. *Capital City: Gentrification and the Real Estate State*. Jacobin.

Stewart, Quincy Thomas. 2008. "Swimming Upstream: Theory and Methodology in Race Research." in *White Logic, White Methods*.

Sudharsanan, Nikkil, and Maarten J. Bijlsma. 2019. "A Generalized Counterfactual Approach to Decomposing Differences Between Populations." *MPIDR Working Papers* 49(0):0–49.

Suglia, Shakira Franco, Shakira Franco Suglia, Cristiane S. Duarte, Megan T. Sandel, and Rosalind J. Wright. 2010. "Social and Environmental Stressors in the Home and Childhood Asthma." *Journal of Epidemiology and Community Health* 64(7):636–42.

Sugrue, Thomas J. 1996. *The Origins of the Urban Crisis*. Princeton University Press.

T., Leventhal, and Brooks-Gunn J. 2003. "Moving to Oppurtunity: An Experimental Study of Neighborhood Effects on Mental Health." *American Journal of Public Health* 93(9):1576–82.

Taylor, Keeanga-Yamahtta. 2016. *From #BlackLivesMatter to Black Liberation*.

Haymarket Books.

Taylor, Keeanga-Yamahtta. 2019. *Race for Profit: Black Homeownership and the End of the Urban Crisis*. University of North Carolina Press.

Tesfai, Rebbeca. 2016. "The Interaction between Race and Nativity on the Housing Market: Homeownership and House Value of Black Immigrants in the United States." *International Migration Review* 50(4):1005–45.

Thomas, Melvin, Cedric Herring, Hayward Derrick Horton, Moshe Semyonov, Loren Henderson, and Patrick L. Mason. 2020. "Race and the Accumulation of Wealth: Racial Differences in Net Worth over the Life Course, 1989-2009." *Social Problems* 67(1):20–39.

Vanderweele, Tyler J. 2014. "A Unification of Mediation and Interaction: A 4-Way Decomposition." *Epidemiology* 25(5):749–61.

VanderWeele, Tyler J. 2009. "Marginal Structural Models for the Estimation of Direct and Indirect Effects." *Epidemiology* 20(1):18–26.

VanderWeele, Tyler J. 2016. "Mediation Analysis: A Practitioner's Guide." *Annual Review of Public Health* 37(1):17–32.

VanderWeele, Tyler J., and Whitney R. Robinson. 2014. "On the Causal Interpretation of Race in Regressions Adjusting for Confounding and Mediating Variables." *Epidemiology* 25(4):473–84.

VanderWeele, Tyler J., and Eric J. Tchetgen Tchetgen. 2017. "Mediation

Analysis with Time Varying Exposures and Mediators." *Journal of the Royal Statistical Society. Series B: Statistical Methodology* 79(3):917–38.

Vanderweele, Tyler J., and Stijn Vansteelandt. 2009. "Conceptual Issues Concerning Mediation, Interventions and Composition." *Statistics and Its Interface* 2(4):457–68.

Vanderweele, Tyler J., Stijn Vansteelandt, and James M. Robins. 2014. "Effect Decomposition in the Presence of an Exposure-Induced Mediator-Outcome Confounder." *Epidemiology* 25(2):300–306.

Vangen-Lønne, Anne M., Peter Ueda, Pablo Gulayin, Tom Wilsgaard, Ellisiv B. Mathiesen, and Goodarz Danaei. 2018. "Hypothetical Interventions to Prevent Stroke: An Application of the Parametric g-Formula to a Healthy Middle-Aged Population." *European Journal of Epidemiology* 3456789:1–10.

Wacquant, Loïc J. D. 2008. *Urban Outcasts: A Comparative Sociology of Advanced Marginality*. Cambridge ; Polity,.

Wang, Aolin, and Onyebuchi A. Arah. 2015. "G-Computation Demonstration in Causal Mediation Analysis." *European Journal of Epidemiology* 30(10):1119–27.

Watson-Daniels, Jamelle, Yeshimabeit Milner, Nicole Triplett, Irene Headen, Dominique Day, Zinzi Bailey, Meme Styles, Lisa Clinton, Courtni Andrews,

Michelle Wilson, Nchedochukwu Ezeokoli, Stacy Jebbett Bullard, and Lucas Mason-Brown. 2020. *Data for Black Lives COVID-19 Movement Pulsecheck and Roundtable Report*.

Wen, Ming, Christopher R. Browning, and Kathleen A. Cagney. 2003. *Poverty, Affluence, and Income Inequality: Neighborhood Economic Structure and Its Implications for Health*. Vol. 57.

Williams, D. R., and C. Collins. 2001. "Racial Residential Segregation: A Fundamental Cause of Racial Disparities in Health." *Public Health Reports (Washington, D.C. : 1974)* 116(5):404–16.

Williams, Daedric. 2019. "A Call to Focus on Racial Domination and Oppression: A Response to "Racial and Ethnic Inequality in Poverty and Affluence, 1959–2015″." *Population Research and Policy Review* (July).

Williams, David R. 2012. "Miles to Go before We Sleep: Racial Inequities in Health." *Journal of Health and Social Behavior* 53(3):279–95.

Williams, David R., Naomi Priest, and Norman B. Anderson. 2016. "Understanding Associations among Race, Socioeconomic Status, and Health: Patterns and Prospects." *Health Psychology* 35(4):407–11.

Willms, Douglas. 2010. "School Composition and Contextual Effects on Student Outcomes." *Teachers College Record.* 112(4).

Wilson, William J. 1987. *The Truly Disadvantaged : The Inner City, the*

Underclass, and Public Policy. University of Chicago Press.

Winship, Christopher, and Stephen L. Morgan. 1999. "The Estimation of Causal Effects from Observational Data." *Annual Review of Sociology* 25:659–707.

Winship, Christopher, and Michael Sobel. 2004. "Causal Inference in Sociological Studies." in *Handbook of Data Analysis*. SAGE Publications.

Wodtke, Geoffrey T., David J. Harding, and Felix Elwert. 2011. "Neighborhood Effects in Temporal Perspective: The Impact of Long-Term Exposure to Concentrated Disadvantage on High School Graduation." *American Sociological Review* 76(5):713–36.

Wodtke, Geoffrey T., and Matthew Parbst. 2017. "Neighborhoods, Schools, and Academic Achievement: A Formal Mediation Analysis of Contextual Effects on Reading and Mathematics Abilities." *Demography* 54(5):1653–76.

Wodtke, Geoffrey T., and Xiang Zhou. 2020. "Effect Decomposition in the Presence of Treatment-Induced Confounding: A Regression-with-Residuals Approach." *Epidemiology* (435):369–75.

Zaimi, Rea. 2020. "Making Real Estate Markets: The Co-Production of Race and Property Value in Early 20th Century Appraisal Science." *Antipode* 52(5):1539–59.

Zewde, Naomi. 2020. "Universal Baby Bonds Reduce Black-White Wealth Inequality , Progressively Raise Net Worth of All Young Adults."

Zuberi, Tukufu. 2001. *Thicker than Blood: How Racial Statistics Lie*. University of Minnesota Press.

Zuberi, Tukufu., and Eduardo. Bonilla-Silva. 2008. *White Logic, White Methods: Racism and Methodology*. Rowman & Littlefield Publishers.

Zuberi, Tukufu, Evelyn J. Patterson, and Quincy Thomas Stewart. 2015. "Race, Methodology, and Social Construction in the Genomic Era." *Annals of the American Academy of Political and Social Science* 661(1):109–27.