Networked resources and health in South Africa: Towards a Materialist Social Network Epidemiology

HARVARD UNIVERSITY

Graduate School of Arts and Sciences

DISSERTATION ACCEPTANCE CERTIFICATE

The undersigned, appointed by the
Committee on Higher Degrees in Population Health Sciences,
have examined a dissertation entitled

**"Networked Resources and Health in South Africa:
Towards a Materialist Social Network Epidemiology"**

presented by

KELETSO MAKOFANE

candidate for the degree of Doctor of Philosophy
and hereby certify that it is worthy of acceptance.

Lisa Berkman

Dr. Lisa F. Berkman, Ph.D., Committee Chair, Harvard T.H. Chan School of Public Health

Mary J. Bassett
Mary Bassett (Sep 3, 2021 10:40 EDT)

Dr. Mary Travis Bassett, M.D., Harvard T.H. Chan School of Public Health

Eric Tchetgen Tchetgen (Sep 8, 2021 17:38 EDT)

Dr. Eric Tchetgen Tchetgen, Ph.D., Harvard T.H. Chan School of Public Health

Date: 05 August 2021

Networked Resources and Health in South Africa:

Towards a Materialist Social Network Epidemiology

Keletso Makofane

August 2021

Dissertation Advisor: Prof. Lisa Berkman Keletso Makofane

Networked Resources and Health in South Africa:

Towards a Materialist Social Network Epidemiology

Abstract

People share resources with each other across meaningful social relationships. It has been demonstrated that increased resources cause better health for individuals, yet there is no epidemiologic research investigating the impact on population health of material resources embedded in social networks. This knowledge gap is related to the use in epidemiologic studies of theories which do not account explicitly for material resources, the assumption of independence in causal inference, and the dearth of socio-centric network data linked to health outcomes and exposures. Drawing on newly developed methods in causal inference, the project examines whether: a) a financial incentive for HIV testing has effects across family relationships; b) the wealth held by one household affects the physical function of family members living in another; and c) the wealth held by one household affects mortality in the households of family members.

To answer these questions, routinely collected data from two South African health and demographic surveillance systems were used to construct socio-centric family networks. Overlaid on these networks were data from HITS — a cluster randomized controlled trial studying the effect of a financial incentive on HIV testing, and data from HAALSI — a cohort study of older South Africans investigating physical function among other health outcomes.

Related to a), financial incentives for HIV testing are shown to not only influence the person who is offered the incentive, but their family members as well. The offer of a financial

incentive affects only the behavior of individuals who were themselves offered the incentive, however. With respect to b), wealth held by family members do not appear to affect the physical function of older individuals. Regarding c), the wealth held by a household is shown to lighten the mortality of its members as well as that of non-household family members. This effect was pronounced among working age (16-59) adults.

Overall, this dissertation finds evidence that the resources held in family networks shape health across the networks. It demonstrates that it is feasible to conduct similar analyses using available theory and data and newly developed analytic methods.

Table of Contents

Acknowledgements

In the rush to finalize my defense presentation, I failed to thank some valued members of my chosen family — Renugan Raidoo, Pedro Carneiro and Ken Nadolski — for seeing me through the best and worst of the past decade of my life. Alex Warburton and Renugan have been part of my Boston family, granting me access to a dependable supply of plastic-cup wine and cubed cheese in the Tozzer Anthropology Building.

My friend and husband, Nick Diamond, took the lion's share of the care work of our household during my studies. He also helped to edit this work, along the way trying to teach me (with limited success) how to use commas. My parents-in-law, Michele and John, have helped us to solve all kinds of practical problems and warmly welcomed me into their family. My sister, Reabetswe, has taken care of our parents and helped me to meet my financial obligations back home. My parents, Mohlolo and Mokitlane, have led by example, learning unceasingly. Over the short time I have been studying, my mother has become a self-taught milliner. My diploma will join the certificates, degrees, and graduation photographs on display in the nook of our family home, which has come to be a shrine to education. My parents recruited a contingent of their friends to my defense: Mr. Sekgopi Malebo, Dr. Herbert Ngakane, Mrs. Malikhang Masia, Mr. Eric Phayane, Mrs. Tumi Phayane, Dr. Mpho Jama, Ms. Winnie Ngakane, Dr. Refilwe Seane, and Dr. Livo Seane, and others who were a part of my life before I was.

The closest that humans can come to a cloud computing service is the Library Steps group consisting of Brigette Davis, Onisha Etkins, Ashley Gripper, and Jourdyn Lawrence. This group of friends has been an invaluable intellectual and emotional resource, helping me to notice the strange ideas many members of the elite academy have of Black people and, particularly, women. Brigette Davis brought us together before the program began. She is also one of the

three other people in the world who experienced the Library Steps incident. For the past year, Tori Cowger and Emily Unger have continuously engaged my work, making many insightful suggestions which strengthened it. All these friends helped me to prepare my defense presentation.

This work would not be possible without collaborations with researchers in the U.S., South Africa, Germany, and the U.K. The idea of using surveillance data to construct networks came to me in the middle of a small-group discussion during a workshop for HAALSI collaborators in Cambridge. By the end of the workshop, Dr. Kathy Kahn, Dr. Steve Tollman, and Dr. Cho Kabadula of the Agincourt Unit at Wits University had graciously agreed to support this effort. Dr. Kabadula shared the data I used in Chapters 2 and 3 and helped me to make sense of it. Near the end of another HAALSI meeting, this time in Johannesburg, Dr. Till Baernighausen initiated a collaboration, which opened a new avenue of research involving networks in the context of randomized controlled trials. He and Dr. Frank Tanser welcomed me into the HITS Study Working Group, with whom I am grateful to have presented a version of Chapter 1 at the 2020 International AIDS Conference. Dr. Hae-Young Kim has been a wonderful thought partner in puzzling through epidemiologic methods related to HITS data. My thanks, finally, to Dr. Julia Rohr and Dr. Livia Montana, who each supported my work using HAALSI data.

I found home in the FXB Center for Health and Human Rights and the Center for Population and Development Studies, where I have been a graduate affiliate. I was fortunate to have a desk at the FXB Center and friends to help me to prepare my thesis proposal presentation and to think together about small and large methodological problems. I am grateful for travel support from the Pop Center, even if the pandemic upended all plans. Ms. Anjli Patel, Ms. Rebecca Shin and Mr. Sam Peisch of the FXB Center and Ms. Claudette Agustin of the Pop Center are responsible for the successful scheduling of my progress meetings and dissertation defense,

coordinating across the busy schedules of my committee members. My thanks also to Ms. Elizabeth Solomon, Ms. Whitney Waddel, Mr. Matthew Boccuzi and Mr. Bruce Villineau, who do the daily work of keeping the Social and Behavioral Sciences Department and the Population Health Sciences program running.

Applying to graduate school is a team sport. Over the years, I enlisted the help of Dr. Francesca Little, Dr. David Gay, Mr. Dave Strugnell, Ms. Heather McLeod, Dr. Shivani Ranchod, Dr. Iain Macdonald, and Ms. Billy Enderstein of University of Cape Town; Dr. Theo Sandfort of Columbia University; Dr. George Ayala of MPACT Global Action (while it was called MSMGF), and Dr. Stef Baral of Johns Hopkins University to write letters of reference. Dr. Ayala, Dr. Sandfort, and Dr. Baral have continued to be valued mentors and collaborators. Dr. Sonya Arreola taught me how to write PhD application essays and how to collaborate in community-based participatory action research.

All my teachers, stretching from junior school through to PhD coursework have left an indelible mark. I include tutors and teaching assistants among them. I am especially grateful to Dr. Jarvis Chen, Dr. Nancy Krieger, and Dr. Karen Emmons who helped me to clarify my interests and have been models for good citizenship in the academy. Dr. Mary Brinton provided thoughtful comments and suggestions on an earlier version of the introduction. From my high school, St. Andrew's School in Bloemfontein, Mrs. Kerry Gower's English classes and debating club; Mrs. Petro Devilliers's after-school "First Team Accounting" practice; Mr. Ettienne Fourie's mathematics classes; and Mrs. Tilla Henkins's cello lessons all taught me to cultivate and take joy in my talents.

Finally, I owe a profound debt of gratitude to the members of my PhD committee — Dr. Mary Bassett, Dr. Eric Tchetgen Tchetgen, and Dr. Lisa Berkman — who have supported this project

with the utmost kindness and attention. I left every meeting with them energized and excited about doing the work. Dr Bassett helped me to achieve clarity without sacrificing precision and challenged me to think about my project's implications for practice. Dr. Tchetgen Tchetgen patiently helped me to learn unfamiliar statistical methods and to identify and overcome methodological roadblocks. As my academic advisor, Dr. Berkman helped to whittle down the expansive ideas I came to the program with into a feasible project. She also created many valuable learning opportunities, including those that yielded this project. Each of my mentors inspired me to become a useful mentor to others.

As many as there are people I have named here, I have the feeling that there some that I have omitted. My apologies. I am grateful to each person who ever said a kind word about the audacious task documented in these pages.

From Mohlolo, Neo, Mmaneo, Maletsatsi, Moditsane, and Aria

Motho ke motho ka batho

Figures and Tables

Introduction

Networked resources and the people’s health

In a recent national household survey in South Africa, 10% of South Africans reported having experienced hunger in the preceding year because they had run out of food. Whereas a small minority of South Africans — only 1.6% — reported having poor health, over 90% reported good, very good, or excellent health (Statistics South Africa, 2021a). Yet, of about 60 million South Africans, only 15 million people were employed in March 2020 (Statistics South Africa, 2021c; United Nations, Department of Economic and Social Affairs, Population Division, 2019). Among the unemployed and looking for work, more than 70% had been without a job for more than a year (Statistics South Africa, 2019). How is it possible, then, for the 45 million jobless South Africans to obtain basic necessities — food, shelter, clothing, education — well enough to maintain good self-reported health on average?

The aim of this project is to examine how the material resources that flow through people's social networks shape their health. I focus on connections among relatives; the connections that link the 15 million South Africans who are employed to the 45 million who are not. In the epidemiologic social network and social capital literatures, the project advances our understanding of relationships among family social networks, socio-economic status, and population health outcomes. In applied causal inference, it furthers our understanding of interference — the idea that one person's treatment or exposure can affect another person's health outcome — a phenomenon usually treated as an impediment to valid inference. A more complete understanding of the effect of networked material resources on population health potentially provides insights that can inform welfare policy.

In this chapter, I review theory on social networks and health and social capital and health, highlighting useful concepts as well as gaps in these literatures. I further outline the theoretical perspective that undergirds the project, offer some relevant context on South Africa, describe

the central methodological challenge of the project, and close with a brief overview of the remaining chapters.

Psychosocial roots of social network epidemiology

The formal study of social networks is often traced back to the work of psychiatrist Jacob Moreno who in 1932 studied the relationships among residents of the New York State Training School for Girls. The facility was established in the 1880s to confine and punish unwed mothers. By the time Moreno conducted his investigations, it had been converted into a school for "younger girls loosely defined as 'wayward,' 'incorrigible' or 'in need of supervision'" (Bernstein, 1996). There had been a spate of escapes: over the space of 2 weeks, 14 girls had run away — "a rate 30 times higher than the norm" (Borgatti et al., 2009, p. 892). Moreno believed that the reasons for escape had less to do with the girls' individual characteristics than their position in an underlying social network. He studied these networks and theorized about their evolution over time, inventing the field of sociometry and its most abiding tool, the sociogram (Moreno, 1934; Moreno & Jennings, 1938).

By the 1980s, social network analysis was well-established as a sub-field within the social sciences (Carrington et al., 2005). In the 1940s and 50s, matrix algebra and graph theory were enlisted to precisely describe social constructs such as groups and social circles. In the 1960s, social network approaches were taken up in anthropology — society was analyzed as networks, and kinship systems were described through relational algebras. For sociologists and anthropologists, social networks promised to move the field beyond the then-dominant structural-functionalist approach to analyzing society. In the 1970s, sociology deepened the analysis of social structure using networks, contributing one of the most cited articles to date:

the “Strength of Weak Ties” by Granovetter, whose arguments continue to shape social network research.

On the back of these developments, the field of social networks and health emerged (L. Berkman & Krishna, 2014). Early work in the 1970s and 80s established the association between structural features of the personal social network and mortality “from almost every cause of death” (L. Berkman & Krishna, 2014). Following the pathbreaking work by Berkman and Syme (1979) showing a strong and protective relationship between social network strength and size and mortality, there burgeoned a literature measuring ego-centric networks,[1] finding in them explanations for health outcomes from physiological stress to mental health (Aiello, 2017). In the 80s and 90s the field took a psychosocial bent, addressing itself to the role of networks in delivering social support and therefore shaping health. For example, related to survival after myocardial infarction (MI), not only did observational studies find robust associations between social support and health (L. F. Berkman, 1992; Gorkin et al., 1993; Orth-Gomer et al., 1988; Ruberman et al., 1984), some intervention studies did as well (Friedman et al., 1986; Kallio et al., 1979; Rahe et al., 1979).

Most recently, invigorated by the emergence of network science as an ecumenical discipline bringing together physicists, statisticians, sociologists, and others (Carrington et al., 2005), scholars have turned towards the emergent properties of networks as a determinant of health, studying the structural properties of both sociocentric and egocentric networks. This latest turn

---

[1] Egocentric – A social network data structure consisting of participants (egos) and their direct connections (alters). There is no information collected on relationships between one ego and another ego, or one ego’s alters and another ego’s alters.

Sociocentric – A social network data structure consisting of a defined group of people as well as all possible connections between them. e.g. a network detailing all friendships between the Population Health Sciences PhD program in 2018.

is exemplified by the investigation of *social contagion* (Christakis & Fowler, 2013), an area which was spurred on by a ground-breaking study arguing that obesity is transmissible along network ties (Christakis & Fowler, 2007).

By and large, the literature on social networks and health has not been concerned with material resources — a fundamental cause of individual and population health (Link & Phelan, 1995). Berkman and Krishna (2014), in a wide-ranging review, consider networks as a mediating structure for the effects of macrosocial causes on individual health. This mediation is conducted through six mechanisms which I group here into four; those related to cognition, affect, biomedical causation, and material resources. Related to cognition are mechanisms such as social influence, which involves the diffusion of norms relevant to behavior. Related to affect are mechanisms such as social engagement, negative social interactions such as conflict, and social support. Related to biomedical causation of illness are contagious processes that happen in the network through person-to-person contact. Finally, those related to material resources involve the provision of housing, financial support, etc. This last set of mechanisms that focusses on material resources is relatively neglected, however (L. Berkman & Krishna, 2014).

To be sure, some aspects of social support as it is currently theorized and operationalized in social epidemiology involve the transfer of material resources from one person to another. For instance, financial support and instrumental support are often measured in egocentric social network studies which aim to assess social support. In the Berkman and Krishna definition, “Instrumental support refers to help, aid, or assistance with tangible needs such as getting groceries, getting to appointments, phoning, cooking, cleaning, or paying bills” (2014, p. 244). This and the concept of financial social support relate to the central thesis of this project: that resources move from person to person as a result of their relationship, potentially with effects on health.

An important gap in social networks and health research, however, is that in assessing the amount of support social contacts give to one another, no account is taken of the resources that contacts have at hand. To put it crudely, it makes no difference to the usual measures of support whether the social contact whose support is being reported is penniless or a millionaire. Ironically, despite roots in social stratification and social conflict scholarship (Song et al., 2018), conceptualizations of social capital in social epidemiologic research suffer a similar limitation.

Mixed origins of social capital epidemiology

At first blush, the term "social capital" appears to suggest its meaning. "Capital" is commonly understood to be a resource that enables action, particularly production. "Social" locates these resources in some aspect of society or social organization. Indeed, social epidemiologic research often uses definitions that are consistent with what this term appears to mean. Kawachi and Berkman's (2014) review of social capital and health studies defines social capital drawing on Pierre Bourdieu: "the aggregate of the actual or potential resources which are linked to possession of a durable network of more or less institutionalized relationships…" (Bourdieu, 1986; Kawachi & Berkman, 2014). But social epidemiologic research in this area regularly uses measures that seem unrelated to the ideas of *networks* and *resources*. Most studies of social capital and health use what has been termed the "cohesion-based" approach to social capital. In this approach, social capital is measured using "(1) individual attitudes, perceptions, and cognitions about the group to which they belong, also referred to as cognitive social capital; and (2) actual behaviors (e.g. whether individuals participate in informal and formal social organizations), also referred to as structural social capital" (Kawachi & Berkman, 2014, p. 301). In seeking to locate this project in the social capital and health literature, I initially attempted to reconcile the measures with the seemingly unrelated definitions. I was not successful. What I

have found instead is that *social capital* is a syncretism - signifying multiple sets of ideas drawn from different literatures. I discuss the two sets which are most closely related to social epidemiologic research: one centered on the work of Pierre Bourdieu, and the other on Robert Putnam's.

Bourdieu and the beginnings of social capital

In the 1983 essay "The Forms of Capital", French sociologist Pierre Bourdieu set out to explain why and how life chances are unevenly distributed across society and how this uneven distribution is maintained (Bourdieu, 1986). He found his explanation in the uneven distribution of capital, a thing that "contains a tendency to persist in its being… a force inscribed in the objectivity of things so that everything is not equally possible or impossible" (Bourdieu, 1986, p. 241). Bourdieu posited three kinds of capital: economic, cultural, and social. Whereas extant economic theory had concerned itself with economic capital, the latter two were of interest since in Bourdieu's view, it was not possible to account for the persistence of social structure without them. Cultural capital was defined as "symbols and meanings [of the dominant class], which are misrecognized and internalized by the dominated class as their own" (Lin, 2017, p. 29), and social capital was defined with reference to material resources available to a person through members of their social network: "capital (economic, cultural or symbolic) possessed … by each of those to whom [one] is connected" (Bourdieu, 1986, p. 29).

Bourdieu defined capital as labor accumulated in commodities, positing that each of the three kinds of capital can be converted into the other two forms. This *transubstantiation* of capital among the three forms was theorized to obscure the flow of economic capital from view. Cultural capital, for instance, might hide the reproduction of class structure through the intergenerational transmission of elite tastes by disguising those tastes as competence or intelligence and rewarding them with economic capital. Strongly influenced by Marx, Bourdieu's

account of social capital emphasizes its unequal distribution and its role in the reproduction of social class (Bourdieu, 1986; Lin, 2017). Social capital epidemiology, however, has tended to disregard this aspect of social capital. That is, social capital in social epidemiology, though it uses definitions that resemble Bourdieu's, is not understood to be fundamentally entangled with, and animated by, economic capital.

If Bourdieu provided the definitions that circulate in this academic literature, Robert Putnam provided the concepts and measures. But after Bourdieu and before Putnam, American sociologist James Coleman laid the groundwork for the explosively popular ideas that would come to be entrenched in public health research.

Coleman, Putnam, and the American development of social capital theory

In his 1988 article "Social Capital in the Creation of Human Capital" Coleman found in the idea of social capital a vehicle to advance his project of methodological individualism within sociology (J. Coleman, 1990; J. S. Coleman, 1988). He saw social capital as a way of describing the influence of social structure on individual action while affording the actor an "engine of action" and avoiding an over-socialized conception of individuals (Wrong, 1963). Coleman defined social capital capaciously, as "a variety of different entities… with two elements in common: they all consist of some aspect of social structures, and they facilitate certain actions of actors… within the structure" (J. S. Coleman, 1988, p. S98). Coleman warned that his notion of social capital is an "unanalyzed concept" that communicates to the analyst that an actor taking an action benefits from some resource that depends on social structure. He advised that there be "a second stage in the analysis to unpack the concept, to discover what components of social organization contribute to the value produced" (J. S. Coleman, 1988, p. S101).

Despite his tentative definition, however, Coleman went on to describe “forms of social capital” which could be conceived as domains for the analyst to explore. These were (a) obligations, expectations, and trustworthiness engendered by social relations; (b) information flows that are shaped by relationships; and (c) norms and sanctions which help to regulate behavior. In addition, Coleman pointed out two aspects of social structure which would engender social capital: closure — the propensity for a friend’s friend to be a friend, and appropriable social organizations — the tendency for voluntary organizations formed for one purpose to be useful in organizing collective action for another purpose. Coleman closed with a discussion likening social capital to a public good: “the kinds of social structures that make possible social norms and the sanctions that enforce them do not benefit primarily the person or persons whose efforts would be necessary to bring them about, but benefit all those who are part of such a structure” (J. S. Coleman, 1988, p. S116).

Citing Coleman as the authority who “put the term firmly and finally on the intellectual agenda” (R. D. Putnam, 2000, p. 24) and homing in on the public good aspect of social capital, Robert Putnam wrote an essay[2] (1995) and wildly popular book (2000) entitled “Bowling Alone”. Using the image of the lone bowler of contemporary America in contrast with the socially engaged league bowler of a golden, nostalgic past, Putnam argued that the stock of social capital in the US was in decline. Though the book did not offer an explicit definition, the essay echoed Coleman’s “forms of capital”[3], defining social capital as “features of social organization such as networks, norms, and social trust that facilitate coordination and cooperation for mutual benefit” (1995). Without conducting Coleman’s “second stage” of the analysis, Putnam presented an array of indicators ranging from political participation and volunteerism to reciprocity, honesty,

[2] This journal seems to be out of print, but there are copies of the article circulating on the internet. It is difficult to verify that they reflect the content of the published article
[3] That is, the domains which Coleman steered analysts of social capital toward

and trust as the traces of a social capital in decline. This approach has garnered some critique, most notably, Portes' observation that Putnam's attempt to demonstrate the existence and operation of social capital by documenting its supposed effects is tautological (Portes, 1998). Despite these shortcomings, Putnam's conception of social capital has left a lasting imprint on social epidemiologic literature, as evidenced by the measures most commonly used in this literature (Ehsan et al., 2019; Kawachi & Berkman, 2014).

In empirical studies these measures sometimes lead to findings which appear to be paradoxical. Kawachi and Berkman (2014) cite a study in an impoverished racial minority inner-city community in Birmingham. Investigators in the study found that "while high bonding capital — as measured by the level of trust and strength of ties between members of the community who share similar race and socioeconomic backgrounds — was associated with more mental distress, the opposite was true for network ties to others who came from different race/class backgrounds" (Mitchell & LaGory, 2002; Kawachi & Berkman, 2014). Findings like these, if not an artifact of confounding, seem related to the likely uneven distribution of resources across race and class among the population surveyed. Yet, no explicit account is taken of these under this strain of social capital theory.

More unsatisfying is the conceptual basis of social capital as deployed in epidemiologic studies. Bringing attention to the definition furnished by Coleman, "features of social organization such as networks, norms, and social trust that facilitate coordination and cooperation for mutual benefit," reveals some striking implicit assumptions. First, the definition assumes that social capital helps to explain the success or failure of collective, goal-oriented action. This appears to imply one of two things: either there are goals that are not only common to all community members, but also relevant to the health outcome under study; or this definition restricts the attention of the analyst to situations and health outcomes for which such goals exist and are

widely shared. Similarly, the definition appears to posit that whatever the health-relevant goals of the community are, they can be met by cooperation rather than, say, competition or contestation.

Critically, if social capital is primarily concerned with collective action, then it is difficult to understand how individuals in one community could be ascribed different levels of social capital, as is often done in epidemiologic studies. It is also difficult to apply this set of ideas to the study of health disparities, which may or may not be driven by conflicting interests in and across communities.

There have been some attempts to rescue Putnam-style social capital from its ontological challenges. Notably, Szreter and Woolcock (2004a), as part of a lively exchange with leading social capital scholars (Ellaway, 2004; Gakidou et al., 2004; Kawachi et al., 2004; Muntaner, 2004; Navarro, 2004; R. D. Putnam, 2004; G. D. Smith & Lynch, 2004; Szreter, 2004; Szreter & Woolcock, 2004a, 2004b), promised to explicate a "comprehensive but grounded theory of social capital" (2004a, p. 650), unifying what they saw as the most prevalent competing interpretations of social capital theory among social epidemiologists. They make a distinction between bridging, bonding, and linking social capital. Bonding capital is defined as "trusting and co-operative relations between members of a network who see themselves as being similar" (2004a, p. 654) and bridging capital as "relations of respect and mutuality between people who know that they are not alike in some socio-demographic (or social identity) sense (differing by age, ethnic group, class, etc)" (2004a, p. 655). Linking social capital, defined as a special case of bonding social capital, is defined as "norms of respect and networks of trusting relationships between people who are interacting across explicit, formal or institutionalized power or authority gradients in society".

Regarding the abovementioned paradoxical finding from Birmingham, that bonding capital was associated with mental distress while bridging capital was not could be interpreted using the idea of linking capital. i.e. When poor black people form links with, say, wealthy white people, they are interacting across a power gradient and that kind of interaction might have different health effects than interactions with peers. It is not clear, though, that using this concept helps us to understand the etiology of mental distress in the community. In particular, it is not explained why seeing oneself as similar or different would shape the effect of a given relationship on one's health.

In the section of their article focused on the explication of their theory of social capital, Szreter and Woolcock do not offer an explicit definition of social capital. Instead, it is left to the reader to glean from two prior mentions of Putnam's definition that Szreter and Woolcock themselves endorse the view that social capital is "the nature and extent of networks and associated norms of reciprocity" (2004, p. 654; 2004a, p. 650). Rather than offer a comprehensive theory, Szreter and Woolcock restate and renovate an old one, missing an opportunity to clarify the theorized relationships between social capital, its measures, and population health and leaving fundamental assumptions unexamined.

Alejandro Portes (1998) offers clearer conceptual guidance. His theory of social capital begins by making the observation, consistent with Bourdieu, that "[t]o possess social capital, a person must be related to others, and it is those others, not himself, who are the actual source of his or her advantage" (1998, p. 7). He distinguishes the resource that underpins this advantage from the ability to obtain the resource through the relationship — a distinction that is necessary to make to avoid Putnam's tautology. Portes makes a further distinction between the motivations of the person potentially receiving the resource from the motivations of the person potentially gifting it or making it available on concessionary terms - parties I will call the "receiver" and

“sender” for this discussion. The motivations of the receiver are obvious as she benefits from resources in her network; it is those of the sender that Portes examines, calling them “sources” of social capital.

Under Portes’ account, a person might provide resources to a social contact for consummatory reasons or instrumental reasons. The former are reasons related to a sense of obligation arising either from norms internalized through childhood or from bounded solidarity — solidarity based on the recognition of shared fate and restricted to the group whose fate one shares. The latter are reasons related to the expectation that one will be paid back the resource in full, though at some undetermined future date and in some undetermined form — i.e. one will be able to “call in the favor”. A second instrumental reason for making a resource available is if the sender and receiver are both embedded in a social structure such that the sender will be rewarded with status or prestige by the collectivity, and/or that the collectivity will act as a guarantor of the “debt” in some way. In summary, to make sense of social capital, Portes argues that one must distinguish resources from relationships and explain why a sender might make a resource available to receivers she is related to.

These ideas prove useful in my project and will be reiterated in a later section, though in a different guise. In making the case that the resources held by an individual’s family members potentially shape the health of that individual, it is necessary to explain whether and why one family member might make resources available to another. To measure the impact of these networked resources, it is necessary to make a distinction between the network and the resource. The concepts and measures used in social capital epidemiology, by contrast, are not useful in this project.

An epidemiology of care

In this project, I aim to study the relationship between individual health and the resources embedded in kinship networks in two community surveillance sites in South Africa. These sites are largely rural, one being in Kwa-Zulu Natal Province, and the other in Mpumalanga Province on the border with Limpopo Province. I take a materialist view, focusing on the network as an apparatus through which resources (Link & Phelan, 1995) flow from individual to individual, shaping health. In this section I set out the substantive theory that I will draw on to make this case.

The notion of social capital as theorized by Bourdieu, since it is concerned with material resources and their unequal distribution, provides a framework for thinking through how social capital relates to economic capital. In the context of my proposed study, the framework proposes a relationship between economic resources held by family members, the individual's economic resources, and the individual's health. There are two limitations to this strain of theory, however. The first is that in the course of framing resources held by social contacts as *capital*, Bourdieu and others in his tradition implicitly claim that rational actors wield the capital strategically, maximizing some future or present utility (Bourdieu, 1986; Lin, 2017). This view forecloses the possibility of benefitting from relationships without taking goal-oriented actions — the possibility that a child benefits from the wealth of his mother whether or not he chooses to and whether or not he recognizes his mother as a source of material benefits. The second limitation is that Bourdieu seems to take for granted the resources that circulate among members of the family not for the reproduction of social class, but the reproduction of life itself. Cooking, cleaning, mending clothes and repairing appliances for family members is a kind of work that depends on resources, is unevenly distributed, and has an impact on the health of

both workers and beneficiaries. Since I am interested in how resources shape health along family ties, it is important to explicitly account for this work.

To do so I draw on the idea of reproductive labor or *care* as conceptualized by socialist feminists: “the work involved in reproducing and sustaining both biological and social life over time and across generations” (Buch, 2015) and, in particular, work that is produced outside of the market (Bhattacharya, 2017). This work is performed and experienced through personal relationships that are themselves embedded in moral economies — obligations and entitlements that arise as a result of repeated interactions in sustained personal relationships (Carrier, 2018). A central concern of socialist feminist theory is the role of care in deepening inequalities: reproductive labor is unequally distributed by gender, race, and nationality. It is persistently undervalued when it is exchanged for wages (Tronto, 2014). I draw on this strain of theory to make visible the work that is constantly conducted and exchanged in service of maintaining human life. I make the assumption that this work depends on material resources.

I posit *care* as the mechanism through which the health outcomes of an ego are determined by the material resources of alters[4]. By doing this I assume that material resources enable care so that the amount that a son can provide to his mother is partially determined by his own access to material resources. This assumption is consistent with that made in social capital literature that “network members resources are valuable social resources non-redundant with personal capital” (Song et al., 2018, p. 241).

[4] Ego – A focal node in a network. e.g. a person under consideration in the context of a kinship network.

Alter – The nodes connected to the ego. e.g. a person’s direct relations in a kinship network.

My focus on the material aspects of care does not deny its affective dimensions, which have been the basis of psychosocially oriented social network and health research since the 1980s. Instead, I aim simply to bring attention to a mechanism that has not been as well investigated and that is likely to have an important influence on health outcomes.

At a macro-scale, care can be understood as a “scarce resource that circulates through complex webs of kinship and intergenerational relations” (Buch, 2015). This circulation is shaped by institutions such as the state, the family, the labor market, and the non-profit sector (Pfau-Effinger, 2005). Each of these has a role in the provision of care, particularly to the vulnerable such as the very young and the very old. The role that each institution plays is dialectically determined by the roles that others play. The configuration of institutions and their role with respect to care in a particular society are collectively termed the *care regime* (Buch, 2015). Because of this mutual determinism, to explain population patterns of health and illness in terms of care in South Africa, it is useful to describe the South African care regime. I focus, here, on the role of the family and of the state.

The South African care regime

The form of the current South African care regime can be traced back to the establishment of a welfare state in the 1920s which extended to white people “enormous assistance — in terms of public schooling, public health and psychiatric care, social welfare and social work programs… - to rise up the economic and social hierarchy” (Seekings & Moore, 2014). Through these programs, members of parliament sought to protect the white and coloured populace from destitution and to preserve the racial income hierarchy (Seekings, 2020). From the late 1800s, there had been social upheaval: black people were alienated from their land, new schemes of taxation forced peasants to become wage laborers, and the burgeoning mining industry sucked

working age black men into their employ and away from their rural homes (Jochelson, 2001; Sooryamoorthy & Makhoba, 2016). There were colonial restrictions on the ability of black people to migrate with their families, resulting in “stretched households” — households whose members resided in different homes. These restrictions evolved through the Apartheid era beginning in the 1940s and were finally lifted by the end of the 1980s (Posel, 2016).

In the dying throes of the Apartheid regime, black South Africans were extended the same entitlements as their white and colored counterparts while the basic design of the welfare system remained the same as it was in the 1920s (Seekings, 2020). This change coincided with major demographic and economic transitions. The repeal of apartheid laws allowed black people to migrate across the country and participate in the economy without any legal constraint (Seekings, 2008). New patterns of temporary labor migration, however, continued into the post-apartheid period and in response to dizzying levels of structural unemployment (D. Casale & Posel, 2020; Madhavan & Brooks, 2015; Posel, 2016). With women’s increased labor market participation and widespread poverty, marriage rates had declined beginning in the 1980s, decoupling childrearing from marriage and the nuclear family (D. Casale & Posel, 2020; Madhavan et al., 2014). Finally, by the end of the first decade of the 2000s, the HIV/AIDS epidemic had killed over 2 million working age South Africans, leaving their children and adult dependents in its wake (Seekings, 2008).

One strategy for living in this environment has been for families to organize households around the availability of resources rather than around the nuclear family unit. For instance, in Mpumalanga, the old age pension drives the composition and structure of households: economically independent adult children leave their parents’ homes to establish their own smaller households while indigent family members and dependent children are absorbed into the households of pension-eligible family members (Makiwane et al., 2017). More generally,

South African households have been described as fluid and porous — household structure changes in response to economic exigencies and individuals may simultaneously belong to multiple households (Kelly, 2018; Seekings, 2008). Households have diverse and complex structures spanning multiple generations (Hoddinott et al., 2018; Madhavan et al., 2017, 2017; Madhavan & Brooks, 2015; Schatz et al., 2015; Wittenberg & Collinson, 2007). Their degree of complexity has increased over time (Schatz et al., 2015; Wittenberg & Collinson, 2007).

Another strategy families adopt in the face of widespread poverty and unemployment has been to share resources across households. A 2005 nationally representative survey found that over half of the respondents reported having recently given money, goods, food, or other items to members of their family who lived in a different household (Everatt et al., 2005). Another study found that in 2014, about one in five households received cash transfers from family members and about two in five households participated in private cash transfers either as sender or receiver (Posel, 2016). Among receiving households, transfers were substantial, amounting to as much as 40% of household income (Posel, 2016; Ssebagala, 2021). These transfers are progressive — they tend to move from wealthier to poorer households and from employed to unemployed people (Ssebagala, 2021). They also tend to be targeted at households that are headed by women and that include children, non-resident parents, or labor migrants (Posel, 2016).

The welfare system of South Africa is a crucial resource for the care that circulates among family members. The government spends about 3% of GDP on social grants which take the form of unconditional cash transfers. About 50 percent of households in the country receive these every month (Kelly, 2019; Moore, 2020). Three types of grant account for the majority of grant spending: the child support grant, which is paid to the primary caregiver of a child under the age of 18; the old age grant, which is paid to adults over the age of 60, and the disability

grant, which is paid to people whose health or disability renders them unable to work. These grants have been found to alleviate extreme poverty (Neves et al., 2009). When received by women, they have been shown to be of benefit to family members, not just the direct recipients (Button & Ncapai, 2019; Case & Deaton, 1998). Old age and disability grants have been shown to enhance the agency of direct recipients (Button & Ncapai, 2019; Kelly, 2018; Madhavan et al., 2017), even if they sometimes occasion family conflict over resources (Kelly, 2018, 2019).

The care regime that is produced by the combination of welfare policy and cultural norms regarding kinship and caregiving has been described as ‘familialist’ (Button & Ncapai, 2019; Mahon, 2018). It lays the responsibility for providing care for children, people with disabilities, and older people on their families rather than mounting state institutions to directly provide care (Kelly, 2018). Though family members do tend to take up the responsibility of, for instance, caring for deceased relatives’ children (Wittenberg & Collinson, 2007), Seekings (2008) argues that obligations based on kinship should not be assumed to be uniform and immutable. Rather, kinship-based claims and obligations are negotiated, contested, and cultivated. The transfer of resources between family members, is based both on norms of obligation and expectations of reciprocity. For close kin, norms of obligation dominate decisions around caregiving and for distant kin, expectations of reciprocity do. Like state welfare support, individuals might also make distinctions among potential recipients of care based on whether they appear to be deserving or not — a destitute child might be more likely to receive support than an unemployed adult with a substance abuse problem.

Adopting a familialist welfare policy orientation has implications for gender equity. Grants are a partial safety net for poor and unemployed people (Moore, 2020; Seekings & Moore, 2014). This safety net is indirect, however. It is comprised of the disability, old age, and child support grants along with the care work of family members who receive these grants and other income (Button

& Ncapai, 2019). In households across the country, women do a disproportionate amount of care work including cooking, housework and household shopping, and taking care of the ill or disabled (Fakier & Cock, 2009; Makiwane et al., 2017; Moore, 2020; L. Patel & Mavungu, 2016; Razavi, 2015). This gendered division of care, when reinforced by welfare policy, maintains gender inequity (Moore, 2020).

***

In the chapters to follow, I draw on Ecosocial theory to investigate the relationship between health and health behavior outcomes and resources held by family members (Krieger, 2001). In interpreting these analyses, I assume that differences in health outcomes between the sub-groups of people represented in the datasets are not produced by innate characteristics but are a result of biological processes that ensue as a result of various chemical, psychological, biological, and other exposures that are themselves arrayed by the conditions under which people live, work, and play. The conditions are not fixed, but ever changing under the constantly evolving social, political, and ecological milieu of the regions of South Africa that host the participants of this study as well as their families, friends, and neighbors. Care by family members is a crucial aspect of these conditions.

Just as epidemiologic theory delimits what questions are useful to ask, and what explanations are available to answer them (Krieger, 2001), statistical theory delimits what quantities are useful or even possible to estimate (Schwartz et al., 2016). Traditional epidemiologic methods for assessing causal effects typically depend on several crucial statistical assumptions: that the measurements made on each individual or observational unit represent an independent and random draw from some underlying joint probability distribution and that all confounders of the

relationship between exposure and outcome of interest have been measured and appropriately accounted for in the analysis.

By assumption, social networks represent systems where a given person's individual treatment or outcome might affect the outcomes of those individuals connected to her, violating the first assumption. Even when all relevant causes of exposures, and outcomes have been measured, however, conducting an analysis that fails to account for potential influence across individuals violates the second assumption and might lead to biased results. In investigating the role of family connections in shaping health, it is necessary to define causal estimands and use statistical models that account for the structure of dependence among observations. These estimands, models, and assumptions are described in detail in the appendices.

Chapter 1 examines the spillover effect of a randomized intervention — the offer of a financial micro-incentive — on HIV testing among 15,000 people in Kwa-Zulu Natal, South Africa. We ask whether it makes a difference to one person's HIV testing behavior whether their family members were offered the incentive. Chapter 2 examines the relationship between household wealth and physical functioning among a cohort of 5,000 older adults, and Chapter 3 investigates the relationship between household assets and mortality across a region of about 100,000 people. In the latter two chapters, we ask what proportion of the effect of wealth on physical function (or mortality) is accounted for by people's own household wealth, and what proportion is accounted for by the wealth of family members living in other homes. In these chapters, we take advantage of auto-g-computation, a relatively new set of tools for dealing with causal inference under network interference. In the first, we borrow a technique from spatial econometrics to reason about spillover effects in a family network.

Chapter 1

Family network spillover of micro-incentives for HIV testing: evidence from a community-randomized controlled trial

Background

Despite the promise of HIV treatment as prevention (TasP) (Rodger et al., 2016, 2019), the rapid expansion of HIV treatment programs in East and Southern Africa has not led to a commensurate reduction in new HIV infections (Joint United Nations Programme on HIV/AIDS (UNAIDS), 2019). There are a number of community-randomized trials that test interventions to close this gap. The Home-based Intervention to Test and Start (HITS) cluster randomized trial aims to assess whether the offer of a small once-off financial incentive and a male-targeted HIV-specific decision support application increase uptake of HIV testing and linkage to HIV care. Using data from this study, we examined the family spillover effect of the financial incentive on HIV testing; that is — we examined the effect of offering a financial incentive to an individual's family members on that individual's HIV testing behavior.

HIV testing coverage is high and increasing — it is estimated that only 13% of people living with HIV in east and southern Africa do not know their status (Joint United Nations Program on HIV/AIDS, 2020). Given the stubbornly high rate of new infections globally, this group is likely to increasingly be comprised of people who are not currently reached by HIV prevention, testing, and treatment programs and who face a high risk for HIV acquisition and transmission (Baral et al., 2019; Ortblad et al., 2019). It is imperative, therefore, to continuously design and implement diverse strategies for reaching people living with HIV with testing services in order to maximize HIV treatment coverage among people who are currently unreached.

Several randomized controlled trials have demonstrated that financial incentives can improve the uptake of HIV testing for the individual receiving the incentive (Bassett et al., 2015; Choko et al., 2019; Kranzer et al., 2018; Montoy et al., 2018). The effectiveness of these incentives depends on their value and form, and they operate through multiple causal mechanisms

(Chamie et al., 2018, 2020; Korte et al., 2019). For instance, a recent qualitative study in Uganda showed that among men, monetary incentives simultaneously addressed structural, interpersonal, and individual-level barriers to HIV testing (Ndyabakira et al., 2019). Several randomized controlled trials have investigated spillover of the effect of financial incentives offered to one person on the health of another, but these tend to focus on caregiver-child relationships (Handa et al., 2014; Liu et al., 2019; Sherr et al., 2020; Stoner et al., 2021; Yotebieng et al., 2016). Our study investigates spillover of financial incentives among adult family members.

There is evidence that social networks are an important determinant of HIV-related health. Social contacts may influence each other's health outcomes by transmitting information, norms, resources, and support (L. F. Berkman et al., 2000; M. Casale et al., 2015, 2019; Chanda et al., 2017; Christakis & Fowler, 2013; Conserve et al., 2019; Hermanstyne et al., 2018; Kawachi & Berkman, 2014; Mulawa et al., 2016; Musheke et al., 2013; Nelson et al., 2015; Perkins et al., 2015; Takada et al., 2019; Yamanis et al., 2016). This body of evidence, however, is largely observational. As a result, it is difficult to measure the effect of one person's characteristics on another person's outcomes (influence) as distinguished from the effect of two people's shared characteristics on their likelihood of being connected (homophily) (Shalizi & Thomas, 2011). This study overcomes such methodological limitations by examining the effect on individual HIV testing uptake of offering a randomly assigned financial incentive to family members.

We investigated the question of the impact of social network on HIV testing by constructing a sociocentric family network using routinely collected demographic and health surveillance data. We hypothesized that subsequent to having family members offered the financial incentive, the likelihood that an individual consents to an HIV test increases whether or not that individual was offered the incentive. We tested this hypothesis using a structural mean model estimated using

the Generalized Method of Moments (GMM) (Hansen, 1982; Hernán & Robins, 2018). Our use of the GMM approach, which is relatively novel for epidemiologists, is motivated by a desire to take advantage of the randomization scheme in HITS while making as few additional assumptions about the data as possible.

This research makes a substantive and methodological contribution. Substantively, we offer an approach to improving the effectiveness of incentive schemes for HIV testing. Methodologically, we offer a causal inference framework for the evaluation of network effects in population-based research. By studying the impact of relationships that span randomization clusters, we show that the assumption of independence of clusters is violated in this setting, with implications for the future conduct of cluster-randomized trials.

Methods

Setting and Participants

The HITS study is an ongoing community-randomized controlled trial in the Hlabisa sub-district of the uMkhanyakude district in northern KwaZulu-Natal, South Africa. The study is nested in the Africa Health Research Institute's (AHRI) population-based demographic and HIV surveillance platform including over 60,000 residents living in an area of 432 km^2. In this platform, trained field workers visit all households annually and interview a key resident informant each time. The survey records demographic information including the parents, co-residents, and conjugal partners of each individual in the household. In addition, all residents aged 15 years or older are offered home-based rapid HIV testing annually.

Individuals were eligible for the HITS study if they were 15 years or older at the time of the surveillance visit, resided within the AHRI surveillance area, agreed to participate in the annual

HIV surveillance, and provided written informed consent for trial participation. Individuals were not eligible to participate in the trial if they refused to participate in AHRI HIV surveillance, reported being already on ART, or were mentally or physically unable to provide consent. The study is registered at the National Institute for Health's ClinicalTrials.gov (# NCT03757104). Enrolment started in February 2018 and follow up will be completed in December 2021. Further details are available in previous publications (Mathenjwa et al., 2019; Tanser et al., 2021).

Randomization

The AHRI surveillance area was divided into 45 communities which were randomized to the study interventions using a 2x2 factorial design. The interventions were a financial micro-incentive for HIV Testing, and a male-targeted HIV-specific decision support program called EPIC-HIV. In this study we consider the effect of the micro-incentive only, ignoring EPIC-HIV since the latter intervention was randomly assigned independently of the former.

For the micro-incentive, stratified randomization (see Figure 1.1) was conducted in order to ensure that both study arms had similar baseline HIV incidence rates among women aged 15 — 30 years — a group with disproportionately high incidence (Chimbindi et al., 2018; Simbayi et al., 2019). The 45 communities were grouped into four strata based on these rates. With respect to the micro-incentive, the intervention arm consisted of four randomly selected communities from each of the four strata (16 communities total). The control arm consisted of seven communities from each stratum, with an additional community in the stratum with second highest incidence (29 communities total). Although the financial incentive intervention was randomly assigned, the order in which households were visited was not. The study is an open label trial.

Intervention

All HITS-eligible individuals in control communities (control arm) were offered rapid HIV testing as per the AHRI HIV surveillance protocol. Those living in intervention communities (intervention arm) were offered a micro-incentive for rapid HIV testing consisting of a food voucher valued at ZAR 50 (~USD 3).

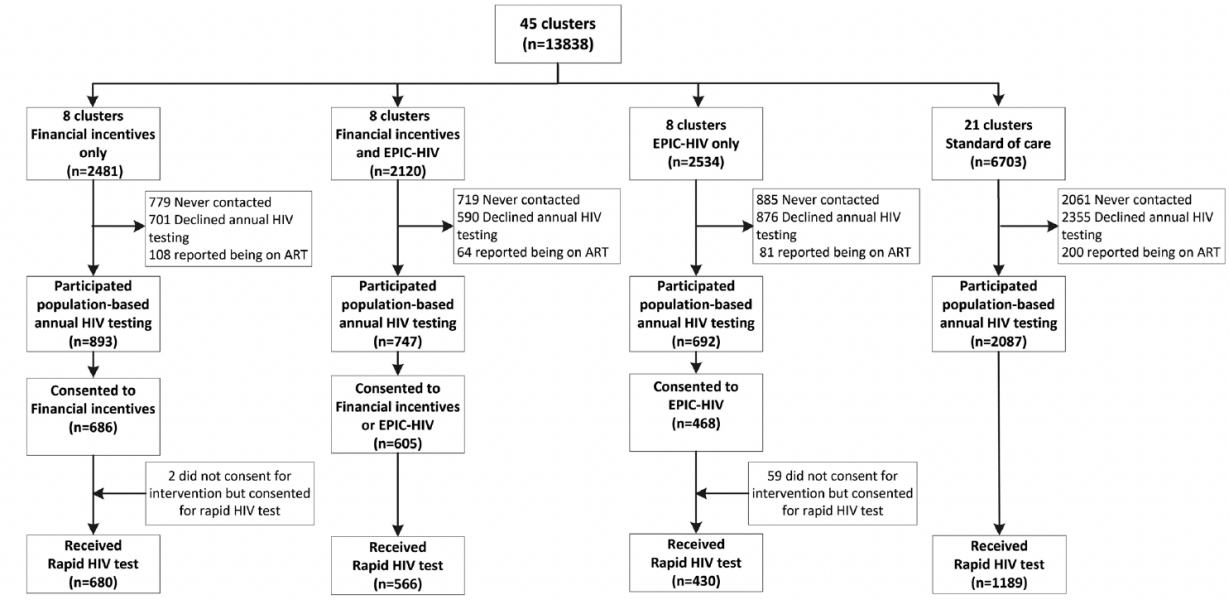

Figure 1.1: Flow diagram for HITS trial (Tanser et al., 2021)

Social Network

A sociocentric family network was constructed among all HITS-eligible individuals in the study area. Nodes in this network represent individuals, and ties represent family relationships. Ties were ascertained using information on household relationships as follows.

First, an undirected network (i.e. a network where ties are bi-directional) was constructed using parent-child and conjugal relationships as recorded in the AHRI surveillance data. Then ties were added between each pair of nodes that had a node in common. This means that, for instance, each person was directly connected to her parents as well as her grandparents. Similarly, each person was connected to their conjugal partner as well as their conjugal partner's parents. To this network, we added a connection between two individuals if they were currently or previously members of the same household and had a family relationship. We call this the *whole family* network — this network was used for calculating descriptive statistics about connections across intervention communities.

Finally, we converted the resulting undirected network into a directed one. Undirected ties were replaced with directed ties that capture the time ordering of the testing offer. Ties point from individuals who had an earlier study visit towards individuals who had a later study visit. This precluded the possibility that a person's future intervention status would affect his family members' current study outcomes. For each pair of individuals whose study visits were on the same day, we added ties in both directions. We refer to this network as the *effective family network — this network was used for calculating causal effect estimates*.

We use *family member* to mean any person that the focal person is connected to in the given sociocentric family network.

Measures

The outcome of interest was consent for rapid HIV testing at the study visit. Participants who consented to rapid HIV testing were considered to have obtained the outcome. The exposures of interest were *individual offer of financial incentive* and *family offer of the incentive* (abbreviated as *individual treatment* and *family treatment*, respectively). The *family treatment*

value is defined as the count of family members in a given network who were offered the financial incentive. In the effective family network, only family members who were offered the financial incentive before a given individual were counted towards *family* treatment for that individual. In the whole network, all family members who were offered the incentive were counted. A *network size* variable was calculated as the count of family members each individual has. These quantities were calculated using the whole family network and the effective family network separately.

Descriptive Analysis

In the descriptive analysis, we report sample characteristics, show the pattern of network connections between communities, and describe the composition of network connections for study participants. These variables are calculated using the whole family network. The exception is that family treatment is calculated using the effective network.

Using the effective family network, we also examine heterogeneity of the effect of individual treatment on HIV testing uptake across strata defined by dichotomized *family treatment* (*≥1 vs. 0)* and strata defined by ordinal *family treatment* (levels 0,1, 2, 3, 4, ≥5). In each case we fitted a semiparametric log-binomial model with two-way multiplicative interaction terms encoding the extent to which the causal effect of individual treatment is modified by family treatment. This analysis allows one to understand how the causal effect of individual treatment varies depending on the level of family treatment, but it does not specify nor estimate the causal effect of family treatment since this estimate would likely be confounded by network size and spatial distribution of family members.

Models

Using the effective family network, we carried out GMM estimation to fit a semiparametric multiplicative structural mean model of potential HIV testing consent ($Y_i^{t,f}$) under the intervention of individual treatment (t) and family treatment (f), conditional on network size and spatial distribution of family members (X). i.e. We model the person's potential HIV testing consent had, possibly contrary to fact, individual and family treatment been set to values t and f:

$$log\frac{P[Y_i^{t,f} = 1|X]}{P\left[Y_i^{0,0} = 1|X\right]} = \theta_1 t + \theta_2 f + \theta_3 t \times f$$

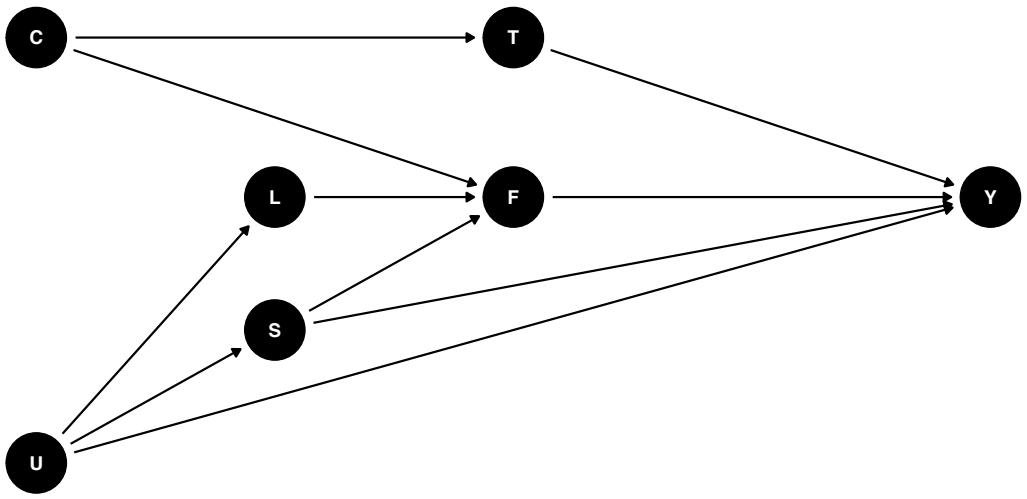

Figure 1.2: Directed Acyclic Graph (DAG)

DAG showing confounding structure for the relationship between family treatment (F) and consent for HIV testing (Y). C is Intervention status of community. T is Intervention status of individual. S is Family network size. L is spatial distribution of family members. U is Unmeasured potential confounders. The graph shows that the only potential confounders of F are mediated by network size S and spatial distribution of family members L.

While outcome regression, i.e. estimation of thetas in the model in displayed above by also assuming a model for $P\left[Y_i^{0,0} = 1|X\right]$, is far more common in epidemiology, its use is only valid under correct specification of the model for $P\left[Y_i^{t,f} = 1|X\right]$ and provided that X includes all confounders for the joint effects of t and f, a necessary condition to recover unbiased estimates of theta parameters even in a randomized trial such as HITS. In our setting, this condition is likely to fail (See Figure 1.2). Whether or not one's family member receives a financial incentive partially depends on the size of one's family network, and the spatial distribution of family members. People who have more family members and whose family members are dispersed across communities are more likely to have family members that were offered the financial incentive. The size and spatial distribution of one's family network, in turn, might be related to HIV testing through pathways other than the one involving the financial incentive.

Using outcome regression requires conditioning on network size to adjust for confounding, which in turn would require assumptions about the functional form of the relationships among network size, family treatment, individual treatment, and the outcome. On the other hand, using a structural mean model only requires assumptions about the relationships among family treatment, individual treatment and the outcome. The semiparametric structural mean model puts fewer restrictions on the observed data distribution, allowing for randomization-based inference.

Estimation

We carried out design-based statistical inference to estimate the parameters of the structural model, using the fact that treatment is randomized, and therefore independent of counterfactual

outcomes (Hernán & Robins, 2018). We constructed a number of *moment conditions* — moments of the population distribution of the data — involving the structural mean models, which are equal to 0 only at the true population value of the model's theta parameters. We estimated these moments using observed data by finding the parameter values that render the empirical analog of moment equations approximately zero. The set of model parameters that achieve this goal define GMM estimates.

Specifically, we define $L_i^{0,0}(\theta) = Y_i e^{-\theta_1 T_i - \theta_2 F_i - \theta_3 T_i F_i}$ where T_i and F_i are person i's observed individual treatment and family treatment, respectively, and $\theta = (\theta_1\ \theta_2\ \theta_3)^T \in \mathbb{R}^3$. $L_i^{0,0}(\theta)$ can be thought of as an approximation of $Y_i^{0,0}$ only at the true theta, as their means coincide. Because of randomization, $L_i^{0,0}$ is mean-independent of T_i and F_i. This means that for any function $f(T_i, F_i)$ that has a mean of zero, we know that, in expectation, the product of $L_i^{0,0}(\theta)$ and $f(T_i, F_i)$ is zero. i.e.:

$$E\left[L_i^{0,0}(\theta) \times f(T_i,\ F_i)|X\right] = 0.$$

Because we estimate three unknown parameters, f must be a vector of at least three functions, all with mean zero. We could choose any set of three functions that each have mean zero, but for simplicity, we constructed f as:

$$f(T_i,\ F_i)\begin{pmatrix} T_i - \mu_{T_i} \\ F_i - \mu_{F_i} \\ T_i \times F_i - \mu_{T_i \times F_i} \end{pmatrix}$$

with

$$E[f(T_i,\ F_i)|X] = \begin{pmatrix} E\left[T_i - \mu_{T_i}\right] \\ E\left[F_i - \mu_{F_i}\right] \\ E\left[T_i \times F_i - \mu_{T_i \times F_i}\right] \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

where $\mu_{T_i} = E[T_i]$, $\mu_{F_i} = E[F_i]$, and $\mu_{T_i \times F_i} = E[T_i \times F_i]$.

Note that unbiasedness of the moment equation displayed above holds by virtue of the mean parameters $\mu_{T_i}, \mu_{F_i}, \mu_{T_i \times F_i}$ can be computed exactly from study design. We then estimated $\hat{\theta}$ as the value of θ which satisfies the following equation:

$$\frac{1}{n}\sum_{i=1}^{n} L_i^{0,0}(\theta) \times f(T_i,\ F_i) \approx \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Statistical Inference

GMM estimators are consistent and asymptotically normal under certain regularity conditions (Hansen, 1982). We used their asymptotic distribution for statistical inference, estimating the variance under two different assumptions:

a) that all observations are independent
b) that observations within communities are correlated, and observations between clusters are independent

Both a) and b) are not consistent with our understanding of the data generating process. We assume that observations might be correlated across family ties and possibly within communities. If this assumption is correct a) will tend to under-estimate standard errors for model estimates. On the other hand, it is likely that observations are more strongly correlated among household members and between individuals and their spatial neighbors, but weakly correlated between people in the same community who are not socially or spatially proximate. Assumption b), therefore might overstate correlation between observations. On the other hand, since network connections are much denser within communities than between communities, assumption b) likely accounts for most of the correlation that arises because of spatial and social proximity. In summary, assumption a) is likely too lenient and yields inferences that may

be anti-conservative, and assumption b) is likely too conservative. Further details on models and statistical inference are given in Appendix A.

We report model coefficients and 95% confidence intervals. These were calculated using robust standard errors under a) and b) respectively. Data analysis was conducted using the effective family network. Data visualization was conducted using both effective and whole networks. All data manipulation, visualization, and analysis was carried out using RStudio. Network computations were conducted using *packages* igraph and tidygraph and model-based analysis was conducted using *geepack* and *momentFit* (Csardi & Nepusz, 2006; Halekoh et al., 2006; Pedersen, 2019, 2020; Rich, 2018, p. 1; RStudio Team, 2018; Yan, 2002; Yan & Fine, 2004; Zeileis, 2004) .

Results

Descriptive analysis

Of 37,068 eligible residents, 15,675 participated in AHRI HIV surveillance. Among study participants, almost all cases (15,665/15,675) had complete outcome and exposure data. Study arms were balanced on age, gender, and network size (see Table 1.1). Being offered a financial incentive led to increased testing uptake. Almost two-thirds (64.7%, 3,647/5,637) of participants in the intervention arm consented to an HIV rapid test, whereas half (50.7%, 5,087/10,028) of the participants in the control arm consented. It was common for participants to have family members in different households (60.4%, 9,468/15,675) and different communities (42.2%, 6,613/15,675) (see Figure 1.3 and Table 1.1). These proportions were similar across study arms. Compared to people living in control communities, people in intervention communities were more likely to have family members living in an intervention community (85.1%,

4,799/5,638 vs. 9.4%, 945/10,037). Further descriptive results have been reported previously (Mathenjwa et al., 2019; Tanser et al., 2021).

Not all family relationships were ascertained: 37.9%, 14,059/37,068 of HITS-eligible individuals did not have their mother listed in AHRI demographic records and 63.7%, 23,625/37,068 did not have their father listed. Most of these relationships were likely not ascertained because study participants' parents were either not living in the surveillance area or not alive any longer, causing them to be ineligible for the surveillance. Ascertainment of these relationships was strongly patterned by age. Among individuals aged 15 - 25, 15.0% (2,321/15,458) were missing information on their mother and 48.3% (7,471/15,458) were missing information on their father. In contrast, among those over 55 years of age, these proportions were 87.9% (4,795/5,458) and 97.3% (5,310/5,458) respectively. For the overwhelming majority of individuals, it was possible to identify family members in the AHRI data. Only 2.7% (424/15,675) of HITS participants were not linked to any family members in the whole family network.

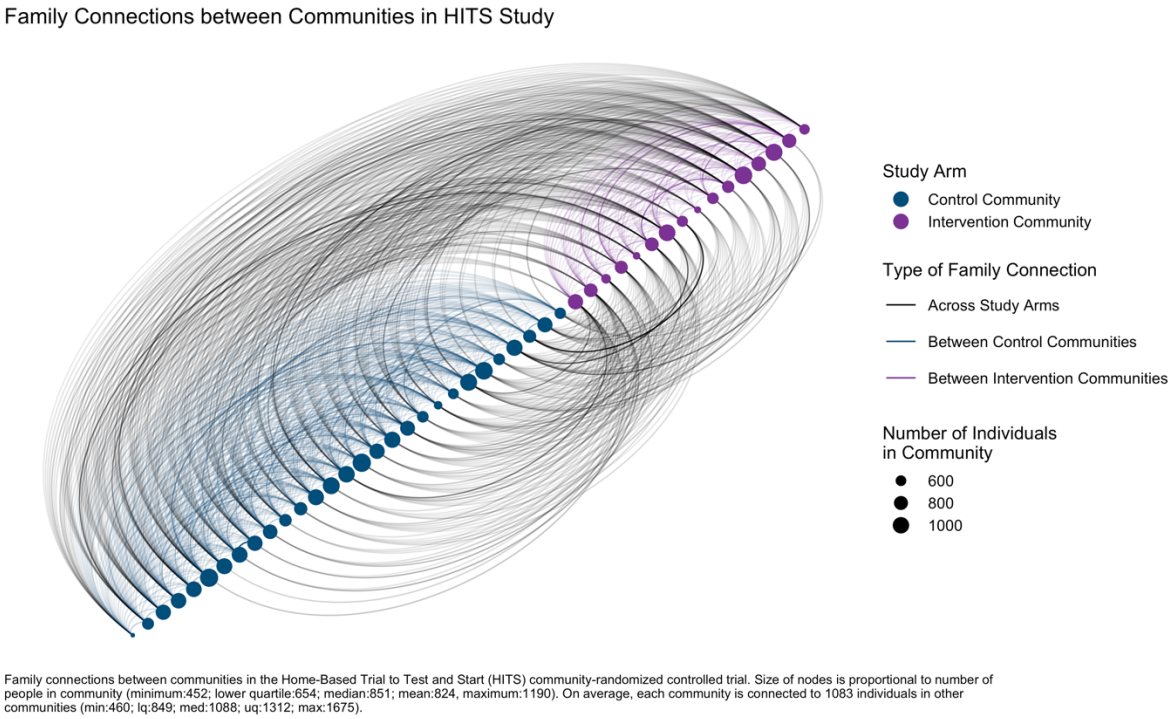

Figure 1.3: Family connections between communities

Family connections between communities in the Home-Based Trial to Test and Start (HITS) community-randomized controlled trial. Size of nodes is proportional to number of people in community (minimum:452; lower quartile:654; median:851; mean:824, maximum:1190). On average, each community is connected to 1083 individuals in other communities (min:460; lq:849; med:1088; uq:1312; max:1675).

The descriptive analysis shows that with the offer of a micro-incentive, the likelihood of consent to HIV testing increased substantially when participants had at least one family member who was in the intervention arm. In contrast, the offer was associated with a more modest effect on consent for HIV testing among individuals who had no family members who were in the treatment arm (Figure 1.4 A). Among people with no family members in the intervention arm the micro-incentive increased uptake of HIV testing by 13% (Risk Ratio: 1.13, 95% CI: 1.05-1.21). Among people with at least one family member in the treatment arm, the micro-incentive increased testing uptake by 46% (RR: 1.46, 95% CI: 1.3-1.64).

Table 1.1: Baseline Characteristics (based on whole-family network)

	Control Arm (N=10037)	Incentive Arm (N=5638)	Overall (N=15675)
Age			
15-25	3920 (39.1%)	2292 (40.7%)	6212 (39.6%)
26-35	1690 (16.8%)	922 (16.4%)	2612 (16.7%)
36-45	1194 (11.9%)	679 (12.0%)	1873 (11.9%)
46-55	1247 (12.4%)	695 (12.3%)	1942 (12.4%)
>55	1986 (19.8%)	1050 (18.6%)	3036 (19.4%)
Gender			
Female	6974 (69.5%)	3829 (67.9%)	10803 (68.9%)
Male	3063 (30.5%)	1809 (32.1%)	4872 (31.1%)
Ever Tested HIV+			
Yes	1796 (17.9%)	1029 (18.3%)	2825 (18.0%)
No	6344 (63.2%)	3716 (65.9%)	10060 (64.2%)
Refused	113 (1.1%)	61 (1.1%)	174 (1.1%)
Missing	1784 (17.8%)	832 (14.8%)	2616 (16.7%)
Family Network Size			
0	290 (2.9%)	134 (2.4%)	424 (2.7%)
1-5	5365 (53.5%)	3029 (53.7%)	8394 (53.6%)
6-10	3064 (30.5%)	1723 (30.6%)	4787 (30.5%)
11-15	961 (9.6%)	539 (9.6%)	1500 (9.6%)
16+	357 (3.6%)	213 (3.8%)	570 (3.6%)
Percentage of Family Members in Different Household			
0%	3990 (39.8%)	2217 (39.3%)	6207 (39.6%)
0-20%	820 (8.2%)	438 (7.8%)	1258 (8.0%)
20-40%	1484 (14.8%)	909 (16.1%)	2393 (15.3%)
40-60%	1390 (13.8%)	792 (14.0%)	2182 (13.9%)
60-80%	1343 (13.4%)	727 (12.9%)	2070 (13.2%)
80-100%	571 (5.7%)	313 (5.6%)	884 (5.6%)
100%	439 (4.4%)	242 (4.3%)	681 (4.3%)
Percentage of Family Members in Different Community			
0%	5831 (58.1%)	3231 (57.3%)	9062 (57.8%)
0-20%	1237 (12.3%)	655 (11.6%)	1892 (12.1%)
20-40%	1269 (12.6%)	759 (13.5%)	2028 (12.9%)
40-60%	803 (8.0%)	486 (8.6%)	1289 (8.2%)
60-80%	537 (5.4%)	320 (5.7%)	857 (5.5%)
80-100%	189 (1.9%)	88 (1.6%)	277 (1.8%)
100%	171 (1.7%)	99 (1.8%)	270 (1.7%)

Table 1.1 (continued). Baseline Characteristics (based on whole-family network)

Network Treatment (# Family Members in Incentive Arm and who have Prior Study Visit)			
0	9092 (90.6%)	839 (14.9%)	9931 (63.4%)
1	551 (5.5%)	1205 (21.4%)	1756 (11.2%)
2	175 (1.7%)	1088 (19.3%)	1263 (8.1%)
3	77 (0.8%)	810 (14.4%)	887 (5.7%)
4	50 (0.5%)	570 (10.1%)	620 (4.0%)
5+	92 (0.9%)	1126 (20.0%)	1218 (7.8%)

The descriptive analysis further provides some evidence of a dose-response relationship between the number of family members in the intervention arm and the strength of the effect of the micro-incentive on HIV testing (Figure 1.4 B). The effect size increased from 13% (RR: 1.13, 95% CI: 1.05-1.21) among people with no family members in the intervention arm to 62% (RR: 1.62, 95% CI: 1.15-2.27) among people who have 3 family members in the intervention arm, and appeared not to change substantially for people with 4 family members (RR: 1.60, 95% CI: 0.80-3.21) or five members (RR: 1.6, 95% CI: 0.90-2.86) in the intervention arm.

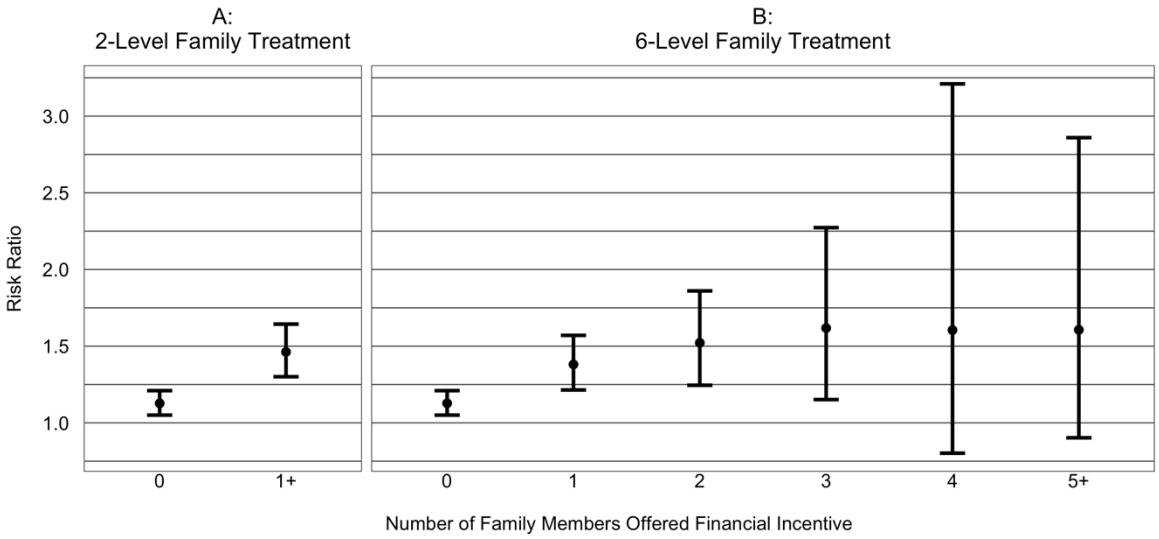

Figure 1.4: Effect Heterogeneity of HITS Intervention

Primary Analysis

The primary analysis shows evidence of a spillover effect. According to the GMM estimates, not being offered the incentive and having an additional family member offered the incentive did not lead to a significantly different probability of consenting for an HIV test compared with not being offered the incentive and having no additional family member offered the incentive (RR=0.90 95% CI: 0.77-1.06). On the other hand, being offered the incentive and having an additional family member offered the incentive led to an 8% increase in the probability of consenting for an HIV test (RR=1.08 95% CI: 1.00-1.16) compared to being offered the incentive and having no additional family member offered he incentive. To put this in context, for a person with two family members treated (the average number of family members), having those family members treated prior to the individual would increase her likelihood of testing uptake by 17%. The overall effect of the financial incentive (accounting for main and spillover effects) was about 20%.

Discussion

We found empirical evidence that being offered an HIV testing incentive not only affects the recipient's chances of taking up HIV testing, but it affects the recipient's family members' chances. Having an additional family member who is offered a financial micro-incentive leads to higher rates of uptake if the individual also receives the offer. If the individual does not receive an offer, there is no evidence that having an incentive offered to a family member affects the individual's HIV testing uptake. Our descriptive analysis suggests that there are diminishing returns for each additional family member who is offered the incentive after about 4 family members have received the offer. In summary, we observed spillover of the effect of offering a micro-incentive on HIV testing uptake, but only among individuals in the intervention arm.

There are a few possible mechanisms connecting the financial incentive received by one person and the HIV testing behavior of another. Family members might discuss HIV testing as a result of undergoing testing themselves. Individuals might pool the incentive and encourage each other to undergo testing in order to maximize this resource. Finally, the fact of being offered an incentive might cause family members to talk about HIV regardless of their own HIV testing behavior.

There have been numerous observational studies that investigated the impact of social networks on HIV-related health, largely supportive of the idea that social networks can promote health (M. Casale et al., 2015, 2019; Chanda et al., 2017; Conserve et al., 2019; Hermanstyne et al., 2018; Mulawa et al., 2016; Musheke et al., 2013; Nelson et al., 2015; Perkins et al., 2015; Takada et al., 2019; Yamanis et al., 2016). Notably, a recent pilot randomized trial assessed the effect on HIV treatment adherence of sending sms reminders to social contacts nominated by participants, finding no evidence of such an effect (Haberer et al., 2016). Ours is the first use of large-scale trial data to assess family network spillover of an HIV intervention.

Conducting a sociocentric analysis of a population-based community-randomized trial allowed us to test a central assumption associated with this study design — that communities are statistically independent of one another. We find that to the extent that they are connected by relationships that are relevant to the exposure and outcome under study, communities are not independent. Given that there are a number of surveillance platforms that collect data on kinship, this study demonstrates a promising approach for assessing family spillover effects in the context of other community-randomized trials nested in these platforms.

Our study offers several methodological contributions. We leverage the study design to produce randomization-based inferences for heterogenous causal effects by characterizing the

distribution of the number of family members who received the intervention — the network exposure of interest. Since the network exposure in our case is linear in the treatment status of individual units, we did this analytically. For more complicated network exposures, the same could be accomplished using simulation. By analyzing relationships that were established prior to the randomized intervention in question, we avoid the problem of disentangling homophily from causal spillover (Shalizi & Thomas, 2011). Further, we conducted this analysis making weak distributional assumptions about the data, allowing this approach to be applied in a wide variety of settings. Finally, we highlight a new usage of routinely collected data on kinship. Processing and analyzing these data is possible using software that is already widely used in epidemiology and data on kinship are readily accessible through several population surveillance platforms.

A limitation we faced is that although the financial incentive intervention was randomly assigned, the order in which households were visited was not. The order of visits affected the network treatment value assigned to each individual. It is difficult, however, to construct a situation where the ordering would bias effect estimates. It was not possible to assess differences in the strength of within- vs. between- household effects. It was also not possible to assess differences in within- vs. between-community effects. The community-randomized trial was not powered to assess these differences. Future trials should consider including both main effects as well as potential spillover effects in power calculations.

Our study suggests that financial incentives for HIV testing could have a multiplier effect across familial relationships (VanderWeele & Christakis, 2019), improving the efficiency of this intervention. The stark difference between individual- and population-level effects of ART on HIV transmission suggests that HIV transmission continues between people who are not yet virologically suppressed and their HIV-negative sexual partners (Baral et al., 2019). The

expansion of ART coverage coupled with stubbornly high incidence signals the progressive concentration of transmission events among groups who have poor access to HIV services (Ortblad et al., 2019). By adding an additional approach to reaching these groups, family-based financial incentive programs for HIV testing could help to close the gap between the fast rate of increase in HIV treatment coverage and the slow rate of decrease in new infections.

Chapter 2

Networked wealth and baseline physical function among HAALSI cohort members

Background

A demographic transition is underway in South Africa. The population of adults over 50 years of age is growing, spurring on age-related non-communicable diseases, including those that interact with widely prevalent HIV disease (Tollman et al., 2016). It is projected that between 2020 and 2030, the proportion of the population that is over 65 will grow by one fifth — from 5.5% in 2020 to 6.7% in 2030 (United Nations, Department of Economic and Social Affairs, Population Division, 2019). Depending on how the patterns of mortality and morbidity evolve along with this demographic transition, there could be profound changes in the independence and overall quality of life of older adults as well as their family members (J. M. Guralnik, 2004). An increased burden of disability would occasion an increase in the demand for formal and informal care.

Guided by the Disablement Process (Pope & Tarlov, 1991; Verbrugge & Jette, 1994), we investigate the relationship between socioeconomic status and three outcomes related to disability: Activities of Daily Living (ADL), grip strength, and gait speed. The Disability Process is a conceptual model that posits a causal pathway moving from *pathology* to *impairment* to *functional limitation* and finally to *disability*. Pathology refers to physiological abnormalities such as diagnosed or undiagnosed medical conditions; impairments are dysfunctions or structural abnormalities that happen as a result of pathologies; functional limitations are experienced difficulty or inability to perform basic actions needed in daily life such as climbing or walking; and disability is defined as a gap between the demands of one's social and physical environment, and one's functional capabilities. For example, if a person cannot walk but lives in an environment and has access to assistive technology such that the inability to walk does not impede his ability to accomplish tasks, then he is not disabled.

The ADL scale was designed in the 1960s to assess the severity of limitations among elderly and chronically ill patients in performing tasks that are “habitually and universally” performed: bathing, dressing, going to the toilet, transferring, continence, and feeding (Katz et al., 1963). It is now widely collected among community-dwelling older adults (Fieo et al., 2011). From the 1980s to date, there have been at least three criticisms of measures like ADL (Brach et al., 2002; Carp, 1977; J. M. Guralnik et al., 1989; Melzer et al., 2004; Miller et al., 2018; Minkler et al., 2006; Schoenmaeckers, 2013; Xu et al., 2019). The first is that ADL, commonly measured by self-report, is subjective; different groups of people systematically appraise the difficulty of performing certain tasks differently. The second is that ADL attempts to measure disability within individuals whereas disability, per the Disablement Process model and its precursors, is a concept relating an individual to her environment. Finally, measures like ADL only identify people once they have sustained significant losses in function. Individuals who sustain minor losses and adapt to them might not experience the type of physical limitation measured by ADL. Individuals with this “preclinical disability” (Fried et al., 1991), however, are important to identify in order to study the etiology of functional limitations and disability and to potentially mount interventions mitigating these.

In response to these criticisms, objective performance-based measures such as hand grip strength and gait speed have been developed and taken up in epidemiologic studies (J. M. Guralnik et al., 1989). They can detect relatively small changes in function and are not susceptible to the systematic biases associated with self-report (J. M. Guralnik et al., 1996). Though it has been argued that they present advantages over measures such as ADL (Elam et al., 1991; Rozzini et al., 1997; Sherman & Reuben, 1998), we take the view that they measure different constructs (Reuben et al., 1995; Sager et al., 1992; Savino et al., 2014). Hand grip strength can be thought of as a measure of impairment — it is a means of measuring overall losses in muscle and muscle strength (S. M. Patel et al., 2020). Gait speed can be considered a

functional limitation — it goes beyond impairment by assessing the ability to perform the action of walking. ADL can be considered a flawed measure of disability. It assesses the capability of individuals to accomplish basic tasks of self-care presumably demanded by their social environment but fails to do so in relation to their actual environments (J. M. Guralnik, 1997; Jette & Keysor, 2003; Verbrugge & Jette, 1994).

The Disablement Process does not represent a fixed causal process dictating a progression from loss of grip strength to slowing gait speed to limitations in the activities of daily living. It is possible for a person experiencing a certain disability to recover, sometimes with the aid of medical or behavioral intervention (J. M. Guralnik et al., 1996; Institute of Medicine (US) Committee on Assessing Rehabilitation Science and Engineering, 1997). “Predisposing characteristics” can shape the severity of pathology, impairment, functional limitation, and disability and can delay or accelerate transitions between these states (Verbrugge & Jette, 1994). Crucially, a particular impairment or functional loss can have discrete causes that have immediate deleterious effects or causes that gradually accumulate over the life course and act synergistically with others. We investigate socioeconomic status as an example of the latter.

Prior studies have established a robust positive association between higher socioeconomic status and performance-based measures of physical function (Borges et al., 2020; Demakakos et al., 2013; Mohd Hairi et al., 2010; Sanderson & Scherbov, 2014; Tampubolon, 2015) and a negative association with self-reported limitations in basic and instrumental activities of daily living (Andrade et al., 2018; d’Orsi et al., 2014; Giacomin et al., 2019; Gjonca et al., 2009; Gong et al., 2020; Qian & Ren, 2016; Serrano-Alarcón & Perelman, 2017; Tang et al., 2021; Torres et al., 2016; Wahrendorf et al., 2013; Zhong et al., 2017). There is evidence that lower childhood socioeconomic status causes lower physical function in later life (Birnie et al., 2011; Landös et al., 2019), and mixed evidence that these effects persist even after accounting for adult

socioeconomic status (Vable et al., 2019). For example, the development of the impairment of sarcopenia in later life is associated with low birth weight and other biological indicators in early life which are themselves shaped by the socioeconomic status of the individual and his mother (Dodds et al., 2012; Sayer et al., 2008). These findings are based largely on studies that were conducted in high income countries, though there is a growing body of evidence from low- and middle-income countries (Brennan-Olsen et al., 2019) including South Africa (Payne et al., 2017). These studies lend support to the idea that socioeconomic status arrays multiple exposures which act over the life course. Through various biological causal pathways that are set off by these exposures, individuals embody their socioeconomic status and manifest health outcomes including disability (Coppin et al., 2006; J. M. Guralnik et al., 1996; Krieger, 2001).

A gap in this literature is that although there is a wealth of evidence demonstrating that individuals' access to material resources shapes physical function and disability, there is very little evidence on the impact of family members' resources on these outcomes. This is an important gap. Evidence from Agincourt, South Africa, the geographical location of our study, shows that households are embedded in local kinship-based networks of support, exchanging labor, food, and other resources with one another (Madhavan et al., 2014). Because of this interdependence, the level of wealth held by family members could be an important determinant of the physical function of the individual. If true, this would be of public health significance. To understand and address socioeconomic health inequities in physical function and disability, it would be important to not only consider the socioeconomic status of aging individuals, but also the socioeconomic status of their family members.

In our study we examine the relationship between physical function as measured in 2014 with household wealth as measured in the 13 preceding years. We ask: on average, how much of a change in *average physical function* would be caused by a one-standard-deviation increase in

household wealth in a particular year? We further ask what proportion of this effect could be attributed to an *average direct effect* and what proportion could be attributed to an *average spillover effect*. i.e. we investigate how much of the effect of wealth on physical function, if it exists, is due to the household's own wealth, and how much is due to the wealth embedded in the family network. Physical functioning is measured using hand grip, gait speed, and limitations in activities of daily living. We examine the level of function rather than change in function even though change scores would automatically adjust for time-invariant confounding. Change scores are challenging to measure since peak physical function will likely have occurred prior to study enrollment for most participants. Past studies have produced robust evidence that physical function predicts later disability and mortality and mixed evidence that measured changes in physical function are as predictive (Gill et al., 1997; Hicks et al., 2012; Miller et al., 2018; Prasitsiriphon & Pothisiri, 2018; Thorpe et al., 2011; Xue et al., 2015).

To answer the above questions, we constructed a network of individuals in the Agincourt Health and Demographic Surveillance System (AHDSS) over the period 2001 - 2017. We restricted the network to individuals who are in the Health and Aging in Africa: A Longitudinal Study of an INDEPTH Community in South Africa (HAALSI) cohort or directly related to HAALSI cohort members. Separately, we regressed average grip strength, gait speed, and ADL on household wealth, network wealth, and covariates. We did this for each year that household wealth was measured. We establish a causal interpretation for these regression results, drawing on auto-g-computation, a recently proposed statistical approach to quantify spillover causal effects on a network of interconnected units (Tchetgen Tchetgen et al., 2020).

Methods

Thought Experiment

To define the causal estimands of interest, we conduct the following thought experiment. We have a network such that individuals are nodes and family connections are ties between the nodes. The ties between individuals are assumed to potentially convey the material resources held by the households that the individuals are members of. In other words, if individual a in Household A is connected to individual b in Household B, we investigate whether physical function for individual a is caused by the resources held by Household B after accounting for those held by Household A. We further assume that network ties possibly allow person b's physical function to affect person a's physical function, even after accounting for resources held by Household A and Household B. We restrict our attention to individuals who are in the HAALSI cohort or directly connected to HAALSI cohort members.

Given this network, we imagine measuring the average level of *grip strength*, *gait speed*, and *limitations in activities of daily living*. We then instantaneously increase each household's wealth by 1 standard deviation and measure average physical function using the same three variables, but this time under the new level of wealth.

For each outcome, the difference between the first and second measure of average physical function would be the Average Total Effect (ATE) — the full impact of increasing everybody's household wealth by 1 standard deviation. But since, by assumption, each individual potentially benefits from not only her own household's wealth, but also from the wealth of households she is connected to through family relationships, it is possible to decompose ATE into two quantities: the average effect of the individual's household wealth on her own physical function (Average Direct Effect — ADE), and the average effect of the wealth held by connected households on

her physical function (Average Spillover Effect — ASE). We calculate these effects separately for wealth as measured in 2001, 2003, 2005, 2007, 2009, 2011, and 2013, and for physical function (gait speed, grip strength, ADL) as measured in 2014.

We can think of ADE as capturing the average effect of increasing every household's wealth by 1 unit while not allowing family members' households to benefit directly from one's own household wealth.[5] ASE is the additional effect of wealth on physical function that results from allowing households to benefit from each other's wealth. ATE is the total effect of increasing household wealth by 1 unit, accounting for both each individual's household wealth, as well as the wealth of the households of family members. These quantities are more formally defined in Appendix B.

---

[5] Say we have a network of 3 individuals where Individuals 1 and 2 live in Household 1 and Individual 3 lives in Household 2. Observed wealth in these households is A_1 and A_2 respectively. We increase wealth in Household 1 (a_1) and wealth in Household 2 (a_2) by an infinitesimal amount.

Consider the contribution of Individual 1 to ATE, ADE, ASE. We label this contribution DE_1, SE_1 and TE_1.

By definition (see Appendix B):

$$TE_1 = \frac{1}{3}\left(\frac{\partial E[Y_1(\boldsymbol{a})]}{\partial a_1} + \frac{\partial E[Y_1(\boldsymbol{a})]}{\partial a_2}\right)\Bigg|_{\boldsymbol{a}=\boldsymbol{A}}$$

$$DE_1 = \frac{1}{3}\left(\frac{\partial E[Y_1(\boldsymbol{a})]}{\partial a_1}\right)\Bigg|_{\boldsymbol{a}=\boldsymbol{A}}$$

$$SE_1 = \frac{1}{3}\left(\frac{\partial E[Y_1(\boldsymbol{a})]}{\partial a_2}\right)\Bigg|_{\boldsymbol{a}=\boldsymbol{A}}$$

Though we intervene by raising wealth in both households, for DE_1 we only consider the change in Individual 1's mean potential outcome that is attributable to the change in Individual 1's household's wealth. For TE_1 we add to this first quantity the change in Individual 1's mean potential outcome that is attributable to the change in the income of the households that Individual 1 is not a member of, in this case, Household 2.

Measures

Setting

The Agincourt Health and Demographic Surveillance System (AHDSS) dataset contains data from Agincourt, South Africa, and is based on an annual census that has been conducted since 1992 (Kahn et al., 2012). In this platform, trained field workers visit all households in the surveillance area annually and interview a key resident informant each time. The survey records demographic information including the parents, co-residents, and conjugal partners of each individual in the household. These data uniquely identify all households and all individuals in the surveillance area. Starting in 2001, household wealth was measured among households once every two years (Kabudula, Houle, Collinson, Kahn, Tollman, et al., 2017; *MRC/Wits Agincourt Unit, Rural Public Health and Transitions Research Unit*, 2021).

The Health and Aging in Africa: A Longitudinal Study of an INDEPTH Community in South Africa (HAALSI) dataset was collected among a sample of AHDSS residents. It contains in-depth data on health outcomes collected from 5,059 people who were over 40 years of age at entry (Gómez-Olivé et al., 2018). Two waves of data were collected: the baseline wave was collected beginning in 2014, and the first follow-up wave beginning in 2018. We use physical functioning data collected during the baseline wave.

Family Network

A sociocentric family network was constructed among all individuals in the AHDSS who are members of the HAALSI cohort or connected to members of the cohort. In this network, nodes represent individuals and ties represent two kinds of family relationships: parent-child relationships, and conjugal (or formalized romantic) relationships. If a pair of nodes were ever

connected over the period 1993-2018, they are connected in this family network. In addition, a tie was added between each pair of nodes that was connected to a common node.

In each round of the annual census, each individual is associated with a uniquely identified household. Individuals can change households over time, so each node in the family network is associated with a repeated (annual) measure of *household*.

Household Wealth

Household wealth is measured using an asset index calculated according to DHS methodology. The measure combines information on household infrastructure and goods (Payne et al., 2017; Rutstein et al., 2004). Measurements of household wealth were made every two years from 2001 to 2013.

Network Wealth

Network wealth for a given individual in a given year is calculated as the sum of household wealth among households whose members include direct relatives of the individual, according to the sociocentric family network (See Figure 2.1).

Individual Physical Functioning

Individual grip strength was measured using a Smedley digital hand dynamometer, taking two measurements per hand. Following the analysis by Payne et al. (2017) we use the average of the grip strength measures on the participant's self-reported dominant hand. For participants who reported being ambidextrous, we take the average of the two highest measures, regardless of which hand they were measured on. Measures above 75kg were treated as out of range and therefore missing.

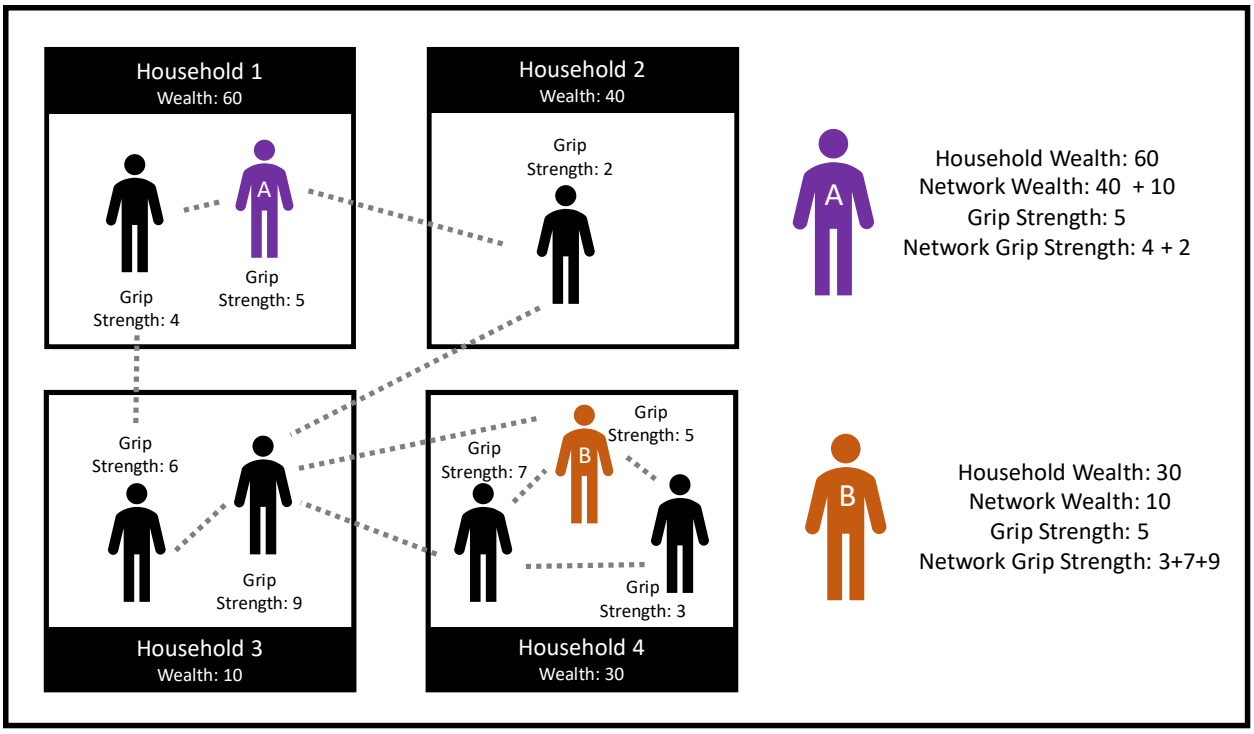

Figure 2.1: AHDSS Network Schematic

Dotted lines indicate family relationships. Boxes indicate households. Each individual has an individual Grip Strength score and each household has a Household Wealth value. Network Wealth is the sum of Household Wealth values for connected households. Network Grip Strength is the sum of Grip Strength scores among connected individuals.

Individual gait speed was measured using a timed walk. Interviewers marked a length of 2.5 meters on an obstacle-free floor. The respondent was asked to walk from one end to the other, and she was timed. The respondent was then asked to turn around and return to the point of origin while being timed. Gait speed was calculated by dividing 5 by the sum of the times (in seconds). Gait speeds below 0.2 m/s or above 2 m/s were treated as out of range and therefore missing.

Individual limitations in activities of daily living (*ADL*) is a measure of functional limitation measured through a set of questions asking if the respondent is unable (or finds it difficult) to bathe, eat, get out of bed, toilet, or walk across the room unaided. ADL is quantified as the

count of activities the individual is unable or finds difficult to accomplish. We dichotomized this variable (0 vs 1+).

Network physical Functioning

For each individual, we calculated *network grip strength*, *network gait speed*, and *network activities of daily living* as the sum of grip strength, gait speed and activities of daily living, respectively, among people she is connected to in the family network who are above 40 years of age. By so doing, we assume that under 40s are not able to acquire and transmit physical functioning. This assumption was necessary to make since physical functioning was only measured among people over the age of 40 — an inclusion criterion for the HAALSI cohort.

Statistical Analysis

Descriptive Analysis and Validity Checks

We tabulated demographic characteristics of HAALSI cohort members. We also examined the validity of the physical function outcomes by regressing death between HAALSI waves 1 and 2 on each of the outcomes. We expected Grip Strength and Gait Speed to be negatively associated with the risk of death and we expected ADL to be positively associated with the risk of death.

Statistical Model

We assume the outcome data arose from a conditional Markov Random Field (Besag, 1974; Tchetgen Tchetgen et al., 2017). This statistical model is defined by a local Markov assumption: one individual's outcome is independent of any other individual's outcome, treatment and covariate observations given the exposures and covariates of the first individual as well as the exposure, covariate, and outcomes of the first individual's family members.

In addition, we assume a conditional mean model relating each individual's physical function to their household wealth, network wealth, and network physical function:

$$E\left[Y_{i,2014}^{ind} \middle| \boldsymbol{L} = \boldsymbol{l}, \boldsymbol{A} = \boldsymbol{a}, \boldsymbol{Y} = \boldsymbol{y}\right] = \beta_0 + \beta_1 a_{it}^{hh} + \beta_2 l_i^{ind} + \beta_3 a_{it}^{net} + \phi y_i^{net}$$

where for individual i, $Y_{i,2014}^{ind}$ is individual physical function as measured in HAALSI, $y_{i,2014}^{net}$ is network physical function, a_{it}^{hh} is household wealth for individual i's household in year t, a_{it}^{net} is network wealth in year t, and l_i^{ind} is an individual-level covariate value. $t \in \{2001, 2003, 2005, \ldots, 2013\}$. We fitted this model separately for each year in which household and network wealth were measured (see Figure 2.2), and for each physical function outcome.

MODEL	EXPOSURES						OUTCOME
	Household & Network Wealth 2001	Household & Network Wealth 2003	Household & Network Wealth 2005		Household & Network Wealth 2013	Network Physical Function 2014	Individual Physical Function 2014
Model 1	■					■	■
Model 2		■		…		■	■
Model 3			■			■	■
	⋮				⋮		
Model 7					■	■	■

Figure 2.2: Schematic showing exposures and outcome of fitted conditional mean models for a given physical function outcome

Using β coefficients from these models along with information about network structure, for each year in which wealth was measured we calculated ADE, ASE, and ATE defined as follows.

(These β coefficients depend on the year in which wealth was measured, and the physical function outcome. We suppress time and outcome indices below):

First, we calculate an $n \times h$ Jacobian matrix $\boldsymbol{J}$:

$$\boldsymbol{J} = \frac{d}{d\boldsymbol{a}} E[\boldsymbol{Y}|\boldsymbol{L} = \boldsymbol{l}, \boldsymbol{A} = \boldsymbol{a}]$$
$$= (\boldsymbol{I} - \phi \boldsymbol{M}_{\boldsymbol{40+}})^{-1}(\beta_1 \boldsymbol{H} + \beta_3 \boldsymbol{HN})$$

where $\boldsymbol{M}_{\boldsymbol{40+}}$ is the adjacency matrix for the induced subnetwork of the family network, consisting only of individuals over 40 years of age. $\boldsymbol{H}$ is the household membership matrix defined as a $n \times h$ matrix with n rows representing individuals and h columns representing households. In this matrix, entry (i, j) is 1 if individual i is a member of household j. $\boldsymbol{N}$ is the adjacency matrix associated a network whose nodes are households, and ties represent family connections between households.

The causal estimands are defined as follows:

$$ATE = \frac{1}{n} \text{grandsum}\{\boldsymbol{J}\}$$
$$ADE = \frac{1}{n} \text{trace}\{\boldsymbol{H}'\boldsymbol{J}\}$$
$$ASE = ATE - \text{ADE}$$

We estimated these estimands separately for each year that household wealth was measured. In Appendix B we provide analytical expressions for the above estimands in terms of coefficients of the conditional mean model.

Estimation

Estimation of the causal estimands of interest proceeds in two steps. First we estimate the coefficients of the conditional mean model. Following that, we calculate the causal estimates: ATE, ADE, and ASE, which are a function of these coefficients.

Conditional Models

We fitted three separate models for each outcome — entering different sets of covariates each time. Model 1 has as predictors *household wealth* and *network wealth*. Model 2 includes these predictors as well as *network physical function*, and Model 3 includes the preceding predictors as well as age and gender.

Estimating the parameters of the conditional mean model is complicated by the fact that the observations belonging to a pair of individuals who are connected in the network are possibly correlated with each other. Proceeding as if the observations were independent may lead to biased estimates.

To account for correlation among observations, we use the coding estimator described in Tchetgen Tchetgen et al. (2017). Briefly, we find a stable set — a sub-network of the family network such that no two individuals in the sub-network are connected in the family network and such that each individual in the sub-network is a member of HAALSI (see Appendix B). Because of the local Markov assumption, this subset consists of conditionally independent observations, given their own treatments and covariates, and their family members' treatments, covariates, and outcomes. We then estimate parameters of the conditional mean model shown above using standard generalized estimating equations with robust standard errors (Zeileis, 2006).

The coding estimator can be inefficient if outcomes among connected units are truly uncorrelated after conditioning on their individual and network exposures and covariates — an assumption we test empirically.

Under a sub-model of the conditional mean model where $\phi = 0$, each individual's physical function outcome is mean-independent of other individuals' physical function outcomes. If there were no empirical evidence against such hypothesized submodel, it would be reasonable to assume that outcomes of directly connected units are independent, so that one can estimate the parameters of the mean model using a standard linear regression for independent outcomes, fitted using all available data. In this sub-model, we would not have to account for correlation between observations meaning that we would not need to use a stable-set.

Model Selection

We conducted a Wald hypothesis test (using robust standard errors), assessing whether $\phi = 0$. This was done by fitting the linear regression model implied by the conditional mean model while using all HAALSI data. This model is correct under the null hypothesis that $\phi = 0$ conditional on covariates. Where the null hypothesis was not rejected, we assumed that the data arise from the sub-model defined by conditional independence.

For conditional mean models, statistical inference was based on the asymptotic distribution of estimated coefficients. We constructed Wald confidence intervals for each of the coefficients of interest.

Causal Estimates

The causal estimates of interest (ADE, ASE, and ATE) are calculated as a deterministic function of the coefficients of the conditional mean models (shown above) and the structure of the family network.

We used the parametric bootstrap to conduct statistical inference for these quantities. We simulated 5000 realizations from the asymptotic joint normal distribution of the regression coefficients of the conditional mean model. We calculated ADE, ASE, and ATE each time, taking the mean of their empirical distributions as the point estimate and the 0.025 and 0.975^{th} quantiles as the bounds of the 95% confidence interval.

Potential Bias

ATE is only identified under the assumption that there is no unaccounted-for confounding. To identify ASE and ATE, it is necessary to further assume that there are no unmeasured causes of physical function which are correlated among family members. If the latter assumption had failed, physical function outcomes would be correlated among family members — a condition we test for (see sub-section on Model Selection). Finally, we assume that the functional form we chose for the conditional mean model is correct.

To account for potential confounding, we adjusted for age (40-50; 51-60; 61-70; >70) and gender (male; female) at HAALSI baseline since these potentially determine household wealth and physical function. Adjusting for age and gender only is consistent with past studies that examine the relationship between socioeconomic status and health over long time horizons (Birnie et al., 2011; Breeze et al., 1999; Chandola et al., 2007; Chandola, 2012; Frankel et al., 1999; Lawlor et al., 2004; Minkler et al., 2006; Mohd Hairi et al., 2010; Osler et al., 2009; Rautio

et al., 2005; G. D. Smith et al., 1997; Turrell et al., 2007). These studies tend to adjust for other variables only when attempting to assess their strength as mediators.

Though we had access to two waves of physical function data, we did not adjust for baseline physical function and assess the relationship between household wealth and physical function at the follow up wave. For individuals who were experiencing decline in physical function at baseline, adjusting for baseline physical function and examining follow-up physical function may induce collider stratification bias (Glymour et al., 2005).

Finally, we enter only one measure of socio-economic status instead of multiple measures since these measures tend to be tightly correlated. Following Link and Phelan (1995), we conceptualize household wealth as one of several resources that can be used to avoid or minimize the impact of illness. These resources include “money, knowledge, power, prestige, and other kinds of interpersonal resources embodied in the concepts of social support and social network” (Link & Phelan, 1995, p. 87).

Missing Data

We accounted for missing data by conducting a modified complete case analysis and separately by conducting multiple imputation using chained equations. Some individuals were missing data on *household wealth*, inducing missingness in *network wealth*. In addition, physical function was only measured among HAALSI participants and not among all AHDSS residents. This meant that among HAALSI participants, *network physical function* measures were right-censored if we summed over valid values of physical function. This kind of censoring possibly introduces bias when estimating parameters of the conditional mean model.

Modified Complete Case Analysis

In the modified complete case analysis, an individual was excluded from the analysis if she was missing the physical function outcome, or *household wealth*, or if all the households in which her family members live were missing *household wealth*, or if all connected individuals were missing the physical function outcome.

If at least one connected household had valid *household wealth*, then *network wealth* was calculated by summing the valid values of *household wealth* and upweighting the sum by the ratio of connected households to valid household wealth measurements.

i.e. we computed network wealth as:

$$\frac{\#\ \text{Connected Households}}{\#\ Valid\ HH\ Wealth\ Measures} \times sum\ of\ valid\ HH\ Wealth\ Measures$$

Similarly, if at least one connected individual had valid physical function, we calculated network physical function by summing the valid values of *physical function* and upweighting the sum by the ratio of # connected individuals to # valid physical function values

i.e. we computed network physical function as:

$$\frac{\#\ \text{Family Members}}{\#\ Valid\ Phys\ Func\ Meausres} \times sum\ of\ valid\ Phys\ Func\ Measures$$

Multiple Imputation using Chained Equations (MICE)

We imputed missing values of *household wealth* and *physical function* using a custom R package based on the multiple imputation using chained equations (MICE) algorithm (Buuren & Groothuis-Oudshoorn, 2011).

In each iteration of this algorithm, we imputed missing *household wealth* values using household-level measures as predictors and calculated *network wealth* values using these. We then imputed individual level *physical function* using individual-level measures in addition to household-level measures as predictors. We calculated *network physical function* measures using these. We iterated these steps 60 times for each imputed dataset, creating 64 such datasets.

Dichotomous variables were imputed using a random draw from the Bernoulli distribution with probability of success given by predicted values from a logistic regression. Continuous variables were imputed using predictive mean matching (See Appendix B). The imputation model only produces valid inferences under the assumption that data were "missing at random" (Rubin, 1976) and that the functional forms we specified for predictive mean matching and logistic regression are correct.

Results

Descriptive Analysis

At entry into the HAALSI cohort, a quarter of respondents were over 70 years of age and half were over 60 years of age (Table 2.1). Slightly more than half of the respondents were women. Over half had some formal education and a small minority were employed. About a third of participants received pension income. About half were married and/or living with a romantic

partner, and the vast majority had children. Respondents whose household wealth was above the median tended to have higher formal education and employment, were more likely to receive pension income and were more likely to have children. They also tended to have higher wealth embedded in their family network. Wealthier members of the HAALSI cohort had higher *grip strength*, lower *gait speed*, and fewer *limitations in activities of daily living* than poorer members.

Within households, household and network wealth was strongly correlated over time. Measurements of wealth that were closer in time, were more strongly correlated than those which were further apart (See Appendix B). All three measures of physical function predicted mortality as expected. HAALSI cohort members who died between the baseline and first follow-up waves of HAALSI had lower grip strength, lower gait speed, and more limitations in activities of daily living. Household wealth also predicted mortality as expected. Those with below median wealth were more likely to die than those with above median wealth.

Conditional Mean Model

In Figure 2.3, Figure 2.4, and Figure 2.5, we show results from the conditional mean models for limitations in *activities of daily living*, *grip strength*, and *gait speed*. In each figure, the top row shows results from the modified complete-case analysis described above and the bottom row shows results from the multiple imputation analysis. The first column shows the point estimate and confidence interval for *household wealth* (β_1 in the conditional mean model above), the second column shows results for *network wealth* (β_3), and the final column shows the coefficient for the network value of *physical function* (ϕ).

Table 2.1: Baseline Characteristics of HAALSI Cohort

	Above Median Wealth (N=2374)	Below Median Wealth (N=2378)	Overall (N=5048)
Age			
40-49	463 (19.5%)	491 (20.6%)	1040 (20.6%)
50-59	637 (26.8%)	656 (27.6%)	1391 (27.6%)
60-69	680 (28.6%)	529 (22.2%)	1260 (25.0%)
70+	593 (25.0%)	702 (29.5%)	1356 (26.9%)
Missing	1 (0.0%)	0 (0%)	1 (0.0%)
Gender			
Male	1098 (46.3%)	1084 (45.6%)	2336 (46.3%)
Female	1276 (53.7%)	1294 (54.4%)	2712 (53.7%)
Any Formal Education			
Yes	1533 (64.6%)	1019 (42.9%)	2728 (54.0%)
No	832 (35.0%)	1353 (56.9%)	2303 (45.6%)
Missing	9 (0.4%)	6 (0.3%)	17 (0.3%)
Employment			
Yes	403 (17.0%)	314 (13.2%)	801 (15.9%)
No	1964 (82.7%)	2058 (86.5%)	4233 (83.9%)
Missing	7 (0.3%)	6 (0.3%)	14 (0.3%)
Receives Pension Income			
Yes	859 (36.2%)	826 (34.7%)	1762 (34.9%)
No	1515 (63.8%)	1552 (65.3%)	3286 (65.1%)
Has Children			
Yes	2283 (96.2%)	2183 (91.8%)	4733 (93.8%)
No	89 (3.7%)	194 (8.2%)	311 (6.2%)
Missing	2 (0.1%)	1 (0.0%)	4 (0.1%)
Married			
Yes	1438 (60.6%)	1014 (42.6%)	2569 (50.9%)
No	933 (39.3%)	1363 (57.3%)	2475 (49.0%)
Missing	3 (0.1%)	1 (0.0%)	4 (0.1%)
Grip Strength			
Mean (SD)	24.3 (8.74)	23.2 (8.64)	23.8 (8.73)
Median [Min, Max]	23.1 [0, 73.6]	22.3 [0, 74.6]	22.7 [0, 74.6]
Missing	218 (9.2%)	230 (9.7%)	486 (9.6%)
Gait Speed			
Mean (SD)	0.672 (0.243)	0.689 (0.275)	0.681 (0.260)
Median [Min, Max]	0.625 [0.200, 1.67]	0.625 [0.200, 1.67]	0.625 [0.200, 1.67]
Missing	178 (7.5%)	193 (8.1%)	410 (8.1%)
Any ADL			
Yes	176 (7.4%)	237 (10.0%)	445 (8.8%)
No	2191 (92.3%)	2141 (90.0%)	4595 (91.0%)
Missing	7 (0.3%)	0 (0%)	8 (0.2%)
Network Wealth			
Mean (SD)	12.2 (6.26)	10.1 (6.79)	11.0 (6.69)
Median [Min, Max]	11.4 [0, 41.0]	9.96 [0, 42.2]	10.6 [0, 43.8]
Missing	720 (30.3%)	603 (25.4%)	1424 (28.2%)

There was no evidence of conditional network dependence for limitations in *activities of daily living* and *grip strength*. i.e. Since we did not reject the hypothesis that $\phi = 0$, we concluded that knowing about one individual's *ADL* or *grip strength* does not provide any information about the *ADL* or grip strength of individuals directly connected to them after adjusting for individual and network exposures and covariates. As a result, we fitted these models on the whole data set.

By contrast, there was some evidence for positive conditional network dependence of *gait speed* among directly connected individuals. As Figure 2.5 shows, for the complete case analysis estimated $\phi > 0$ meaning that if one person's *gait speed* is high, it is likely that their directly connected family members will also have high *gait speed*. We used the coding estimator to fit this model.

For all three physical function outcomes as measured at HAALSI baseline in 2014, and for each of the preceding years in which household wealth was measured, we failed to reject the null hypothesis that $\beta_3 = 0$ at type 1 error level of 0.05. We conclude that there is no significant statistical evidence to reject the hypothesis that wealth of connected households is not associated with each individual's physical function outcome after accounting for that individual's own wealth.

Finally, we found weak evidence of a negative association between *household wealth* and *ADL*, and weak evidence of a negative association between *household wealth* and *Gait Speed*. That is, greater wealth predicted slower gait speed. In both cases, imputation results showed a stronger negative association than complete case analysis results. We found relatively strong evidence of a positive association between *household wealth* and *grip strength*, both in complete case and multiple imputation results.

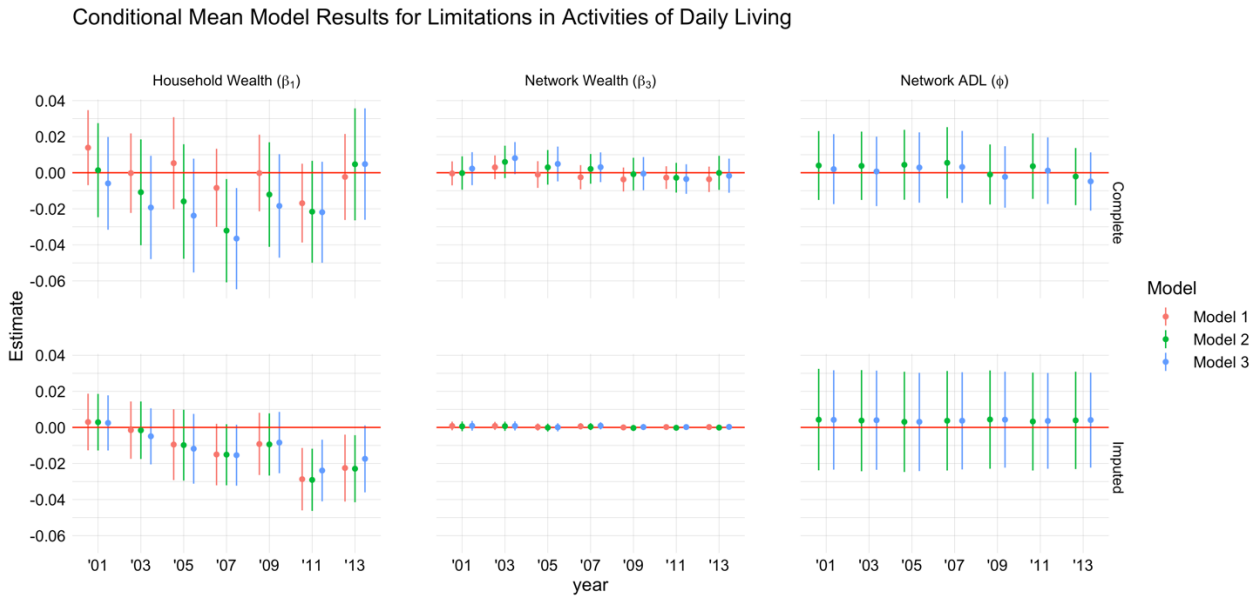

Figure 2.3: Conditional Mean Model Results for Activities of Daily Living Limitations (ADL)

Model 1 includes Network Wealth and Household Wealth as predictors. Model 2 includes these as well as Network ADL as predictors. Model 3 includes these as predictors and adjusts age and gender as potential confounders.

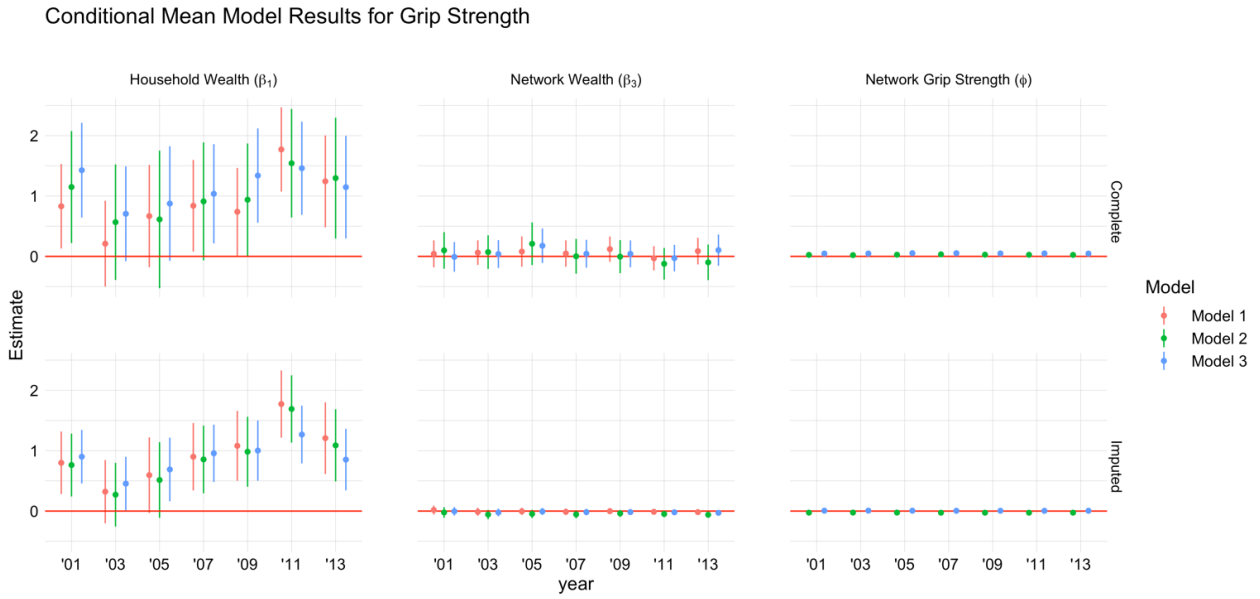

Figure 2.4: Conditional Mean Model Results for Grip Strength

Model 1 includes Network Wealth and Household Wealth as predictors. Model 2 includes these as well as Network Grip Strength as predictors. Model 3 includes these as predictors and adjusts age and gender as potential confounders.

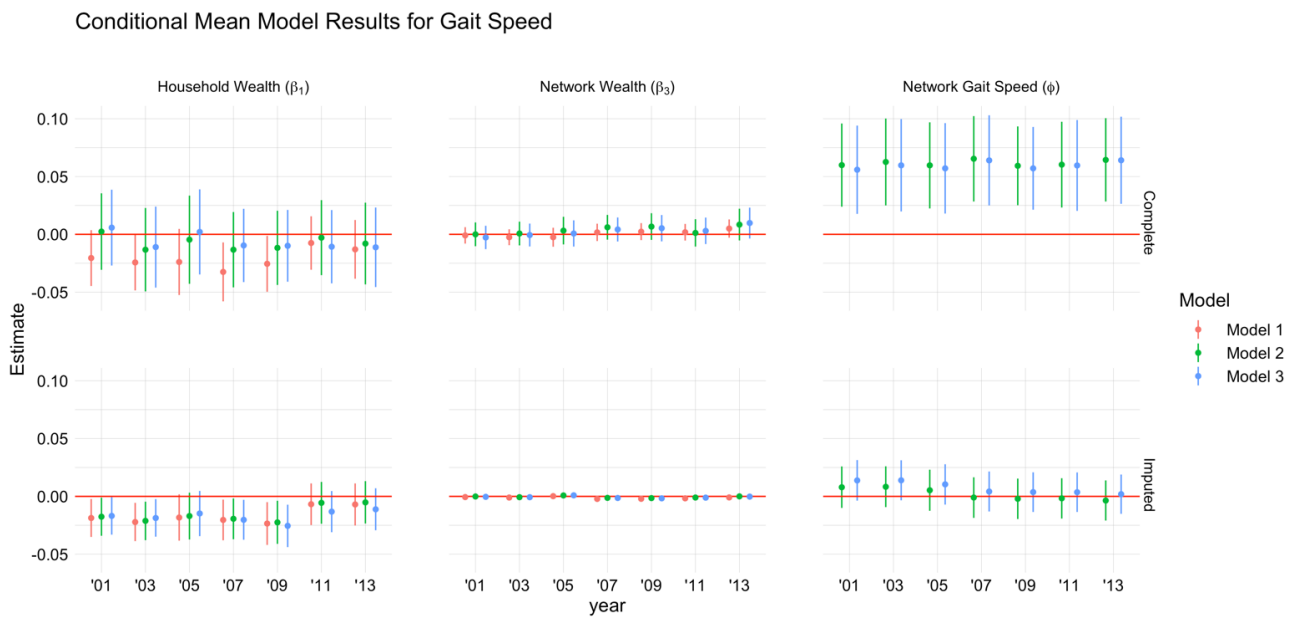

Figure 2.5: Conditional Mean Model Results for Gait Speed

Model 1 includes Network Wealth and Household Wealth as predictors. Model 2 includes these as well as Network Gait Speed as predictors. Model 3 includes these as predictors and adjusts age and gender as potential confounders.

Causal Estimates

Since for *grip strength* and for *ADL*, both the network wealth and network physical function variables were not associated with the outcome, there can be no network spillover of household wealth on these outcomes. (This follows from the fact that if $\phi = \beta_3 = 0$ then $ADE = ATE = \beta_1$).

Figure 2.6 displays the causal estimates for the effect of household wealth on *gait speed*. In the complete case analysis, there is no statistical evidence of a direct or spillover effect of household wealth on gait speed. In the multiple imputation analysis, there is some statistical evidence of a negative direct effect of *household wealth* on *gait speed* in 2001 (-0.02, 95% CI: -0.04:-0.01), 2003 (-0.02, 95% CI: -0.04:-0.01), 2007 (-0.02, 95% CI: -0.04:-0.01), and 2009 (-0.03, 95% CI: -0.05:-0.01), but no statistical evidence of a spillover effect. These would likely not remain statistically significant if we accounted for multiple testing.

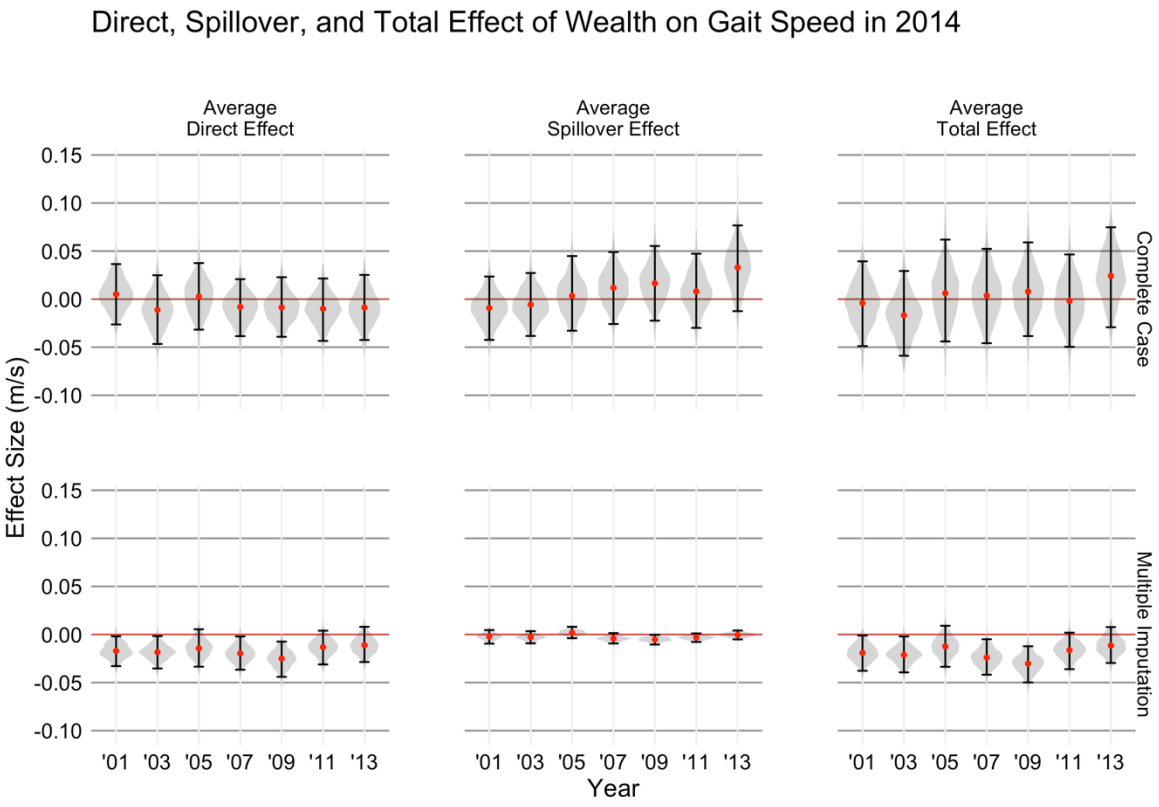

Figure 2.6: Causal Estimates of the Effect of Wealth on Gait Speed

Discussion

Overall, we found no evidence of spillover effects of household wealth on measures of objective physical function or limitations in activities of daily living. This was true for the modified complete case analysis and the multiple imputation analysis. We found evidence of a direct negative effect of household wealth on gait speed and a direct positive effect of household wealth on grip strength. Finally, we found weak evidence of a direct negative effect of wealth on limitations in activities of daily living.

For grip strength and activities of daily living, our results are broadly consistent with the extant literature on functional limitations and disability among older adults. Socio economic status has

consistently found to be positively correlated with performance-based measures of physical function (Borges et al., 2020; Demakakos et al., 2013; Mohd Hairi et al., 2010; Sanderson & Scherbov, 2014; Tampubolon, 2015) and negatively associated with variables based on self-reported limitations in basic and instrumental activities of daily living (Andrade et al., 2018; d'Orsi et al., 2014; Giacomin et al., 2019; Gjonca et al., 2009; Gong et al., 2020; Qian & Ren, 2016; Serrano-Alarcón & Perelman, 2017; Tang et al., 2021; Torres et al., 2016; Wahrendorf et al., 2013; Zhong et al., 2017).

Our finding that household wealth is negatively associated with gait speed is inconsistent with similar studies globally, but consistent with earlier work using the HAALSI dataset. Payne et al. (2017) found that those in the highest quintile of household wealth had lower gait speed than those in the lowest quintile. We found one other study, conducted in Boston, that found in a crude analysis that grip strength was negatively associated with socioeconomic status (Rios et al., 2001). Participants in this study were not randomly sampled and the study used the affluence of two residential neighborhoods as a proxy for socioeconomic status, however, so it is not methodologically comparable to ours. Our null finding echoes a result from a study that is comparable. Using the China Health and Retirement Longitudinal Study data which, like HAALSI, is a Health and Retirement Study (HRS) sister study, Tang et al. (2021) found that "family economic support" — financial support by parents, children, or siblings — was not associated with health outcomes including ADL among older adults. Family economic support was not quantified as in our study; it appears to have been a dichotomous variable indicating whether there was some family support or none.

In Agincourt, people of lower socio-economic status do more walking than those of higher status, possibly accounting for their faster walk speed. Figure 2.7, for instance, shows that HAALSI respondents in the lowest quartile of household wealth walked for 2.3 hours (95%CI:

2.14 - 2.47) per day to move from place to place while those in the second lowest quartile walked for 2 hours (1.87 - 2.09) and those in the highest quartile walked for 2.1 hours (1.97 - 2.32). There is robust evidence suggesting that physical activity plays a role in the prevention of limitations in physical function (S.-Y. Lee et al., 2018; Oliveira et al., 2020). Cross-country comparisons show that not only are population levels of impairment, limitation, and disability variable across settings (Payne et al., 2017; Tavares Milhem Ygnatios et al., 2021; Wahrendorf et al., 2013; Xu et al., 2019), the quantitative relationships among socioeconomic status, physical function, and self-reported limitations are context dependent as well (Chan et al., 2012; J. Guralnik et al., 2020; Miller et al., 2018; Peterson et al., 2017).

Alternatively, this negative association could be a result of collider stratification bias. Higher wealth likely increased the probability of surviving until enrollment in HAALSI through mechanisms other than gait speed, and higher gait speed might have also been positively associated with the probability of being in the HAALSI cohort. This kind of selection bias would induce a negative association between gait speed and wealth. A robustness check (shown in Appendix B) using wave 2 HAALSI physical function measures (as opposed to wave 1, which was used here), and using inverse probability of censoring weights to account for attrition between wave 1 and wave 2, shows a positive non-significant association between household wealth and gait speed. The latter analysis does not account for selective attrition between 2001 and 2014, however, so it does not completely address the potential for selection bias.

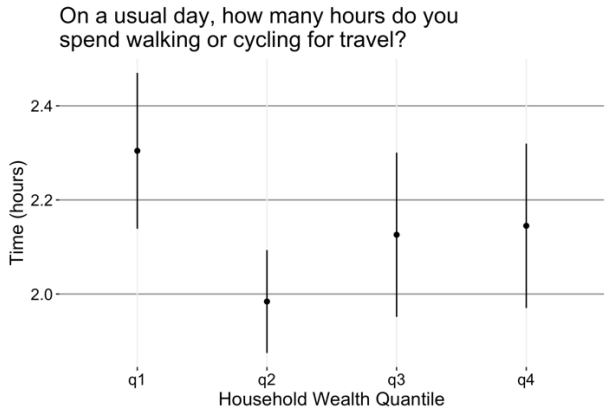

Figure 2.7: Time spent walking by Household Wealth

It is possible that there are spillover effects of wealth on function and disability that we failed to detect in this study. Our earliest measurements of wealth were in 2001 — only 13 years prior to the measurement of the health outcomes. For the youngest members of HAALSI, wealth was measured from their early 30s. Fort the oldest quarter of members of HAALSI, wealth was measured from their late 50s. These measures of wealth might miss a critical window during which shared resources are more important, etiologically. There is robust evidence that early life conditions shape mid-life physical function. Birth weight, pre-pubertal height gain, pubertal growth, infant motor development all predict mid-life grip strength (Sayer et al., 2008).

Our measure of socioeconomic status might be relatively insensitive to social gradients as they affect physical function, and therefore insensitive to networked social gradients. A prior study using HAALSI data showed steeper gradients in health outcomes when using a consumption-based measure of socioeconomic status rather than the wealth-based measure we used here (Riumallo-Herl et al., 2019). Unfortunately, this consumption-based measure was not available for non-HAALSI households in the AHDSS, making its use in this study extremely difficult. Furthermore, physical function and ADL might have been measured with some error, leading to

attenuated associations with household and network wealth and flattening measured social gradients.

Finally, it is not possible in our study to rule out reverse causation: it might be the case that prior ADL limitations, low gait speed and weak hand grip lowered the ability of HAALSI participants to generate income for their households due to unemployment (Ahrenfeldt & Möller, 2021), leading to lower wealth. It might also be that these effects were transmitted across family ties. Agincourt, however, has a chronically high unemployment rate, ranging from 63% as measured in the 2001 census to 52% as measured in the 2011 census (Statistics South Africa, 2021b). As a result, the share of household income that is comprised of government grants, including the disability grant, is high. In 2019, Statistics South Africa reported this figure to be 54% across Mpumalanga Province, with a further 22% of income gained from remittances (Statistics South Africa, 2021a). It is not clear whether functional limitations would decrease income via wages or increase income via grants and remittances, so it is difficult to anticipate the existence and direction of bias due to reverse causation.

We found clear evidence for direct effects of household wealth on physical function. This lends credence to our finding that there are no spillover effects. Even if these direct effects were an artifact of the relationship between early childhood socioeconomic status and current physical function (i.e. if direct effects presented here are an artifact of correlation between wealth at early age, and wealth in adulthood), we would expect to detect spillover effects at childhood (if they existed) using our measure of networked wealth. We found that networked wealth was more strongly correlated over time than household wealth. A further strength of the study is that we conducted this analysis making weak distributional assumptions about the data, allowing this approach to be applied in a wide variety of settings.

Our study provides evidence that physical function and disability are shaped by household wealth but not the wealth of family members in other households. Further research is needed to assess the effect of wealth and family wealth in early life, and if those effects are significant, to assess the different causal pathways that connect exposure and outcome. There is a need to revisit the relative usefulness of measures of consumption vs. measures of wealth in conducting these analyses. To enable this work, it is crucial to cultivate relational datasets, possibly combining data from epidemiologic studies with data that are passively collected from social media and mobile devices.

Chapter 3

Networked wealth and mortality in the Agincourt Health and Demographic Surveillance System 2009 – 2018

Background

There is indirect evidence that health outcomes are shaped not only by the resources held by individuals, but also those held by their social contacts. Much of this evidence has been generated through the social capital and health and social networks and health literatures. Despite the central importance of material resources as a mechanism through which social capital and social networks might affect health, however; evidence thus far tends not to measure resources directly. We address that gap in this study.

Prior research in South Africa has found that social capital is associated with better oral health (Olutola & Ayo-Yusuf, 2012), improved HIV treatment outcomes (Mukoswa et al., 2017), lower risk of tuberculosis infection (Cramm et al., 2011), higher self-rated health and subjective well-being (Chola & Alaba, 2013; Christian et al., 2020; Cramm et al., 2012; Cramm & Nieboer, 2011; Lau & Ataguba, 2015; Ramlagan et al., 2013) and better mental health (Adjaye-Gbewonyo et al., 2019; Tomita & Burns, 2013). Strongly influenced by political scientist Robert Putnam (R. Putnam, 1995; R. D. Putnam, 2000), these studies take what has been termed the *cohesion-based* approach to the measurement of social capital (Kawachi & Berkman, 2014). By assessing study participants' perceptions and behaviors, they measure aspects of social structure that are thought to be conducive to the diffusion of resources across social networks. They do not quantify the amount of resources available to diffuse, however. Similarly, research on social networks in South Africa has found positive associations between health outcomes and social support including financial and instrumental support (Harling et al., 2020) without measuring the quantity of support, but instead measuring the frequency.

In societies with extremely unequal distributions of wealth and income, such as South Africa, this is a crucial omission: if social capital and social networks confer benefits by channeling

resources, the failure to measure these resources might lead to spurious conclusions. For example, the stokvel is a widely prevalent voluntary association where members pay a fixed, regular contribution which is used for some agreed-upon purpose (Verhoef, 2001). The purpose could be, for instance, to pay out an agreed sum upon the death of a stokvel member's relative. Say we had two individuals who joined one community association each. Under the cohesiveness approach, we might conclude that they both have the same amount of social capital since they are both in the same number of community associations. If one had joined a stokvel, though, and the other had joined a choir because they could not afford to join a stokvel, they would have access to different levels of resources from their social networks. One would be a participant in an informal insurance scheme and the other would not be. This difference might be important to consider when investigating the relationship between social capital and health. Crucially, assessing these individuals' social capital using the coherence-based approach would obscure the fact that in this example, they not only have different levels of social capital, but that social capital serves to deepen the inequality between them (Makiwane et al., 2017).

In this study, we examine the relationship between mortality in Agincourt, South Africa, and the level of wealth held in individuals' family networks. Agincourt is a relatively poor area. According to the 2011 national census, the Bushbuckridge municipality, which contains Agincourt, had an average household income of ZAR 37 000. For context, the poorest province in South Africa, Limpopo, had an average income of ZAR 57 000 and the wealthiest province had an average income of ZAR 156 000 (Leholha, 2012a, 2012b).

By attempting to assess the level of wealth in the immediate family network, we take the network-based approach to social capital (Kawachi & Berkman, 2014). We ask: on average, how much of a change in mortality would an increase in *household wealth* cause? We quantify

how much of that effect would be due to the household's own increase in wealth, and how much would be due to the increase in wealth they have access to through their family network across households.

It has been well established that mortality exhibits a social gradient in Agincourt: people with higher socioeconomic status live longer lives (Kabudula, Houle, Collinson, Kahn, Gómez-Olivé, Tollman, et al., 2017). What has not been studied is the extent to which the socioeconomic status of family members affects mortality. We do so by constructing a sociocentric network of households in the Agincourt Health and Demographic Surveillance System (AHDSS) over the period 2009 - 2018. In this network, nodes are households and ties are family connections between households. Using regression analysis we assess the relationship between death in one household and wealth among households connected by family ties. We establish a causal interpretation for these regression results, drawing on auto-g-computation, a recently developed approach to evaluate network causal effects (Tchetgen Tchetgen et al., 2017).

Methods

Thought Experiment

To define the causal estimands of interest, we conduct the following thought experiment: In Agincourt, South Africa, we have a network such that households are nodes and family connections are ties between the nodes. The ties between households convey material resources between them. In other words, if Household A is connected to Households B and C, deaths in Household A are a function of Household A's wealth but also, we hypothesize, the wealth held by Households B and C.

Given this network, we measure the proportion of households that experience death in a given year. By “experience death”, we mean that at least one member of the household dies. We then increase each household’s wealth in the preceding year by 1 unit and measure the proportion of households that experience death under this new level of wealth.

The difference between the first and second proportion of households that experience death is defined as the Average Total Effect (ATE). Since, by assumption, each household potentially benefits from not only its own level of wealth, but from the wealth of households it is connected to, it is possible to decompose ATE into two quantities: the effect of the household’s own wealth on death within the household (Average Direct Effect — ADE), and the effect of the wealth of connected households on household death (Average Spillover Effect — ASE).

We can think of ADE as the average effect of increasing household wealth by 1 unit when all the ties between households are severed.[6] ASE is the additional effect of wealth on household death that results from re-connecting the severed ties. ATE is the total effect of increasing household wealth by 1 unit, accounting for both each household’s own wealth, as well as the wealth of connected households. These quantities are more formally defined in the appendix.

Data and Measures

The Agincourt Health and Demographic Surveillance System (AHDSS) platform contains data from Agincourt, South Africa, and is based on an annual census that has been conducted since 1992 (Kahn et al., 2012). In this platform, trained field workers visit all households in the surveillance area annually and interview a key adult resident informant each time. The survey

[6] This is not true in general. It is true in this example since household deaths are conditionally independent across households, as we argue below.

records demographic information including the parents, co-residents, and conjugal partners of each individual in the household.

Longitudinal Network

A longitudinal sociocentric network was constructed among all households in the AHDSS. To do this, we first constructed a sociocentric family network among residents. In this latter network, nodes represent individuals, and ties represent two kinds of family relationships: parent-child relationships, and conjugal (or formalized romantic) relationships. If a pair of nodes were ever connected over the period 1993-2018, they were connected in this family network. In addition, a tie was added between each pair of nodes that was connected to a common node. For example, a tie would be added between a grandmother and her grandchild (who are both connected through their child/mother), and a tie would be added between a brother and a sister (who are both connected through their mother).

In each round of the annual census, each individual is associated with a uniquely identified household. Individuals can change households over time, so each node in the family network is associated with a repeated measure of *household*.

For each year of the census, a network was created where nodes are households and ties are relationships between individuals who live in those households. i.e. two households (say Household A and Household B) are connected in a particular year if during that year, there was at least one pair of individuals a and b, such that (1) individual a was a member of Household A, (2) individual b was a member of Household B, and (3) individuals a and b are connected in the sociocentric family network described above. A household network was created for each year between 2001 and 2018. Together, these comprise a discrete time longitudinal network over this period.

Household Death

In the AHDSS data, records for people who are deceased are associated with a date of death. Using this field, we identified which individuals died in each year from 2009 and 2018 and assigned those deaths to the household the individual lived in during the year of their death. A household is said to have experienced death in a particular year, if at least one member of the household died during that year.

We constructed four different measures of household death: overall household death, household child death (under age 5), household adult death (ages 15-59), and household elder death (ages 60+). Overall household deaths is 1 if at least one member of the household died, and zero otherwise. Household child death, premature death, and elder death obtained when at least one child, one adult, or one elder, respectively, died. We chose these age ranges to conform with recent mortality studies based on AHDSS data (Byass et al., 2017; D'Ambruoso et al., 2016; Deribew et al., 2016; Kabudula, Houle, Collinson, Kahn, Gómez-Olivé, Clark, et al., 2017; Kabudula, Houle, Collinson, Kahn, Gómez-Olivé, Tollman, et al., 2017; Mee et al., 2016).

Household Wealth

Household wealth is measured by an asset index calculated using DHS methodology which combines information on household infrastructure and goods (Payne et al., 2017; Rutstein et al., 2004). Measurements of household wealth were made every two years from 2001 to 2013, and annually thereafter. For the years between measurements the last value of household wealth was carried forward. As a result, the household wealth measures for 2010 were carried forward from 2009, and the measures from 2012 were carried forward from 2011.

Network Deaths

Network deaths for a given household in a given year is calculated as the number of connected households that experienced death. *Network child deaths*, *network adult deaths*, and *network elder deaths* are calculated as the number of connected households that experienced household child deaths, household adult deaths, and household elder deaths, respectively.

Network Wealth

Network wealth for a given household in a particular year was calculated as the sum of household wealth among connected households.

Statistical Analysis

Descriptive Analysis

We tabulated demographic characteristics of households and computed intra-household correlation in wealth over time. To assess network auto-correlation in the exposure and outcome, we computed the Moran's I statistic for networks (Y. Lee & Ogburn, 2019). This statistic ranges from -1 to 1 and measures the extent of auto-correlation over network ties. A Moran's I statistic value of 1 represents perfect correlation between measures taken on adjacent nodes, and a value of -1 represents perfect dispersion.

Model Selection

We assume our network outcome data arise from a conditional Markov Random Field (Besag, 1974; Tchetgen Tchetgen et al., 2017). We first tested whether the data are compatible with the special case where each household's outcome is independent of the outcomes of connected households, conditional on covariates.

This was implemented by fitting the following conditional mean model using ordinary least squares using all available data:

$$E\left[Y_{i,t+1}^{hh} \middle| \boldsymbol{L}_t = \boldsymbol{l}_t, \boldsymbol{A}_t = \boldsymbol{a}_t, \boldsymbol{Y}_{t+1}\right] = \beta_0 + \beta_1 a_{i,t}^{hh} + \beta_2 l_{i,t}^{hh} + \beta_3 a_{i,t}^{net} + \beta_4 l_{i,t}^{net} + \beta_5 a_{i,t}^{hh} l_{i,t}^{hh} + \beta_6 l_{i,t}^{hh} a_{i,t}^{net} + \beta_7 l_{i,t}^{hh} l_{i,t}^{net} + \phi Y_{i,t+1}^{net}$$

where for household i in year t, $Y_{i,t}^{hh}$ is household death, $Y_{i,t}^{net}$is network death, $a_{i,t}^{hh}$ is household wealth, $a_{i,t}^{net}$ is network wealth, $l_{i,t}^{hh}$ is a household covariate value, and $l_{i,t}^{net}$ is the network covariate value. Network values are calculated by summing over household values associated with households connected to i. The time index (year) is represented by t.

We tested the null hypothesis that $\phi = 0$. Failing to reject the null hypothesis would allow us to proceed with regression analysis under the usual assumption that outcomes are conditionally independent. Rejecting the null, however, would necessitate a more complicated estimation procedure (Tchetgen Tchetgen et al., 2017). As shown in the results section, there was no evidence against the hypothesis that outcomes are conditionally independent.

Estimation

To calculate point estimates for ADE, ASE, and ATE when data are conditionally independent, we estimated the following model using ordinary least squares:

$$E\left[Y_i^{hh} \middle| \boldsymbol{L} = \boldsymbol{l}, \boldsymbol{A} = \boldsymbol{a}\right] = \beta_0 + \beta_1 a_i^{hh} + \beta_2 l_i^{hh} + \beta_3 a_i^{net} + \beta_4 l_i^{net} + \beta_5 a_i^{hh} l_i^{hh} + \beta_6 l_i^{hh} a_i^{net} + \beta_7 l_i^{hh} l_i^{net}$$

Using results from this regression model we calculated ASE, ADE, and ATE as follows (details in Appendix):

$$ASE = \beta_3 \bar{d} + \beta_6 \frac{1}{n} \sum_i l_i d_i$$

$$ADE = \beta_1 + \beta_5 \bar{l}$$

$$ATE = \beta_1 + \beta_3 \bar{d} + \beta_5 \bar{l} + \beta_6 \frac{1}{n} \sum_i l_i d_i$$

Though we show only one covariate for ease of exposition, we entered the following covariates in the regression model: village, # connected households, # people in household, # women in household, # children in household, # number of elders in household, # people in connected households, # women in connected households, # children in connected households, # number of elders in connected households. We fitted interaction terms for a_i^{hh} and a_i^{net} with # women in household, # children in household, and # number of elders in household.

Treating within-household repeated measures as correlated, we used the cluster bootstrap to calculate standard errors. We re-sampled from the list of households 5000 times, each time fitting the above model. To construct confidence intervals for ADE, ASE, and ATE, we took the 0.025^{th} and the 0.975^{th} quantiles of the empirical distribution of each of these measures.

Potential Bias

ATE is only identified under the assumption that there is no unaccounted-for confounding. To identify ASE and ATE, we further assume that there are no unmeasured causes of physical

function which are correlated among family members. If the latter assumption had failed, household death would be correlated across connected households — a condition we test for (see sub-section on Model Selection above). Finally, we assume that the functional form we chose for the conditional mean model is correct.

Missing Data

Some households were missing data on *household wealth*. This induced missingness in *network wealth*. All other variables were fully ascertained. To account for missing data, we used a modified complete case analysis. As a sensitivity analysis, we conducted multiple imputation. The household wealth index measure was missing for 9-34% of households depending on the year (see Table 3.1 and additional tables in Appendix C).

Modified Complete Case Analysis

For the main analysis we used a modified complete case analysis. A household was excluded if it was missing *household wealth* or if all connected households were missing *household wealth*. If at least one connected household had valid *household wealth*, then *network wealth* was calculated by summing the valid values of household wealth and upweighting the sum by the ratio of the number connected households to the number of valid household wealth measurements.

In effect, for the purpose of calculating network wealth, we imputed missing values of household wealth using the mean of valid household wealth within each household's egocentric network i.e. for household i:

$$\text{HH Wealth}_i = \frac{\#\ \text{Connected Households}}{\#\ Valid\ HH\ Wealth\ Measures} \times sum\ of\ valid\ Connected\ HH\ Wealth\ Measures$$

Multiple Imputation

As a sensitivity analysis, we imputed missing values of household wealth using a custom R package based on the Multiple Imputation by Chained Equations (MICE) package (Buuren & Groothuis-Oudshoorn, 2011). Using the variables listed below, we performed predictive mean matching, randomly selecting from 5 of the closest matches. We created 64 imputed datasets, each time iterating the MICE algorithm 4 times. In each iteration, the model predicted missing *household wealth* and computed network wealth, time-lagged household wealth, and time-lagged network wealth. The latter three variables were included in the imputation model.

Dichotomous variables were imputed using a random draw from the Bernoulli distribution with probability of success given by predicted values from a logistic regression. Continuous variables were imputed using predictive mean matching (See Appendix B). The imputation model only produces valid inferences under the assumption that data were "missing at random" (Rubin, 1976) and that the functional forms we specified for predictive mean matching and logistic regression are correct.

Textbox: Variables in Imputation Model

Household Wealth*
Time-Lagged Household Wealth*

Network Wealth*
Time-Lagged Network Wealth*

Village

Number of Connected Households

Household Composition
- # children
- # adult women
- # adult men
- # elder women
- # elder men

Household Deaths
- > 0 children
- > 0 adult women
- > 0 adult men
- > 0 elder women
- > 0 elder men

Household Employment
- # currently working
- # not working

Household Unemployment
- # unemployed
- # not in labor market

* Household Wealth was the only variable with missing values. This induced missingness in *time-lagged household wealth*, *network wealth*, and *time-lagged network wealth*.

Results

Descriptive Statistics

From 2009-2018, the number of households in the Agincourt Health and Demographic Surveillance System grew from 16 725 to 21 291 (see Table 3.1 and additional tables in Appendix C). This increase was driven in part by the inclusion of new villages to the AHDSS over time. Household size and household gender and age composition did not appear to change substantially over this period. A minority of households were not connected by family relationships to other households. This proportion increased from 7.9% in 2009 to 9.4% in 2018. Over 90% of households were connected to fewer than 6 households.

On average, the rate of household death decreased over time. Whereas 94.6% of households did not experience death in 2009, in 2018 this figure increased to 97.5%. Household deaths were dominated by working age adults as compared to children and elders. In 2018, in 60% of households that experienced death, it was an adult who died, compared to 8% of households where it was a child or 33% in which it was an elder who died. Overall, in 2018 elders had the highest probability of death at 2.98% (177/5946) followed by adults (0.48%; 322/66 580) and children (0.10%; 44/41 203).

Average household wealth increased marginally from 2009 to 2018, and the standard deviation decreased. Across the years under study, households with lower wealth had lower household sizes and therefore lower rates of death (Figure 3.1). This is because households with smaller size were less likely to experience household death — there was a smaller number of people who could possibly die in any given year.

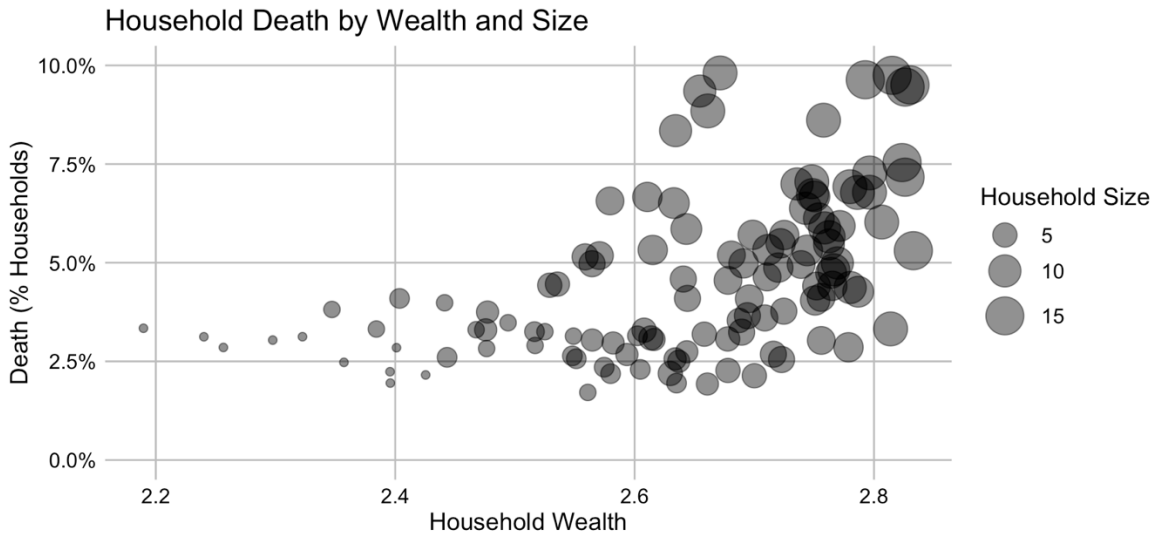

Figure 3.1: Household Deaths by Household Size and Household Wealth

There was no evidence of marginal network dependence in household deaths (Figure 3.2). i.e. Knowing that one household experienced death did not give any information about a connected household's likelihood of also experiencing death that year. By contrast, household wealth was positively network dependent. It is likely that if a given household has high wealth, the households it is connected to also have high wealth.

Household wealth was correlated within households over time (see Appendix C). Measures of SES that were close to each other in time were more strongly correlated than measures further apart in time. Since measures of household wealth for 2010 and 2012 were carried forward from 2009 and 2011, respectively, the correlation between 2009 and 2010 wealth, and 2011 and 2012 wealth is almost 1.

Network wealth was also correlated within households over time (see Appendix C). Like with *household wealth*, the strength of correlation attenuated as the distance (in time) between

measures increased. Overall, correlation among measures of *network wealth* was much stronger than correlation among *household wealth* measure

Model Selection

There was no statistical evidence that household death was correlated across connected households after adjusting for household wealth, network wealth, and covariates (Appendix C). However, there was statistically significant evidence, of within-household correlation of household death over time. To account for this dependence, we allowed for repeated measures of household death to be correlated over time by implementing the cluster bootstrap in order to compute confidence intervals.

Table 3.1: Descriptive statistics for Households in AHDSS

	2009 (N=16275)	2014 (N=21527)	2018 (N=21291)
# Connected Households			
0	4769 (29.3%)	7840 (36.4%)	7645 (35.9%)
1	3372 (20.7%)	4119 (19.1%)	4364 (20.5%)
2	2734 (16.8%)	3107 (14.4%)	3190 (15.0%)
3	1943 (11.9%)	2245 (10.4%)	2254 (10.6%)
4	1332 (8.2%)	1560 (7.2%)	1491 (7.0%)
5	799 (4.9%)	1045 (4.9%)	947 (4.4%)
6+	1326 (8.1%)	1611 (7.5%)	1400 (6.6%)
# Household Members			
1	1647 (10.1%)	2384 (11.1%)	2513 (11.8%)
2-5	7380 (45.3%)	10075 (46.8%)	10053 (47.2%)
6-10	5751 (35.3%)	7200 (33.4%)	7038 (33.1%)
11-20	1458 (9.0%)	1834 (8.5%)	1657 (7.8%)
21+	39 (0.2%)	34 (0.2%)	30 (0.1%)
# Children (<5)			
0	3487 (21.4%)	5012 (23.3%)	5651 (26.5%)
1	2751 (16.9%)	3959 (18.4%)	4033 (18.9%)
2-5	8938 (54.9%)	11375 (52.8%)	10744 (50.5%)
6-10	1041 (6.4%)	1136 (5.3%)	838 (3.9%)
11-20	57 (0.4%)	44 (0.2%)	25 (0.1%)
21+	1 (0.0%)	1 (0.0%)	0 (0%)
# Adults (15-59)			
0	551 (3.4%)	702 (3.3%)	680 (3.2%)
1	2993 (18.4%)	3907 (18.1%)	3701 (17.4%)
2-5	10815 (66.5%)	14245 (66.2%)	14318 (67.2%)
6-10	1851 (11.4%)	2585 (12.0%)	2506 (11.8%)
11-20	65 (0.4%)	88 (0.4%)	84 (0.4%)
21+	0 (0%)	0 (0%)	2 (0.0%)
# Elders (60+)			
0	12546 (77.1%)	16795 (78.0%)	16122 (75.7%)
1	3115 (19.1%)	3999 (18.6%)	4413 (20.7%)
2	590 (3.6%)	705 (3.3%)	735 (3.5%)
3+	24 (0.1%)	28 (0.1%)	21 (0.1%)
Household Wealth Index			
Mean (SD)	2.51 (0.464)	2.65 (0.417)	2.68 (0.387)
Median [Min, Max]	2.56 [0.902, 4.04]	2.69 [0.902, 4.02]	2.71 [0.921, 4.18]
Missing	1498 (9.2%)	4340 (20.2%)	4761 (22.4%)
Network Wealth Index			
Mean (SD)	5.29 (5.66)	5.08 (6.03)	4.97 (5.88)
Median [Min, Max]	3.78 [0, 46.3]	2.94 [0, 52.3]	2.93 [0, 56.7]
Missing	262 (1.6%)	744 (3.5%)	1025 (4.8%)

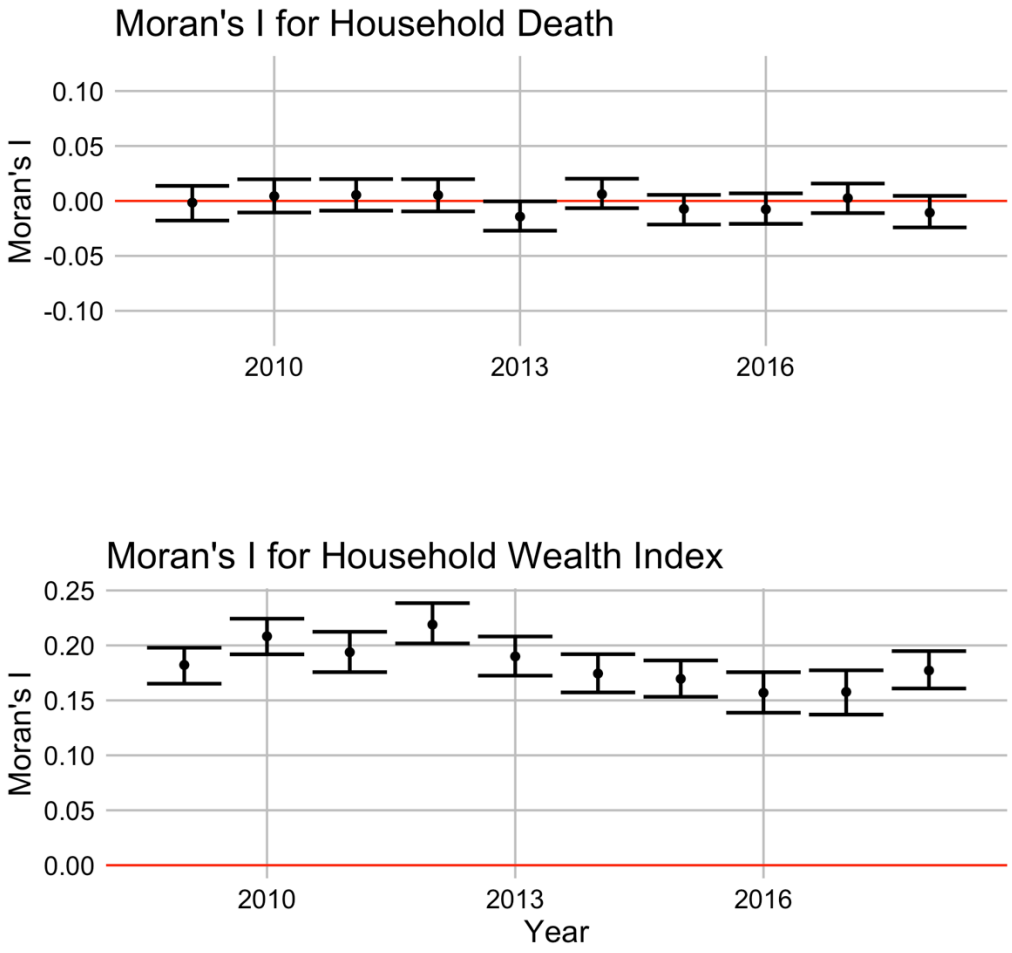

Figure 3.2: Network Autocorrelation in Household Deaths and Household Wealth

Modified Complete Case Analysis

There is significant statistical evidence of a network spillover effect of wealth on deaths (Table 3.2). Overall, a 1- standard deviation increase in household wealth led to a 0.69% (95% CI: 0.57 : 0.82) decrease in household deaths. Most of this was driven by the deaths of adults. A 1-

standard deviation increase in household wealth led to a 0.44% (95% CI: 0.35 : 0.54) decrease in household deaths of adults. There was a significant average direct effect of wealth on deaths overall and for each population. Only death among adults, however, was affected by the network spillover of increased wealth. A 1-standard deviation increase in household wealth led to a 0.1% (95% CI: 0.06 : 0.15) decrease in the proportion of households which experienced the death of an adult.

Overall, about 16% (95% CI: 7.8 : 24.3) of the effect of wealth on deaths can be attributed to spillover (Table 3.3). This proportion was highest for working age adults (23%; 95% CI: 13.9 : 33.5).

Table 3.2: Causal effects of household wealth on household death

Causal Estimands calculated at Observed Value of Covariates (Modified Complete Case Analysis, Full Graph)

	Average Direct Effect		Average Spillover Effect		Average Total Effect	
	Estimate	95% CI	Estimate	95% CI	Estimate	95% CI
Children	-0.17%	-0.21% : -0.12%	0.00%	-0.03% : 0.02%	-0.17%	-0.22% : -0.12%
Adults	-0.34%	-0.43% : -0.25%	-0.10%	-0.15% : -0.06%	-0.44%	-0.54% : -0.35%
Elders	-0.08%	-0.14% : -0.02%	-0.01%	-0.04% : 0.02%	-0.09%	-0.15% : -0.03%
Overall	-0.58%	-0.70% : -0.46%	-0.11%	-0.17% : -0.05%	-0.69%	-0.82% : -0.57%

Multiple Imputation

The multiple imputation analysis supports the main findings (Figure 3.3 and additional tables in Appendix C). According to this analysis, overall, a 1-standard deviation increase in household wealth would lead to a 0.08% (95% CI: -0.01 : 0.18) decrease in household deaths. Again, most of this was driven by the deaths of working age adults. A 1-standard deviation increase in household wealth would lead to a 0.37 (95% CI: 0.26 : 0.47) decrease in household deaths of working age adults.

Table 3.3: Proportion of total effect mediated by spillover

% Spillover calculated at observed Covariate values (Modified Complete Case Analysis)

	Total Effect		Proportion Mediated by Spillover	
	Estimate	95% CI	Estimate	95% CI
Children	-0.17%	-0.22% : -0.12%	1.04%	-14.25% : 14.22%
Adults	-0.44%	-0.54% : -0.35%	23.54%	13.90% : 33.50%
Elders	-0.09%	-0.15% : -0.03%	11.35%	-49.82% : 50.31%
Overall	-0.69%	-0.82% : -0.57%	16.02%	7.78% : 24.27%

Overall, conducting a multiple imputation analysis led to slightly attenuated effect estimates and wider confidence intervals. As a result, whereas the modified complete case analysis shows a statistically significant spillover effect overall and among adults, the multiple imputation analysis shows a marginally non-significant spillover effect overall, and a marginally significant spillover effect among adults.

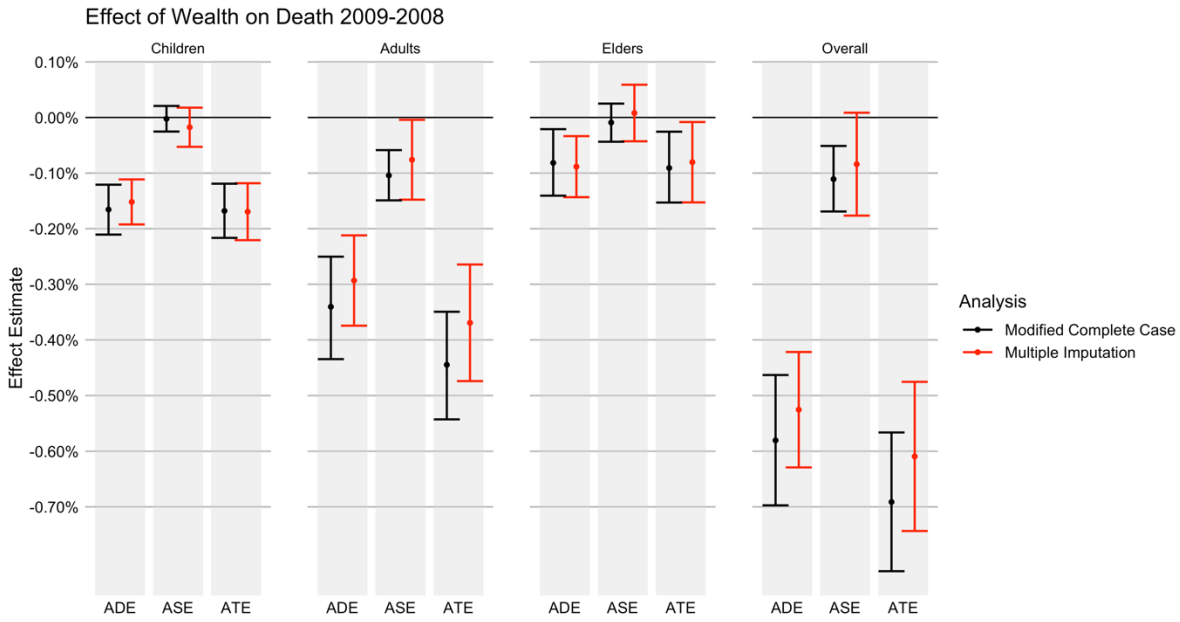

Figure 3.3: Effect of Household Wealth on Household Death

Discussion

Our study shows evidence that network wealth has a protective effect on mortality over and above individual household wealth. In the modified complete case analysis, there was a statistically significant spillover effect when considering premature deaths (i.e. deaths of people aged 15 to 59) as well as when considering all deaths. In the multiple imputation analysis, there was a statistically significant spillover effect when considering deaths before age 60, and a non-significant spillover effect when considering all deaths. Since the modified complete case analysis implicitly mean-imputed values of household wealth for the purpose of calculating network wealth, it likely underestimates uncertainty in the causal estimates. We place more weight, therefore, on the results from the multiple imputation analysis, which properly accounts for variability caused by the imputation itself.

These findings may be explained by transfers between connected households. In a 2005 nationally representative survey on giving in South Africa, over half of the respondents reported having recently given money, goods, food, or other items to members of their family who lived in a different household (Everatt et al., 2005). This giving was motivated by social norms. A more recent Cape Town study of young adults showed that a significant minority of black people reported feeling obligated to financially support some members of their extended family (Harper & Seekings, 2010), suggesting that social norms may be changing among younger, urban residents. In Agincourt, a rural area, households are embedded in local kinship-based networks of support. Members of extended family often live close to one another. “People and food flow constantly between them, and labour, including childcare and supervision, is commonly shared or exchanged” (Madhavan et al., 2014).

Our study makes contributions to several literatures. By measuring material resources available to individuals through their own households as well as the households of their family members, we more directly assess the impact of "resources that are accessed by individuals as a result of their membership in a network" — a definition that is often used in social capital and health research (Kawachi & Berkman, 2014). In contrast to the concepts and measures employed in past social capital studies in South Africa, however, our study measures social capital using the network-based approach that assesses network and household material resources. Whereas the *cohesion-based* approach gives primacy to the behaviors and perceptions that are theorized to facilitate the diffusion of resources through social networks, the *network-based approach* lends primacy to the resources themselves. The measure we use for this purpose — *network wealth* — is related to other measures that have attempted to assess resources, albeit indirectly: Lin's *position generator* (Lin et al., 2017), and van der Gaag and Snijder's *resource generator* (Van Der Gaag & Snijders, 2005).

We show that it is possible to use routinely collected data to measure the ability of members of extended family to provide material support. With respect to the social networks and health literature, we show that it is possible to construct and analyze the structure of sociocentric social networks using existing data. With respect to the family studies literature, we show that whereas most studies only study the effect of family support on health *within* the household, it is possible to measure the potential for family support across households (Harper & Seekings, 2010; Madhavan et al., 2017). This type of support is well-studied with respect to remittances from migrants, but less so with respect to transfers between non-migrant kin.

Our study faced some limitations. We make the assumption that mortality is independent of wealth from prior years after adjusting for wealth in the year immediately preceding the year in which we counted deaths. This assumption might not hold if household wealth in earlier life has

an impact throughout the life course, leading to an impact on mortality in later life, even after adjusting for wealth in the intervening time periods. Some studies have shown that socioeconomic status in early life can affect adult health directly (Birnie et al., 2011; Landös et al., 2019; Vable et al., 2019).

Furthermore, past research shows that household structure and composition are shaped by income and wealth (Madhavan et al., 2014; Makiwane et al., 2017). As such, it is not possible to disentangle the direct effect of wealth on mortality from the indirect effect of wealth on mortality through household structure and composition. For instance, it might be that the reason that the wealth of connected households affects death in a focal household has more to do with the movement of sick people between connected households than it does with the movement of resources between the households (See Figure 3.4 B). Even if this is the case, however, it would remain true that the wealth of connected households is protective against death in the focal household (Figure 3.4 A).

Unmeasured environmental variables may have caused correlation in household deaths among connected households (Figure 3.4 C) or might have shaped both household wealth and household deaths through a mechanism other than household wealth (Figure 3.4 D). We found no evidence, however, that household deaths were correlated across households — neither marginally or conditional on covariates and wealth. This is a necessary condition for there to be environmental confounding.

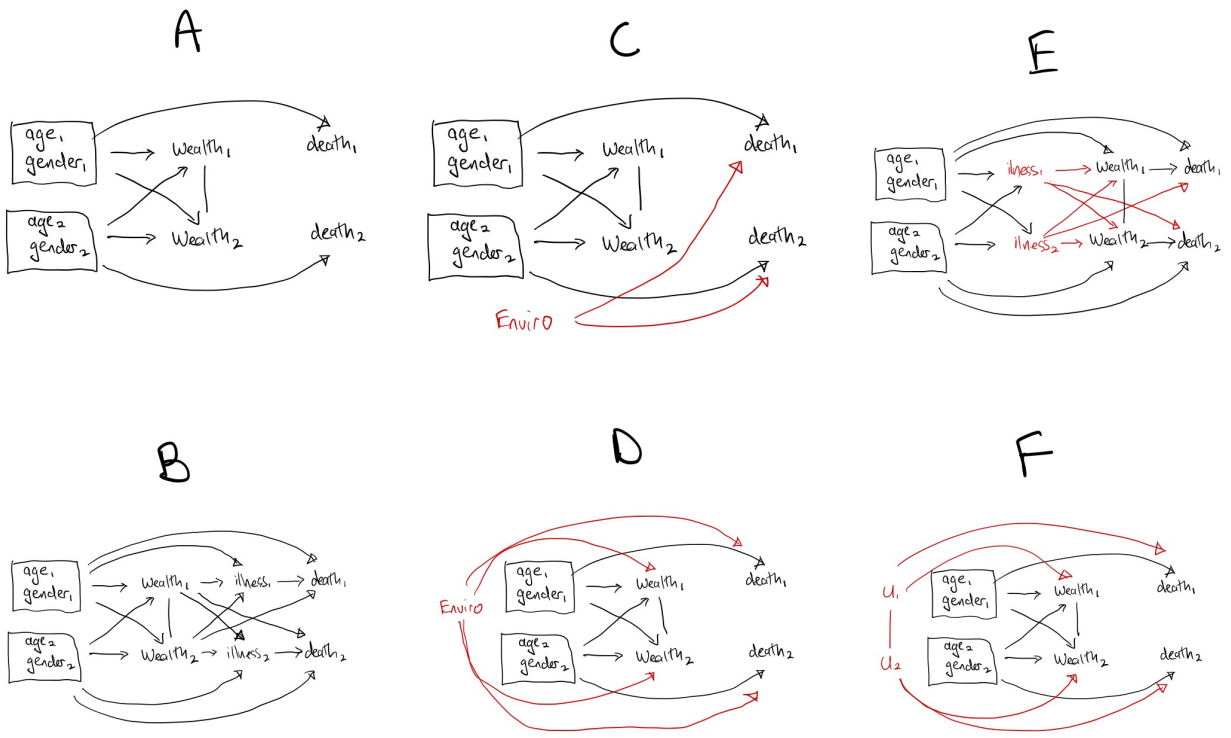

Figure 3.4: Example Chain Graphs Illustrating Bias under the null hypothesis of no treatment effect

Finally, it might be the case that illness in the household decreases household wealth over time (Figure 3.4 E) through loss of wages and costs associated with caring for the ill. In this case, illness would be an unmeasured confounder in the relationship between household wealth and household deaths. It is not clear whether illness would necessarily reduce household wealth since being ill enough to be incapable of work renders one eligible for the disability grant. In the context of widespread unemployment, this might increase household income and household wealth, biasing our estimates of the relationship between wealth and death towards the null. As Kelly et al. found in a recent study, “given the lack of adequate social provisioning for those who are able-bodied and unemployed, disability [grants are] highly valued in households” (Kelly, 2018, p. 1). Since health-related data were not uniformly available among residents of the AHDSS — we are unable to test this hypothesis using available data. More generally, there could be unmeasured confounding (Figure 3.4 F). The adaptation of approaches to sensitivity

analysis such as the E-Value are another avenue to pursue in future work (L. H. Smith & VanderWeele, 2019; VanderWeele & Ding, 2017).

A strength of this study is that, by using family relationship, we avoid the problem of disentangling homophily from spillover (Shalizi & Thomas, 2011). Since family relationship are largely established prior to wealth, we are assured that the statistical relationships we found are not an artefact of homophily. Further, we conducted this analysis making weak distributional assumptions about the data, allowing this approach to be applied in a wide variety of settings.

Our study contributes evidence that networked material resources shape mortality, with strongest effects seen on premature, and therefore potentially preventable, mortality. Further research is needed to assess the effect of networked resources on other health outcomes. Research on socioeconomic status and health should endeavor to incorporate the status of family members as well as that of the individual. To make this possible, it is crucial to find ways of using existing data, including passively collected data, to study social networks for the purpose of understanding population health.

Postscript

Towards a materialist social network epidemiology

In the months leading up to my qualifying exam, I visited an older friend in New York City who happened to be hosting other friends of his from out of town. The four or five of us sat down for lunch at a nearby restaurant and began, as strangers do, searching each other for topics of conversation. Despite my best efforts, one of my new acquaintances found this dissertation, then incipient, and drew the unwanted attention of the table to it. “What is your thesis about?” I explained briefly and vaguely that I intended to show that people share resources over social relationships and that these shared resources shape health. I was met with a polite but discouraging silence which was eventually broken by one of the guests — one who had failed to conceal her puzzlement. “But isn’t that obvious?” It took me a few seconds to recover and a few more to piece together a response. But before I managed to return fire, our conflict-averse host had derailed this invidious line of questioning by directing us to urgently consider our orders before the waiter returned, offering some entirely unsolicited suggestions. An amount of time passed that rendered it too late to respond without appearing to be defensive, so that vicious accusation has, to date, gone unanswered. But I have always had a memory for grudges, and over the past few months, have come to a considered response which I will now outline.

Yes. It is obvious that social relationships occasion the sharing of resources. My interlocutor and I both had the benefit of free accommodation in one of the most expensive cities in the country. The availability of that resource was predicated on our relationship with our friend. The same relationship had yielded me a monetarily free, if emotionally costly, lunch. What is not so obvious, however, is *how* these shared resources might influence health; that is, we do not know which causal pathways might connect the level and type of resources available through relationships to the biophysical processes that unfold in the body to shape health. In addition, there is an epistemic problem: the sub-field of causal inference within epidemiology does not offer widely understood conceptual tools for asking how people affect each other. In fact, the

idea that an individual's exposures and outcomes might be entangled with those of other individuals is generally understood to be a violation of the conditions under which the problem of inferring causality is tractable. Rather than excite innovation, this fact is met with defeatism in teaching and practice: classical causal inference methods underpinned by the SUTVA (or equivalent) assumptions are presented as if they have no alternatives. Finally, even if conceptual tools were widely available, network data are not. This kind of data is notoriously difficult to collect both logistically and in terms of the complex ethical challenges they present to researchers and ethical review boards.

This project has grappled with each of these challenges, making inroads on several fronts. Drawing on newly developed methods in network causal inference, it extends concepts well-understood by epidemiologists who are familiar with causal inference to allow for the explicit consideration of inter-dependence. These extensions allow the specification of useful causal estimands. With these in place, the project demonstrates an approach to estimation that is based on outcome regression — the workhorse of observational epidemiology. It applies these methods to routinely collected data that are not usually treated as network data but that are readily available to researchers. These data can be the basis of future network-analyses. To investigators that have already completed studies in the Agincourt Health and Demographic Surveillance System and the Africa Health Research Institute surveillance platform, they offer a new lens through which to examine collected data.

Empirical results from this project begin to address the question of *how* networked resources shape health. In Chapter 1, financial incentives for HIV testing are shown to not only influence the person who is offered the incentive, but their family members. These effects operate through a surprising pathway; the offer of a financial incentive to family members only affects the behavior of an individual if that individual herself was offered the incentive. In Chapter 3, the

wealth held by a household prolongs the lives of its members as well as those of non-household family members. This effect is pronounced for preventable deaths — deaths among those aged 16 to 59. By contrast, wealth held by family members does not appear to affect the physical function of older individuals in those households according to Chapter 2. In other words, while there is evidence that networked wealth affects health, we found no evidence that this effect is mediated by changes in physical function or disability among older individuals. Of course, there are many more causal pathways to investigate, including those that are initiated in early life.

Taken together, the empirical findings of the project lend evidence for a “multiplier effect” (VanderWeele & Christakis, 2019) of material resources on health: an increase in the wealth or resources of one person affects their own health, as well as the health of individuals closely connected to them. As argued by my classmate Emily Unger in her own doctoral dissertation, these multiplier effects are crucial to integrate in cost-benefit analyses that guide welfare policy. By failing to account for them, policy makers risk under-estimating the positive effects of cash transfers on health in South Africa and underestimating the social costs of retrenching welfare. For the same reasons, multiplier effects are important to understand in the context of health interventions and intervention research. The allocation of resources should be guided by the anticipated benefits of that allocation. We fail to accurately anticipate benefits when we act as if people are independent of one another. They are not.

To improve our knowledge of the role of human relationships in shaping population health, it is not only important to advance the science of the measurement and analysis of network data, but to contextualize analyses using concepts that illuminate, rather than obfuscate, the role of shared material resources. Early social capital literature, as exemplified by Bourdieu, offered an understanding of how social relationships shape the distribution of resources across society and over generations, grounding this understanding in an analysis of the political economy. To this

extent, social capital theory might have been a suitable home for this dissertation. However, subsequent work in this tradition, particularly the work that has pervaded social epidemiology, has all but evacuated the materiality of resources. The idealist assumptions evinced by this body of work are incompatible with the intent of my project. Further, the term 'social capital' itself has been used to such contradictory ends that it is scarcely informative. It signifies at once Bourdieu's analysis of the reproduction of class, Lin's theory of status attainment, Coleman's explication of methodological individualism, and Putnam's concern with social cohesion.

The social networks and health literature, on the other hand, does not carry such heavy a burden. Though it has roots in the psychosocial tradition, it is generally an umbrella category for work that considers how measured relationships shape health. It is this capaciousness that makes it more hospitable to my project. But this is not a strong enough basis alone.

In this moment of extreme and widening inequality, the field must equip itself with concepts that make the role of material *things* visible, illuminate their unequal and inequitable distribution in society, and by so doing, call for action on the causes and consequences of this inequity. That is, rather than assume extreme inequality in material resources to be natural and immutable, or beyond the scope of epidemiologic research, or to deny its salience by focusing exclusively on intangible forces such as feelings, behaviors, and perceptions, we must subject it to analysis.

This analysis should strive to unshroud the connections between individual- and ecological-levels of analysis, helping to clarify, for instance, the multiple pathways through which country- or neighborhood-level socioeconomic status shape individual health. In this project, I explored as a pathway the availability of material resources in social networks.

Finally, the analysis of health inequity should be truthful. It should acknowledge that extreme inequality benefits some groups of people at the expense of others. Those who benefit are aware — efforts to trick them into believing that it is in their interest to stem inequality are in vain. Efforts to equip the vibrant social movements of our time with useful information, on the other hand, might not be. If social epidemiologists are to contribute to social justice, we must be clear minded about our material world. Socialist feminist scholarship on social reproduction is an invaluable resource in this regard, though its thoughtful engagement in U.S. social epidemiology will require that we overcome a deeply held stigma against Marxian thought.

In summary, this project demonstrates that it is feasible to investigate the role of social networks in shaping health inequities using available data and substantive theory. It adapts cutting-edge epidemiologic methods to clarify the inferential target of such investigations and outlines appropriate statistical procedures for estimating the target. Far from being obvious, this research embraces the complexity of interdependent lives, promising new insights into the processes that shape population health.

It takes one step towards a materialist social network epidemiology.

References

Adjaye-Gbewonyo, D., Rebok, G. W., Gross, A. L., Gallo, J. J., & Underwood, C. R. (2019). Assessing urban-rural differences in the relationship between social capital and depression among Ghanaian and South African older adults. *PLoS ONE*, *14*(6). https://doi.org/10.1371/journal.pone.0218620

Ahrenfeldt, L. J., & Möller, S. (2021). The Reciprocal Relationship between Socioeconomic Status and Health and the Influence of Sex: A European SHARE-Analysis Based on Structural Equation Modeling. *International Journal of Environmental Research and Public Health*, *18*(9), 5045. https://doi.org/10.3390/ijerph18095045

Aiello, A. E. (2017). Invited Commentary: Evolution of Social Networks, Health, and the Role of Epidemiology. *American Journal of Epidemiology*, *185*(11), 1089–1092. https://doi.org/10.1093/aje/kwx076

Andrade, F. B. de, Duarte, Y. A. de O., Souza Junior, P. R. B. de, Torres, J. L., Lima-Costa, M. F., & Andrade, F. C. D. (2018). Inequalities in basic activities of daily living among older adults: ELSI-Brazil, 2015. *Revista De Saude Publica*, *52Suppl 2*(Suppl 2), 14s. https://doi.org/10.11606/S1518-8787.2018052000617

Baral, S., Rao, A., Sullivan, P., Phaswana-Mafuya, N., Diouf, D., Millett, G., Musyoki, H., Geng, E., & Mishra, S. (2019). The disconnect between individual-level and population-level HIV prevention benefits of antiretroviral treatment. *The Lancet HIV*.

Bassett, I. V., Wilson, D., Taaffe, J., & Freedberg, K. A. (2015). Financial incentives to improve progression through the HIV treatment cascade: *Current Opinion in HIV and AIDS*, *10*(6), 451–463. https://doi.org/10.1097/COH.0000000000000196

Baum, C. F., Schaffer, M. E., & Stillman, S. (2003). Instrumental variables and GMM: Estimation and testing. *The Stata Journal*, *3*(1), 1–31.

Berkman, L. F. (1992). Emotional Support and Survival after Myocardial Infarction: A Prospective, Population-based Study of the Elderly. *Annals of Internal Medicine*, *117*(12), 1003. https://doi.org/10.7326/0003-4819-117-12-1003

Berkman, L. F., Glass, T., Brissette, I., & Seeman, T. E. (2000). From social integration to health: Durkheim in the new millennium☆. *Social Science & Medicine*, *51*(6), 843–857.

Berkman, L. F., & Syme, S. L. (1979). Social networks, host resistance, and mortality: A nine-year follow-up study of Alameda County residents. *American Journal of Epidemiology*, *109*(2), 186–204.

Berkman, L., & Krishna, A. (2014). Social network epidemiology. *Social Epidemiology*, 234–289.

Bernstein, N. (1996, June 23). Ward of the State; The Gap in Ella Fitzgerald's Life. *The New York Times*. https://advance-lexis-com.ezp-prod1.hul.harvard.edu/api/document?collection=news&id=urn:contentItem:3S89-29F0-0005-G0W7-00000-00&context=1516831

Besag, J. (1974). Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society: Series B (Methodological)*, *36*(2), 192–225.

Bhattacharya, T. (2017). Introduction: Mapping social reproduction theory. *Social Reproduction Theory: Remapping Class, Recentering Oppression*, 1–20.

Birnie, K., Cooper, R., Martin, R. M., Kuh, D., Sayer, A. A., Alvarado, B. E., Bayer, A., Christensen, K., Cho, S., Cooper, C., Corley, J., Craig, L., Deary, I. J., Demakakos, P., Ebrahim, S., Gallacher, J., Gow, A. J., Gunnell, D., Haas, S., … HALCyon study team. (2011). Childhood socioeconomic position and objectively measured physical capability levels in adulthood: A systematic review and meta-analysis. *PloS One*, *6*(1), e15564. https://doi.org/10.1371/journal.pone.0015564

Borgatti, S. P., Mehra, A., Brass, D. J., & Labianca, G. (2009). Network Analysis in the Social Sciences. *Science*, *323*(5916), 892–895. https://doi.org/10.1126/science.1165821

Borges, V. S., Lima-Costa, M. F. F., & Andrade, F. B. de. (2020). A nationwide study on prevalence and factors associated with dynapenia in older adults: ELSI-Brazil. *Cadernos de Saúde Pública*, *36*(4), e00107319. https://doi.org/10.1590/0102-311x00107319

Bourdieu, P. (1986). The forms of capital. In J. G. Richardson, *Handbook of Theory and Research for the Sociology of Education* (pp. 241–258). Greenwood Press New York.

Brach, J. S., VanSwearingen, J. M., Newman, A. B., & Kriska, A. M. (2002). Identifying early decline of physical function in community-dwelling older women: Performance-based and self-report measures. *Physical Therapy*, *82*(4), 320–328.

Breeze, E., Sloggett, A., & Fletcher, A. (1999). Socioeconomic and demographic predictors of mortality and institutional residence among middle aged and older people: Results from the Longitudinal Study. *Journal of Epidemiology and Community Health*, *53*(12), 765–774. https://doi.org/10.1136/jech.53.12.765

Brennan-Olsen, S. L., Bowe, S. J., Kowal, P., Naidoo, N., Quashie, N. T., Eick, G., Agrawal, S., & D'Este, C. (2019). Functional Measures of Sarcopenia: Prevalence, and Associations with Functional Disability in 10,892 Adults Aged 65 Years and Over from Six Lower- and Middle-Income Countries. *Calcified Tissue International*, *105*(6), 609–618. https://doi.org/10.1007/s00223-019-00609-x

Buch, E. D. (2015). Anthropology of aging and care. *Annual Review of Anthropology*, *44*, 277–293.

Button, K., & Ncapai, T. (2019). Conflict and negotiation in intergenerational care: Older women's experiences of caring with the Old Age Grant in South Africa. *Critical Social Policy*, *39*(4), 560–581. https://doi.org/10.1177/0261018319867594

Buuren, S. van, & Groothuis-Oudshoorn, K. (2011). mice: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software*, *45*(3), 1–67.

Byass, P., Collinson, M. A., Kabudula, C., Gómez-Olivé, F. X., Wagner, R. G., Ngobeni, S., Silaule, B., Mee, P., Coetzee, M., Twine, W., Tollman, S. M., & Kahn, K. (2017). The long road to elimination: Malaria mortality in a South African population cohort over 21 years. *Global Health, Epidemiology and Genomics*, *2*, e11. https://doi.org/10.1017/gheg.2017.7

Carp, F. M. (1977). Retirement and physical health. In *Epidemiologic Studies in Psychosomatic Medicine* (Vol. 9, pp. 140–159). Karger Publishers.

Carrier, J. G. (2018). Moral economy: What's in a name. *Anthropological Theory*, *18*(1), 18–35.

Carrington, P. J., Scott, J., & Wasserman, S. (2005). *Models and methods in social network analysis* (Vol. 28). Cambridge university press.

Casale, D., & Posel, D. (2020). Gender and the early effects of the COVID-19 crisis in the paid and unpaid economies in South Africa. *NIDS-CRAM Policy Paper. Recuperado El*, *18*.

Casale, M., Boyes, M., Pantelic, M., Toska, E., & Cluver, L. (2019). Suicidal thoughts and behaviour among South African adolescents living with HIV: Can social support buffer the impact of stigma? *Journal of Affective Disorders*, *245*, 82–90. https://doi.org/10.1016/j.jad.2018.10.102

Casale, M., Cluver, L., Crankshaw, T., Kuo, C., Lachman, J. M., & Wild, L. G. (2015). Direct and Indirect Effects of Caregiver Social Support on Adolescent Psychological Outcomes in Two South African AIDS-Affected Communities. *American Journal of Community Psychology*, *55*(3–4), 336–346. https://doi.org/10.1007/s10464-015-9705-3

Case, A., & Deaton, A. (1998). Large Cash Transfers to the Elderly in South Africa. *The Economic Journal*, *108*(450), 1330–1361. https://doi.org/10.1111/1468-0297.00345

Chamie, G., Ndyabakira, A., Marson, K. G., Emperador, D. M., Kamya, M. R., Havlir, D. V., Kwarisiima, D., & Thirumurthy, H. (2020). A pilot randomized trial of incentive strategies to promote HIV retesting in rural Uganda. *PLOS ONE*, *15*(5), e0233600. https://doi.org/10.1371/journal.pone.0233600

Chamie, G., Schaffer, E. M., Ndyabakira, A., Emperador, D. M., Kwarisiima, D., Camlin, C. S., Havlir, D. V., Kahn, J. G., Kamya, M. R., & Thirumurthy, H. (2018). Comparative

effectiveness of novel nonmonetary incentives to promote HIV testing: *AIDS*, *32*(11), 1443–1451. https://doi.org/10.1097/QAD.0000000000001833

Chan, K. S., Kasper, J. D., Brandt, J., & Pezzin, L. E. (2012). Measurement equivalence in ADL and IADL difficulty across international surveys of aging: Findings from the HRS, SHARE, and ELSA. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, *67*(1), 121–132. https://doi.org/10.1093/geronb/gbr133

Chanda, M. M., Perez-Brumer, A. G., Ortblad, K. F., Mwale, M., Chongo, S., Kamungoma, N., Kanchele, C., Fullem, A., Barresi, L., Bärnighausen, T., & Oldenburg, C. E. (2017). Barriers and Facilitators to HIV Testing Among Zambian Female Sex Workers in Three Transit Hubs. *AIDS Patient Care and STDs*, *31*(7), 290–296. https://doi.org/10.1089/apc.2017.0016

Chandola, T. (2012). Spatial and social determinants of urban health in low-, middle- and high-income countries. *Public Health*, *126*(3), 259–261. https://doi.org/10.1016/j.puhe.2011.12.008

Chandola, T., Ferrie, J., Sacker, A., & Marmot, M. (2007). Social inequalities in self reported health in early old age: Follow-up of prospective cohort study. *BMJ (Clinical Research Ed.)*, *334*(7601), 990. https://doi.org/10.1136/bmj.39167.439792.55

Chimbindi, N., Mthiyane, N., Birdthistle, I., Floyd, S., McGrath, N., Pillay, D., Seeley, J., Zuma, T., Dreyer, J., Gareta, D., & others. (2018). Persistently high incidence of HIV and poor service uptake in adolescent girls and young women in rural KwaZulu-Natal, South Africa prior to DREAMS. *PLoS One*, *13*(10), e0203193.

Choko, A. T., Corbett, E. L., Stallard, N., Maheswaran, H., Lepine, A., Johnson, C. C., Sakala, D., Kalua, T., Kumwenda, M., Hayes, R., & Fielding, K. (2019). HIV self-testing alone or with additional interventions, including financial incentives, and linkage to care or prevention among male partners of antenatal care clinic attendees in Malawi: An adaptive multi-arm, multi-stage cluster randomised trial. *PLOS Medicine*, *16*(1), e1002719. https://doi.org/10.1371/journal.pmed.1002719

Chola, L., & Alaba, O. (2013). Association of Neighbourhood and Individual Social Capital, Neighbourhood Economic Deprivation and Self-Rated Health in South Africa – a Multi-Level Analysis. *PLoS ONE*, *8*(7). https://doi.org/10.1371/journal.pone.0071085

Christakis, N. A., & Fowler, J. H. (2007). The spread of obesity in a large social network over 32 years. *New England Journal of Medicine*, *357*(4), 370–379.

Christakis, N. A., & Fowler, J. H. (2013). Social contagion theory: Examining dynamic social networks and human behavior. *Statistics in Medicine*, *32*(4), 556–577.

Christian, A. K., Sanuade, O. A., Okyere, M. A., & Adjaye-Gbewonyo, K. (2020). Social capital is associated with improved subjective well-being of older adults with chronic non-

communicable disease in six low- and middle-income countries. *Globalization and Health*, *16*. https://doi.org/10.1186/s12992-019-0538-y

Coleman, J. (1990). *Foundations of Social Theory. Cambridge, MA: Belknap Press of Harvard University Press*.

Coleman, J. S. (1988). Social capital in the creation of human capital. *American Journal of Sociology*, *94*, S95–S120.

Conserve, D. F., Bay, C., Kilonzo, M. N., Makyao, N. E., Kajula, L., & Maman, S. (2019). Sexual and Social Network Correlates of Willingness to Self-Test for HIV Among Ever-Tested and Never-Tested Men: Implications for the Tanzania STEP Project. *AIDS Care*, *31*(2), 169–176. https://doi.org/10.1080/09540121.2018.1537466

Coppin, A. K., Ferrucci, L., Lauretani, F., Phillips, C., Chang, M., Bandinelli, S., & Guralnik, J. M. (2006). Low socioeconomic status and disability in old age: Evidence from the InChianti study for the mediating role of physiological impairments. *The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences*, *61*(1), 86–91. https://doi.org/10.1093/gerona/61.1.86

Cramm, J. M., Koolman, X., Møller, V., & Nieboer, A. P. (2011). Socio-economic status and self-reported tuberculosis: A multilevel analysis in a low-income township in the Eastern Cape, South Africa. *Journal of Public Health in Africa*, *2*(2). https://doi.org/10.4081/jphia.2011.e34

Cramm, J. M., Møller, V., & Nieboer, A. P. (2012). Individual- and Neighbourhood-Level Indicators of Subjective Well-Being in a Small and Poor Eastern Cape Township: The Effect of Health, Social Capital, Marital Status, and Income. *Social Indicators Research*, *105*(3), 581–593. https://doi.org/10.1007/s11205-011-9790-0

Cramm, J. M., & Nieboer, A. P. (2011). The influence of social capital and socio-economic conditions on self-rated health among residents of an economically and health-deprived South African township. *International Journal for Equity in Health*, *10*, 51. https://doi.org/10.1186/1475-9276-10-51

Csardi, G., & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal*, *Complex Systems*, 1695.

d'Orsi, E., Xavier, A. J., Steptoe, A., de Oliveira, C., Ramos, L. R., Orrell, M., Demakakos, P., & Marmot, M. G. (2014). Socioeconomic and lifestyle factors related to instrumental activity of daily living dynamics: Results from the English Longitudinal Study of Ageing. *Journal of the American Geriatrics Society*, *62*(9), 1630–1639. https://doi.org/10.1111/jgs.12990

D'Ambruoso, L., Kahn, K., Wagner, R. G., Twine, R., Spies, B., van der Merwe, M., Gómez-Olivé, F. X., Tollman, S., & Byass, P. (2016). Moving from medical to health systems classifications of deaths: Extending verbal autopsy to collect information on the

circumstances of mortality. *Global Health Research and Policy*, *1*, 2. https://doi.org/10.1186/s41256-016-0002-y

Demakakos, P., Cooper, R., Hamer, M., de Oliveira, C., Hardy, R., & Breeze, E. (2013). The bidirectional association between depressive symptoms and gait speed: Evidence from the English Longitudinal Study of Ageing (ELSA). *PloS One*, *8*(7), e68632. https://doi.org/10.1371/journal.pone.0068632

Deribew, A., Ojal, J., Karia, B., Bauni, E., & Oteinde, M. (2016). Under-five mortality rate variation between the Health and Demographic Surveillance System (HDSS) and Demographic and Health Survey (DHS) approaches. *BMC Public Health*, *16*(1), 1118. https://doi.org/10.1186/s12889-016-3786-2

Dodds, R., Denison, H., Ntani, G., Cooper, R., Cooper, C., Sayer, A., & Baird, J. (2012). Birth weight and muscle strength: A systematic review and meta-analysis. *The Journal of Nutrition, Health & Aging*, *16*(7), 609–615.

Ehsan, A., Klaas, H. S., Bastianen, A., & Spini, D. (2019). Social capital and health: A systematic review of systematic reviews. *SSM - Population Health*, *8*, 100425. https://doi.org/10.1016/j.ssmph.2019.100425

Elam, J. T., Graney, M. J., Beaver, T., el Derwi, D., Applegate, W. B., & Miller, S. T. (1991). Comparison of subjective ratings of function with observed functional ability of frail older persons. *American Journal of Public Health*, *81*(9), 1127–1130. https://doi.org/10.2105/ajph.81.9.1127

Ellaway, A. (2004). Commentary: Can subtle refinements of popular concepts be put into practice? *International Journal of Epidemiology*, *33*(4), 681–682; discussion 700-704. https://doi.org/10.1093/ije/dyh184

Everatt, D., Habib, A., Maharaj, B., & Nyar, A. (2005). Patterns of giving in South Africa. *Voluntas: International Journal of Voluntary and Nonprofit Organizations*, *16*(3), 275–291.

Fakier, K., & Cock, J. (2009). A Gendered Analysis of the Crisis of Social Reproduction in Contemporary South Africa. *International Feminist Journal of Politics*, *11*(3), 353–371. https://doi.org/10.1080/14616740903017679

Fieo, R. A., Austin, E. J., Starr, J. M., & Deary, I. J. (2011). Calibrating ADL-IADL scales to improve measurement accuracy and to extend the disability construct into the preclinical range: A systematic review. *BMC Geriatrics*, *11*, 42. https://doi.org/10.1186/1471-2318-11-42

Frankel, S., Smith, G. D., & Gunnell, D. (1999). Childhood socioeconomic position and adult cardiovascular mortality: The Boyd Orr Cohort. *American Journal of Epidemiology*, *150*(10), 1081–1084. https://doi.org/10.1093/oxfordjournals.aje.a009932

Fried, L. P., Herdman, S. J., Kuhn, K. E., Rubin, G., & Turano, K. (1991). Preclinical disability: Hypotheses about the bottom of the iceberg. *Journal of Aging and Health*, *3*(2), 285–300.

Friedman, M., Thoresen, C. E., Gill, J. J., Ulmer, D., Powell, L. H., Price, V. A., Brown, B., Thompson, L., Rabin, D. D., Breall, W. S., & others. (1986). Alteration of type A behavior and its effect on cardiac recurrences in post myocardial infarction patients: Summary results of the recurrent coronary prevention project. *American Heart Journal*, *112*(4), 653–665.

Gakidou, E., Hogan, M., & Lopez, A. D. (2004). Adult mortality: Time for a reappraisal. *International Journal of Epidemiology*, *33*(4), 710–717. https://doi.org/10.1093/ije/dyh099

Giacomin, K. C., Duarte, Y. A. O., Camarano, A. A., Nunes, D. P., & Fernandes, D. (2019). Care and functional disabilities in daily activities – ELSI-Brazil. *Revista de Saúde Pública*, *52*(Suppl 2), 9s. https://doi.org/10.11606/s1518-8787.2018052000650

Gill, T. M., Williams, C. S., Mendes de Leon, C. F., & Tinetti, M. E. (1997). The role of change in physical performance in determining risk for dependence in activities of daily living among nondisabled community-living elderly persons. *Journal of Clinical Epidemiology*, *50*(7), 765–772. https://doi.org/10.1016/s0895-4356(97)00065-6

Gjonca, E., Tabassum, F., & Breeze, E. (2009). Socioeconomic differences in physical disability at older age. *Journal of Epidemiology & Community Health*, *63*(11), 928–935. https://doi.org/10.1136/jech.2008.082776

Glymour, M. M., Weuve, J., Berkman, L. F., Kawachi, I., & Robins, J. M. (2005). When Is Baseline Adjustment Useful in Analyses of Change? An Example with Education and Cognitive Change. *American Journal of Epidemiology*, *162*(3), 267–278. https://doi.org/10.1093/aje/kwi187

Gómez-Olivé, F. X., Montana, L., Wagner, R. G., Kabudula, C. W., Rohr, J. K., Kahn, K., Bärnighausen, T., Collinson, M., Canning, D., Gaziano, T., Salomon, J. A., Payne, C. F., Wade, A., Tollman, S. M., & Berkman, L. (2018). Cohort Profile: Health and Ageing in Africa: A Longitudinal Study of an INDEPTH Community in South Africa (HAALSI). *International Journal of Epidemiology*, *47*(3), 689–690j. https://doi.org/10.1093/ije/dyx247

Gong, B., Shang, S., & Wu, C. (2020). Association between cognitive declines and disability in activities of daily living in older adults with COPD: Evidence from the China health and retirement longitudinal study. *BMJ Open*, *10*(10), e040098. https://doi.org/10.1136/bmjopen-2020-040098

Gorkin, L., Schron, E. B., Brooks, M. M., Wiklund, I., Kellen, J., Verter, J., Schoenberger, J. A., Pawitan, Y., Morris, M., Shumaker, S., & others. (1993). Psychosocial predictors of

mortality in the Cardiac Arrhythmia Suppression Trial-1 (CAST-1). *The American Journal of Cardiology*, *71*(4), 263–267.

Guralnik, J., Bandeen-Roche, K., Bhasin, S. a. R., Eremenco, S., Landi, F., Muscedere, J., Perera, S., Reginster, J.-Y., Woodhouse, L., & Vellas, B. (2020). Clinically Meaningful Change for Physical Performance: Perspectives of the ICFSR Task Force. *The Journal of Frailty & Aging*, *9*(1), 9–13. https://doi.org/10.14283/jfa.2019.33

Guralnik, J. M. (1997). Assessment of physical performance and disability in older persons. *Muscle & Nerve. Supplement*, *5*, S14-16.

Guralnik, J. M. (2004). Robine and Michel's "Looking forward to a general theory on population aging": Population aging across time and cultures: can we move from theory to evidence? *The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences*, *59*(6), M606-608; author reply M616-620. https://doi.org/10.1093/gerona/59.6.m606

Guralnik, J. M., Branch, L. G., Cummings, S. R., & Curb, J. D. (1989). Physical performance measures in aging research. *Journal of Gerontology*, *44*(5), M141-146. https://doi.org/10.1093/geronj/44.5.m141

Guralnik, J. M., Fried, L. P., & Salive, M. E. (1996). Disability as a public health outcome in the aging population. *Annual Review of Public Health*, *17*, 25–46. https://doi.org/10.1146/annurev.pu.17.050196.000325

Haberer, J. E., Musiimenta, A., Atukunda, E. C., Musinguzi, N., Wyatt, M. A., Ware, N. C., & Bangsberg, D. R. (2016). Short message service (SMS) reminders and real-time adherence monitoring improve antiretroviral therapy adherence in rural Uganda. *AIDS*, *30*(8), 1295–1299. https://doi.org/10.1097/QAD.0000000000001021

Halekoh, U., Højsgaard, S., & Yan, J. (2006). The R Package geepack for Generalized Estimating Equations. *Journal of Statistical Software*, *15/2*, 1–11.

Handa, S., Halpern, C. T., Pettifor, A., & Thirumurthy, H. (2014). The government of Kenya's cash transfer program reduces the risk of sexual debut among young people age 15-25. *PloS One*, *9*(1), e85473. https://doi.org/10.1371/journal.pone.0085473

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica: Journal of the Econometric Society*, 1029–1054.

Harling, G., Kobayashi, L. C., Farrell, M. T., Wagner, R. G., Tollman, S., & Berkman, L. (2020). Social contact, social support, and cognitive health in a population-based study of middle-aged and older men and women in rural South Africa. *Social Science & Medicine (1982)*, *260*. https://doi.org/10.1016/j.socscimed.2020.113167

Harper, S., & Seekings, J. (2010). *Claims on and obligations to kin in Cape Town, South Africa*.

Hermanstyne, K. A., Green, H. D., Cook, R., Tieu, H.-V., Dyer, T. V., Hucks-Ortiz, C., Wilton, L., Latkin, C., & Shoptaw, S. (2018). Social Network Support and Decreased Risk of Seroconversion in Black MSM: Results of the BROTHERS (HPTN 061) Study. *Journal of Acquired Immune Deficiency Syndromes (1999)*, *78*(2), 163–168. https://doi.org/10.1097/QAI.0000000000001645

Hernán, M., & Robins, J. M. (2018). *Causal inference*. Chapman & Hall/CRC.

Hicks, G. E., Shardell, M., Alley, D. E., Miller, R. R., Bandinelli, S., Guralnik, J., Lauretani, F., Simonsick, E. M., & Ferrucci, L. (2012). Absolute strength and loss of strength as predictors of mobility decline in older adults: The InCHIANTI study. *The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences*, *67*(1), 66–73. https://doi.org/10.1093/gerona/glr055

Hoddinott, G., Myburgh, H., de Villiers, L., Ndubani, R., Mantantana, J., Thomas, A., Mbewe, M., Ayles, H., Bock, P., Seeley, J., Shanaube, K., Hargreaves, J., Bond, V., & Reynolds, L. (2018). Households, fluidity, and HIV service delivery in Zambia and South Africa – an exploratory analysis of longitudinal qualitative data from the HPTN 071 (PopART) trial. *Journal of the International AIDS Society*, *21*(Suppl Suppl 4). https://doi.org/10.1002/jia2.25135

Institute of Medicine (US) Committee on Assessing Rehabilitation Science and Engineering. (1997). *Enabling America: Assessing the Role of Rehabilitation Science and Engineering* (E. N. Brandt & A. M. Pope, Eds.). National Academies Press (US). http://www.ncbi.nlm.nih.gov/books/NBK233578/

Jette, A. M., & Keysor, J. J. (2003). Disability models: Implications for arthritis exercise and physical activity interventions. *Arthritis and Rheumatism*, *49*(1), 114–120. https://doi.org/10.1002/art.10909

Jochelson, K. (2001). *The Colour of Disease: Syphilis and racism in South Africa, 1880-1950*. Springer.

Joint United Nations Program on HIV/AIDS. (2020). *UNAIDS data 2020*. Joint United Nations Program on HIV/AIDS. https://www.unaids.org/sites/default/files/media_asset/2020_aids-data-book_en.pdf

Joint United Nations Programme on HIV/AIDS (UNAIDS). (2019). *AIDSInfo*. https://aidsinfo.unaids.org/

Kabudula, C. W., Houle, B., Collinson, M. A., Kahn, K., Gómez-Olivé, F. X., Clark, S. J., & Tollman, S. (2017). Progression of the epidemiological transition in a rural South African setting: Findings from population surveillance in Agincourt, 1993–2013. *BMC Public Health*, *17*. https://doi.org/10.1186/s12889-017-4312-x

Kabudula, C. W., Houle, B., Collinson, M. A., Kahn, K., Gómez-Olivé, F. X., Tollman, S., & Clark, S. J. (2017). Socioeconomic differences in mortality in the antiretroviral therapy era in Agincourt, rural South Africa, 2001-13: A population surveillance analysis. *The Lancet. Global Health*, *5*(9), e924–e935. https://doi.org/10.1016/S2214-109X(17)30297-8

Kabudula, C. W., Houle, B., Collinson, M. A., Kahn, K., Tollman, S., & Clark, S. (2017). Assessing changes in household socioeconomic status in rural South Africa, 2001–2013: A distributional analysis using household asset indicators. *Social Indicators Research*, *133*(3), 1047–1073.

Kahn, K., Collinson, M. A., Gómez-Olivé, F. X., Mokoena, O., Twine, R., Mee, P., Afolabi, S. A., Clark, B. D., Kabudula, C. W., Khosa, A., & others. (2012). Profile: Agincourt health and socio-demographic surveillance system. *International Journal of Epidemiology*, *41*(4), 988–1001.

Kallio, V., Hämäläinen, H., Hakkila, J., & Luurila, O. (1979). Reduction in sudden deaths by a multifactorial intervention programme after acute myocardial infarction. *The Lancet*, *314*(8152), 1091–1094.

Katz, S., Ford, A. B., Moskowitz, R. W., Jackson, B. A., & Jaffe, M. W. (1963). Studies of Illness in the Aged—The Index of ADL: A Standardized Measure of Biological and Psychosocial Function. *JAMA*, *185*, 914–919. https://doi.org/10.1001/jama.1963.03060120024016

Kawachi, I., & Berkman, L. F. (2014). Social capital, social cohesion, and health. *Social Epidemiology*, 290–319.

Kawachi, I., Kim, D., Coutts, A., & Subramanian, S. V. (2004). Commentary: Reconciling the three accounts of social capital. *International Journal of Epidemiology*, *33*(4), 682–690; discussion 700-704. https://doi.org/10.1093/ije/dyh177

Kelly, G. (2018). The Influence of Disability-Related Cash Transfers on Family Practices in South Africa. *CSSR Working Paper No*, *414*, 2006.

Kelly, G. (2019). Disability, cash transfers and family practices in South Africa. *Critical Social Policy*, *39*(4), 541–559. https://doi.org/10.1177/0261018319867593

Korte, J. E., Strauss, M., Ba, A., Buregyeya, E., Matovu, J. K., Kisa, R., Musoke, W., Chemusto, H., Vrana-Diaz, C. J., Malek, A. M., Wanyenze, R. K., & George, G. (2019). HIV testing preferences among pregnant women attending antenatal care and their male partners: A discrete choice experiment in Uganda. *African Journal of AIDS Research*, *18*(4), 332–340. https://doi.org/10.2989/16085906.2019.1686032

Kranzer, K., Simms, V., Bandason, T., Dauya, E., McHugh, G., Munyati, S., Chonzi, P., Dakshina, S., Mujuru, H., Weiss, H. A., & Ferrand, R. A. (2018). Economic incentives for

HIV testing by adolescents in Zimbabwe: A randomised controlled trial. *The Lancet HIV*, *5*(2), e79–e86. https://doi.org/10.1016/S2352-3018(17)30176-5

Krieger, N. (2001). *Epidemiology and the people's health: Theory and context*. Oxford University Press.

Landös, A., von Arx, M., Cheval, B., Sieber, S., Kliegel, M., Gabriel, R., Orsholits, D., van der Linden, B. W. A., Blane, D., Boisgontier, M. P., Courvoisier, D. S., Guessous, I., Burton-Jeangros, C., & Cullati, S. (2019). Childhood socioeconomic circumstances and disability trajectories in older men and women: A European cohort study. *European Journal of Public Health*, *29*(1), 50–58. https://doi.org/10.1093/eurpub/cky166

Lau, Y. K., & Ataguba, J. E. (2015). Investigating the relationship between self-rated health and social capital in South Africa: A multilevel panel data analysis. *BMC Public Health*, *15*. https://doi.org/10.1186/s12889-015-1601-0

Lawlor, D. A., Smith, G. D., & Ebrahim, S. (2004). Association between childhood socioeconomic status and coronary heart disease risk among postmenopausal women: Findings from the British Women's Heart and Health Study. *American Journal of Public Health*, *94*(8), 1386–1392. https://doi.org/10.2105/ajph.94.8.1386

Lee, S.-Y., Tung, H.-H., Liu, C.-Y., & Chen, L.-K. (2018). Physical activity and sarcopenia in the geriatric population: A systematic review. *Journal of the American Medical Directors Association*, *19*(5), 378–383.

Lee, Y., & Ogburn, E. L. (2019). Network dependence and confounding by network structure lead to invalid inference. *ArXiv Preprint ArXiv:1908.00520*.

Leholha, P. (2012a). *Census 2011 Municipal report—Mpumalanga*. Statistics South Africa.

Leholha, P. (2012b). *The South Africa I know, the home I understand*. Statistics South Africa.

Lin, N. (2017). Building a network theory of social capital. In *Social capital* (pp. 3–28). Routledge.

Lin, N., Fu, Y., & Hsung, R.-M. (2017). The position generator: Measurement techniques for investigations of social capital. In *Social capital* (pp. 57–81). Routledge.

Link, B. G., & Phelan, J. (1995). Social conditions as fundamental causes of disease. *Journal of Health and Social Behavior*, 80–94.

Liu, J. X., Shen, J., Wilson, N., Janumpalli, S., Stadler, P., & Padian, N. (2019). Conditional cash transfers to prevent mother-to-child transmission in low facility-delivery settings: Evidence from a randomised controlled trial in Nigeria. *BMC Pregnancy and Childbirth*, *19*(1), 32. https://doi.org/10.1186/s12884-019-2172-3

Madhavan, S., & Brooks, A. (2015). Family Complexity in Rural South Africa: Examining Dynamism in Children's Living Arrangement and the Role of Kin. *Journal of Comparative Family Studies*, *46*(4), 483–498. https://doi.org/10.3138/jcfs.46.4.483

Madhavan, S., Clark, S., Beguy, D., Kabiru, C. W., & Gross, M. (2017). Moving beyond the household: Innovations in data collection on kinship. *Population Studies*, *71*(1), 117–132.

Madhavan, S., Mee, P., & Collinson, M. (2014). Kinship in Practice: Spatial Distribution of Children's Kin Networks. *Journal of Southern African Studies*, *40*(2), 401–418. https://doi.org/10.1080/03057070.2014.906211

Mahon, R. (2018). *Rethinking Welfare Regimes: Challenges from the South*.

Makiwane, M., Gumede, N. A., Makoae, M., & Vawda, M. (2017). Family in a changing South Africa: Structures, functions and the welfare of members. *South African Review of Sociology*, *48*(2), 49–69. https://doi.org/10.1080/21528586.2017.1288166

Mathenjwa, T., Kim, H.-Y., Zuma, T., Shahmanesh, M., Seeley, J., Matthews, P., Wyke, S., Mcgrath, N., Sartorius, B., Yapa, H., & others. (2019). Home-based intervention to test and start (HITS) protocol: A cluster-randomized controlled trial to reduce HIV-related mortality in men and HIV incidence in women through increased coverage of HIV treatment. *BMC Public Health*, *19*(1), 969.

Mee, P., Kahn, K., Kabudula, C. W., Wagner, R. G., Gómez-Olivé, F. X., Madhavan, S., Collinson, M. A., Tollman, S. M., & Byass, P. (2016). The development of a localised HIV epidemic and the associated excess mortality burden in a rural area of South Africa. *Global Health, Epidemiology and Genomics*, *1*. https://doi.org/10.1017/gheg.2016.3

Melzer, D., Lan, T.-Y., Tom, B. D. M., Deeg, D. J. H., & Guralnik, J. M. (2004). Variation in thresholds for reporting mobility disability between national population subgroups and studies. *The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences*, *59*(12), 1295–1303. https://doi.org/10.1093/gerona/59.12.1295

Miller, M. E., Magaziner, J., Marsh, A. P., Fielding, R. A., Gill, T. M., King, A. C., Kritchevsky, S., Manini, T., McDermott, M. M., Neiberg, R., Orwig, D., Santanasto, A. J., Pahor, M., Guralnik, J., Rejeski, W. J., & LIFE Investigators. (2018). Gait Speed and Mobility Disability: Revisiting Meaningful Levels in Diverse Clinical Populations. *Journal of the American Geriatrics Society*, *66*(5), 954–961. https://doi.org/10.1111/jgs.15331

Minkler, M., Fuller-Thomson, E., & Guralnik, J. M. (2006). Gradient of disability across the socioeconomic spectrum in the United States. *The New England Journal of Medicine*, *355*(7), 695–703. https://doi.org/10.1056/NEJMsa044316

Mitchell, C. U., & LaGory, M. (2002). Social capital and mental distress in an impoverished community. *City & Community*, *1*(2), 199–222.

Mohd Hairi, F., Mackenbach, J. P., Andersen-Ranberg, K., & Avendano, M. (2010). Does socio-economic status predict grip strength in older Europeans? Results from the SHARE study in non-institutionalised men and women aged 50+. *Journal of Epidemiology & Community Health*, *64*(9), 829–837. https://doi.org/10.1136/jech.2009.088476

Montoy, J. C. C., Dow, W. H., & Kaplan, B. C. (2018). Cash incentives versus defaults for HIV testing: A randomized clinical trial. *PLOS ONE*, *13*(7), e0199833. https://doi.org/10.1371/journal.pone.0199833

Moore, E. (2020). Financing Social Reproduction: Women's Responsibilities in Financing and Undertaking Household Social Reproduction in Multigenerational Households in South Africa. *Revue internationale des études du développement*, *N°242*(2), 37. https://doi.org/10.3917/ried.242.0037

Moreno, J. L. (1934). *Who shall survive?: A new approach to the problem of human interrelations.*

Moreno, J. L., & Jennings, H. H. (1938). Statistics of social configurations. *Sociometry*, 342–374.

MRC/Wits Agincourt Unit, Rural Public Health and Transitions Research Unit. (2021). https://www.agincourt.co.za

Mukoswa, G. M., Charalambous, S., & Nelson, G. (2017). The association between social capital and HIV treatment outcomes in South Africa. *PloS One*, *12*(11), e0184140.

Mulawa, M., Yamanis, T. J., Hill, L. M., Balvanz, P., Kajula, L. J., & Maman, S. (2016). Evidence of social network influence on multiple HIV risk behaviors and normative beliefs among young Tanzanian men. *Social Science & Medicine (1982)*, *153*, 35–43. https://doi.org/10.1016/j.socscimed.2016.02.002

Muntaner, C. (2004). Commentary: Social capital, social class, and the slow progress of psychosocial epidemiology. *International Journal of Epidemiology*, *33*(4), 674–680; discussion 700-704. https://doi.org/10.1093/ije/dyh200

Musheke, M., Ntalasha, H., Gari, S., McKenzie, O., Bond, V., Martin-Hilber, A., & Merten, S. (2013). A systematic review of qualitative findings on factors enabling and deterring uptake of HIV testing in Sub-Saharan Africa. *BMC Public Health*, *13*, 220. https://doi.org/10.1186/1471-2458-13-220

Navarro, V. (2004). Commentary: Is capital the solution or the problem? *International Journal of Epidemiology*, *33*(4), 672–674; discussion 700-704. https://doi.org/10.1093/ije/dyh179

Ndyabakira, A., Getahun, M., Byamukama, A., Emperador, D., Kabageni, S., Marson, K., Kwarisiima, D., Chamie, G., Thirumurthy, H., Havlir, D., Kamya, M. R., & Camlin, C. S.

(2019). Leveraging incentives to increase HIV testing uptake among men: Qualitative insights from rural Uganda. *BMC Public Health*, *19*(1). https://doi.org/10.1186/s12889-019-8073-6

Nelson, L. E., Wilton, L., Agyarko-Poku, T., Zhang, N., Zou, Y., Aluoch, M., Apea, V., Hanson, S. O., & Adu-Sarkodie, Y. (2015). Predictors of condom use among peer social networks of men who have sex with men in Ghana, West Africa. *PloS One*, *10*(1), e0115504. https://doi.org/10.1371/journal.pone.0115504

Neves, D., Samson, M., van Niekerk, I., Hlatshwayo, S., & Du Toit, A. (2009). *The use and effectiveness of social grants in South Africa*.

Oliveira, J. S., Pinheiro, M. B., Fairhall, N., Walsh, S., Franks, T. C., Kwok, W., Bauman, A., & Sherrington, C. (2020). Evidence on physical activity and the prevention of frailty and sarcopenia among older people: A systematic review to inform the world health organization physical activity guidelines. *Journal of Physical Activity and Health*, *17*(12), 1247–1258.

Olutola, B. G., & Ayo-Yusuf, O. A. (2012). Socio-Environmental Factors Associated with Self-Rated Oral Health in South Africa: A Multilevel Effects Model. *International Journal of Environmental Research and Public Health*, *9*(10), 3465–3483. https://doi.org/10.3390/ijerph9103465

Ortblad, K. F., Baeten, J. M., Cherutich, P., Wamicwe, J. N., & Wasserheit, J. N. (2019). The arc of HIV epidemics in sub-Saharan Africa: New challenges with concentrating epidemics in the era of 90–90–90. *Current Opinion in HIV and AIDS*, *14*(5), 354–365.

Orth-Gomer, K., Unden, A.-L., & Edwards, M.-E. (1988). Social Isolation and Mortality in Ischemic Heart Disease: A 10-Year Follow-up Study of 150 Middle-aged Men. *Acta Medica Scandinavica*, *224*(3), 205–215.

Osler, M., Madsen, M., Nybo Andersen, A.-M., Avlund, K., McGue, M., Jeune, B., & Christensen, K. (2009). Do childhood and adult socioeconomic circumstances influence health and physical function in middle-age? *Social Science & Medicine (1982)*, *68*(8), 1425–1431. https://doi.org/10.1016/j.socscimed.2009.01.014

Patel, L., & Mavungu, E. M. (2016). 'Children, families and the conundrum about men': Exploring factors contributing to father absence in South Africa and its implications for social and care policies. *South African Review of Sociology*, *47*(2), 19–39. https://doi.org/10.1080/21528586.2016.1161543

Patel, S. M., Duchowny, K. A., Kiel, D. P., Correa-de-Araujo, R., Fielding, R. A., Travison, T., Magaziner, J., Manini, T., Xue, Q., Newman, A. B., Pencina, K. M., Santanasto, A. J., Bhasin, S., & Cawthon, P. M. (2020). Sarcopenia Definition & Outcomes Consortium Defined Low Grip Strength in Two Cross-Sectional, Population-Based Cohorts. *Journal of the American Geriatrics Society*, *68*(7), 1438–1444. https://doi.org/10.1111/jgs.16419

Payne, C. F., Gómez-Olivé, F. X., Kahn, K., & Berkman, L. (2017). Physical Function in an Aging Population in Rural South Africa: Findings From HAALSI and Cross-National Comparisons With HRS Sister Studies. *The Journals of Gerontology: Series B*, *72*(4), 665–679. https://doi.org/10.1093/geronb/gbx030

Pedersen, T. L. (2019). *tidygraph: A Tidy API for Graph Manipulation*. https://CRAN.R-project.org/package=tidygraph

Pedersen, T. L. (2020). *ggraph: An Implementation of Grammar of Graphics for Graphs and Networks*. https://CRAN.R-project.org/package=ggraph

Perkins, J. M., Subramanian, S., & Christakis, N. A. (2015). Social networks and health: A systematic review of sociocentric network studies in low-and middle-income countries. *Social Science & Medicine*, *125*, 60–78.

Peterson, M. D., Duchowny, K., Meng, Q., Wang, Y., Chen, X., & Zhao, Y. (2017). Low Normalized Grip Strength is a Biomarker for Cardiometabolic Disease and Physical Disabilities Among U.S. and Chinese Adults. *The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences*, *72*(11), 1525–1531. https://doi.org/10.1093/gerona/glx031

Pfau-Effinger, B. (2005). Welfare state policies and the development of care arrangements. *European Societies*, *7*(2), 321–347.

Pope, A. M., & Tarlov, A. R. (1991). *Disability in America: Toward a national agenda for prevention.* ERIC.

Portes, A. (1998). Social capital: Its origins and applications in modern sociology. *Annual Review of Sociology*, *24*(1), 1–24.

Posel, D. (2016). *Inter-household transfers in South Africa: Prevalence, patterns and poverty* (No. 180). SALDRU, University of Cape Town.

Prasitsiriphon, O., & Pothisiri, W. (2018). Associations of Grip Strength and Change in Grip Strength With All-Cause and Cardiovascular Mortality in a European Older Population. *Clinical Medicine Insights. Cardiology*, *12*, 1179546818771894. https://doi.org/10.1177/1179546818771894

Putnam, R. (1995). Bowling alone: The strange disappearance of civic America. *Journal of Democracy*, *6*(1), 65–78.

Putnam, R. D. (2000). *Bowling alone: The collapse and revival of American community*. Simon and schuster.

Putnam, R. D. (2004). Commentary: “Health by association”: some comments. *International Journal of Epidemiology*, *33*(4), 667–671; discussion 700-704. https://doi.org/10.1093/ije/dyh204

Qian, J., & Ren, X. (2016). Association between comorbid conditions and BADL/IADL disability in hypertension patients over age 45: Based on the China health and retirement longitudinal study (CHARLS). *Medicine*, *95*(31), e4536. https://doi.org/10.1097/MD.0000000000004536

Rahe, R. H., Ward, H. W., & Hayes, V. (1979). Brief group therapy in myocardial infarction rehabilitation: Three-to four-year follow-up of a controlled trial. *Psychosomatic Medicine*.

Ramlagan, S., Peltzer, K., & Phaswana-Mafuya, N. (2013). Social capital and health among older adults in South Africa. *BMC Geriatrics*, *13*, 100. https://doi.org/10.1186/1471-2318-13-100

Rautio, N., Heikkinen, E., & Ebrahim, S. (2005). Socio-economic position and its relationship to physical capacity among elderly people living in Jyväskylä, Finland: Five- and ten-year follow-up studies. *Social Science & Medicine (1982)*, *60*(11), 2405–2416. https://doi.org/10.1016/j.socscimed.2004.11.029

Razavi, S. (2015). Care and social reproduction: Some reflections on concepts, policies and politics from a development perspective. *Oxford Handbooks: The Oxford Handbook of Transnational Feminist Movements*, 423–445.

Reuben, D. B., Valle, L. A., Hays, R. D., & Siu, A. L. (1995). Measuring physical function in community-dwelling older persons: A comparison of self-administered, interviewer-administered, and performance-based measures. *Journal of the American Geriatrics Society*, *43*(1), 17–23. https://doi.org/10.1111/j.1532-5415.1995.tb06236.x

Rich, B. (2018). *table1: Tables of Descriptive Statistics in HTML*. https://CRAN.R-project.org/package=table1

Rios, D. A., Abdulah, D. R., Wei, J. Y., & Hausdorff, J. (2001). Disparate effects of socioeconomic status on physical function and emotional well-being in older adults. *Aging Clinical and Experimental Research*, *13*(1), 30–37.

Riumallo-Herl, C., Canning, D., & Kabudula, C. (2019). Health inequalities in the South African elderly: The importance of the measure of social-economic status. *The Journal of the Economics of Ageing*, *14*, 100191.

Rodger, A. J., Cambiano, V., Bruun, T., Vernazza, P., Collins, S., Degen, O., Corbelli, G. M., Estrada, V., Geretti, A. M., Beloukas, A., Raben, D., Coll, P., Antinori, A., Nwokolo, N., Rieger, A., Prins, J. M., Blaxhult, A., Weber, R., Van Eeden, A., … Janeiro, N. (2019). Risk of HIV transmission through condomless sex in serodifferent gay couples with the HIV-positive partner taking suppressive antiretroviral therapy (PARTNER): Final results

of a multicentre, prospective, observational study. *The Lancet*, *393*(10189), 2428–2438. https://doi.org/10.1016/S0140-6736(19)30418-0

Rodger, A. J., Cambiano, V., Bruun, T., Vernazza, P., Collins, S., van Lunzen, J., Corbelli, G. M., Estrada, V., Geretti, A. M., Beloukas, A., Asboe, D., Viciana, P., Gutiérrez, F., Clotet, B., Pradier, C., Gerstoft, J., Weber, R., Westling, K., Wandeler, G., … for the PARTNER Study Group. (2016). Sexual Activity Without Condoms and Risk of HIV Transmission in Serodifferent Couples When the HIV-Positive Partner Is Using Suppressive Antiretroviral Therapy. *JAMA*, *316*(2), 171. https://doi.org/10.1001/jama.2016.5148

Rozzini, R., Frisoni, G. B., Ferrucci, L., Barbisoni, P., Bertozzi, B., & Trabucchi, M. (1997). The effect of chronic diseases on physical function. Comparison between activities of daily living scales and the Physical Performance Test. *Age and Ageing*, *26*(4), 281–287. https://doi.org/10.1093/ageing/26.4.281

RStudio Team. (2018). *RStudio: Integrated Development Environment for R*. RStudio, Inc. http://www.rstudio.com/

Ruberman, W., Weinblatt, E., Goldberg, J. D., & Chaudhary, B. S. (1984). Psychosocial influences on mortality after myocardial infarction. *New England Journal of Medicine*, *311*(9), 552–559.

Rubin, D. B. (1976). Inference and missing data. *Biometrika*, *63*(3), 581–592. https://doi.org/10.1093/biomet/63.3.581

Rutstein, S. O., Johnson, K., MEASURE, O. M., & others. (2004). *The DHS wealth index*. ORC Macro, MEASURE DHS.

Sager, M. A., Dunham, N. C., Schwantes, A., Mecum, L., Halverson, K., & Harlowe, D. (1992). Measurement of activities of daily living in hospitalized elderly: A comparison of self-report and performance-based methods. *Journal of the American Geriatrics Society*, *40*(5), 457–462. https://doi.org/10.1111/j.1532-5415.1992.tb02011.x

Sanderson, W. C., & Scherbov, S. (2014). Measuring the speed of aging across population subgroups. *PloS One*, *9*(5), e96289. https://doi.org/10.1371/journal.pone.0096289

Savino, E., Volpato, S., Zuliani, G., & Guralnik, J. M. (2014). Assessment of mobility status and risk of mobility disability in older persons. *Current Pharmaceutical Design*, *20*(19), 3099–3113. https://doi.org/10.2174/13816128113196660684

Sayer, A. A., Syddall, H., Martin, H., Patel, H., Baylis, D., & Cooper, C. (2008). The developmental origins of sarcopenia. *The Journal of Nutrition Health and Aging*, *12*(7), 427–432.

Schatz, E., Madhavan, S., Collinson, M., Gómez-Olivé, F. X., & Ralston, M. (2015). Dependent or Productive? A New Approach to Understanding the Social Positioning of Older South Africans Through Living Arrangements. *Research on Aging*, *37*(6), 581–605. https://doi.org/10.1177/0164027514545976

Schoenmaeckers, J. (2013). Filling in the gap: How to explain the divergence between subjective and objective health measures. In *Active ageing and solidarity between generations in Europe* (pp. 187–198). De Gruyter.

Schwartz, S., Prins, S. J., Campbell, U. B., & Gatto, N. M. (2016). Is the "well-defined intervention assumption" politically conservative? *Social Science & Medicine (1982)*, *166*, 254.

Seekings, J. (2008). *Beyond'fluidity': Kinship and households as social projects*. Centre for Social Science Research, University of Cape Town Cape Town.

Seekings, J. (2020). The Effects of Colonialism on Social Protection in South Africa and Botswana. In C. Schmitt (Ed.), *From Colonialism to International Aid* (pp. 109–135). Springer International Publishing. https://doi.org/10.1007/978-3-030-38200-1_5

Seekings, J., & Moore, E. (2014). Kin, market and state in the provision of care in South Africa. *Sorge: Arbeit, Verhältnisse, Regime*, 439–454.

Serrano-Alarcón, M., & Perelman, J. (2017). Ageing under unequal circumstances: A cross-sectional analysis of the gender and socioeconomic patterning of functional limitations among the Southern European elderly. *International Journal for Equity in Health*, *16*(1), 175. https://doi.org/10.1186/s12939-017-0673-0

Shalizi, C. R., & Thomas, A. C. (2011). Homophily and contagion are generically confounded in observational social network studies. *Sociological Methods & Research*, *40*(2), 211–239.

Sherman, S. E., & Reuben, D. (1998). Measures of functional status in community-dwelling elders. *Journal of General Internal Medicine*, *13*(12), 817–823. https://doi.org/10.1046/j.1525-1497.1998.00245.x

Sherr, L., Roberts, K. J., Mebrahtu, H., Tomlinson, M., Skeen, S., & Cluver, L. D. (2020). The food of life: An evaluation of the impact of cash grant receipt and good parenting on child nutrition outcomes in South Africa and Malawi. *Global Health Promotion*, *27*(4), 131–140. https://doi.org/10.1177/1757975920957598

Simbayi, L., Zuma, K., Zungu, N., Moyo, S., Marinda, E., Jooste, S., Mabaso, M., Ramlagan, S., North, A., Van Zyl, J., & others. (2019). *South African National HIV Prevalence, Incidence, Behaviour and Communication Survey, 2017: Towards achieving the UNAIDS 90-90-90 targets*.

Smith, G. D., Hart, C., Blane, D., Gillis, C., & Hawthorne, V. (1997). Lifetime socioeconomic position and mortality: Prospective observational study. *BMJ (Clinical Research Ed.)*, *314*(7080), 547–552. https://doi.org/10.1136/bmj.314.7080.547

Smith, G. D., & Lynch, J. (2004). Commentary: Social capital, social epidemiology and disease aetiology. *International Journal of Epidemiology*, *33*(4), 691–700; discussion 705-709. https://doi.org/10.1093/ije/dyh261

Smith, L. H., & VanderWeele, T. J. (2019). Mediational E-values: Approximate Sensitivity Analysis for Unmeasured Mediator-Outcome Confounding. *Epidemiology (Cambridge, Mass.)*, *30*(6), 835–837. https://doi.org/10.1097/EDE.0000000000001064

Song, L., Frazier, C. G., & Pettis, P. J. (2018). Do network members' resources generate health inequality? Social capital theory and beyond. In *Elgar Companion to Social Capital and Health*. Edward Elgar Publishing.

Sooryamoorthy, R., & Makhoba, M. (2016). The family in modern South Africa: Insights from recent research. *Journal of Comparative Family Studies*, *47*(3), 309–321.

Ssebagala, R. A. (2021). Plugging the Welfare Gap: The Role of Kinship Transfers in South Africa. *The Journal of Development Studies*, *57*(7), 1154–1171. https://doi.org/10.1080/00220388.2020.1850698

Statistics South Africa. (2019). *Labour market dynamics in South Africa, 2019* (No. 02-11-02 (2019)). http://www.statssa.gov.za/publications/Report-02-11-02/Report-02-11-022019.pdf

Statistics South Africa. (2021a). *General Household Survey 2019*. http://www.statssa.gov.za/publications/P0318/P03182019.pdf

Statistics South Africa. (2021b). *Key Statistics—Bushbuckridge*. Bushbuckridge. http://www.statssa.gov.za/?page_id=993&id=bushbuckridge-municipality

Statistics South Africa. (2021c). *Quarterly labour force survey—Q1 2021* (No. P0211). http://www.statssa.gov.za/publications/P0211/P02111stQuarter2021.pdf

Stoner, M. C. D., Edwards, J. K., Westreich, D., Kilburn, K., Ahern, J., Lippman, S. A., Gómez-Olivé, F. X., Kahn, K., & Pettifor, A. (2021). Modeling Cash Plus Other Psychosocial and Structural Interventions to Prevent HIV Among Adolescent Girls and Young Women in South Africa (HPTN 068). *AIDS and Behavior*. https://doi.org/10.1007/s10461-021-03158-3

Szreter, S. (2004). Author response: Debating mortality trends in 19th century Britain. *International Journal of Epidemiology*, *33*(4), 705–709. https://doi.org/10.1093/ije/dyh143

Szreter, S., & Woolcock, M. (2004a). Health by association? Social capital, social theory, and the political economy of public health. *International Journal of Epidemiology*, *33*(4), 650–667. https://doi.org/10.1093/ije/dyh013

Szreter, S., & Woolcock, M. (2004b). Rejoinder: Crafting rigorous and relevant social theory for public health policy. *International Journal of Epidemiology*, *33*(4), 700–704. https://doi.org/10.1093/ije/dyh263

Takada, S., Nyakato, V., Nishi, A., O'Malley, A. J., Kakuhikire, B., Perkins, J. M., Bangsberg, D. R., Christakis, N. A., & Tsai, A. C. (2019). The social network context of HIV stigma: Population-based, sociocentric network study in rural Uganda. *Social Science & Medicine (1982)*, *233*, 229–236. https://doi.org/10.1016/j.socscimed.2019.05.012

Tampubolon, G. (2015). Growing Up in Poverty, Growing Old in Infirmity: The Long Arm of Childhood Conditions in Great Britain. *PloS One*, *10*(12), e0144722. https://doi.org/10.1371/journal.pone.0144722

Tang, S., Xu, Y., Li, Z., Yang, T., & Qian, D. (2021). Does Economic Support Have an Impact on the Health Status of Elderly Patients With Chronic Diseases in China? - Based on CHARLS (2018) Data Research. *Frontiers in Public Health*, *9*, 658830. https://doi.org/10.3389/fpubh.2021.658830

Tanser, F., Kim, H.-Y., Mathenjwa, T., Shahmanesh, M., Seeley, J., Matthews, P., Wyke, S., McGrath, N., Adeagbo, O., Sartorius, B., Yapa, H. M., Zuma, T., Zeitlin, A., Blandford, A., Dobra, A., & Bärnighausen, T. (2021). Home-Based Intervention to Test and Start (HITS): a community-randomized controlled trial to increase HIV testing uptake among men in rural South Africa. *Journal of the International AIDS Society*, *24*(2), e25665. https://doi.org/10.1002/jia2.25665

Tavares Milhem Ygnatios, N., de Oliveira, C., Vaz de Melo Mambrini, J., Bof de Andrade, F., Lima-Costa, M. F., & Lustosa Torres, J. (2021). Differences in disability and nutritional status among older Brazilian and English adults: The Brazilian Longitudinal Study of Aging (ELSI-Brazil) and English Longitudinal Study of Aging (ELSA) cohorts. *The American Journal of Clinical Nutrition*, nqab060. https://doi.org/10.1093/ajcn/nqab060

Tchetgen Tchetgen, E. J., Fulcher, I. R., & Shpitser, I. (2020). Auto-g-computation of causal effects on a network. *Journal of the American Statistical Association*, 1–12.

Tchetgen Tchetgen, E. J., Fulcher, I., & Shpitser, I. (2017). Auto-g-computation of causal effects on a network. *ArXiv Preprint ArXiv:1709.01577*.

Thorpe, R. J., Koster, A., Kritchevsky, S. B., Newman, A. B., Harris, T., Ayonayon, H. N., Perry, S., Rooks, R. N., Simonsick, E. M., & Health, Aging, and Body Composition Study. (2011). Race, socioeconomic resources, and late-life mobility and decline: Findings from the Health, Aging, and Body Composition study. *The Journals of Gerontology. Series A,*

Biological Sciences and Medical Sciences, *66*(10), 1114–1123. https://doi.org/10.1093/gerona/glr102

Tollman, S. M., Norris, S. A., & Berkman, L. F. (2016). Commentary: The value of life course epidemiology in low-and middle-income countries: An ageing perspective. *International Journal of Epidemiology*, *45*(4), 997–999.

Tomita, A., & Burns, J. K. (2013). A multilevel analysis of association between neighborhood social capital and depression: Evidence from the first South African National Income Dynamics Study. *Journal of Affective Disorders*, *144*(1–2), 101–105. https://doi.org/10.1016/j.jad.2012.05.066

Torres, J. L., Lima-Costa, M. F., Marmot, M., & de Oliveira, C. (2016). Wealth and Disability in Later Life: The English Longitudinal Study of Ageing (ELSA). *PloS One*, *11*(11), e0166825. https://doi.org/10.1371/journal.pone.0166825

Tronto, J. (2014). The ethics of care, democracy and social inequalities. An interview. *Sorge: Arbeit, Verhältnisse, Regime*, 45–52.

Turrell, G., Lynch, J. W., Leite, C., Raghunathan, T., & Kaplan, G. A. (2007). Socioeconomic disadvantage in childhood and across the life course and all-cause mortality and physical function in adulthood: Evidence from the Alameda County Study. *Journal of Epidemiology and Community Health*, *61*(8), 723–730. https://doi.org/10.1136/jech.2006.050609

United Nations, Department of Economic and Social Affairs, Population Division. (2019). *World Population Prospects 2019, Volume II: Demographic Profiles (ST/ESA/SER. A/427)*.

Vable, A. M., Gilsanz, P., & Kawachi, I. (2019). Is it possible to overcome the “long arm” of childhood socioeconomic disadvantage through upward socioeconomic mobility? *Journal of Public Health (Oxford, England)*, *41*(3), 566–574. https://doi.org/10.1093/pubmed/fdz018

Van Der Gaag, M., & Snijders, T. A. (2005). The Resource Generator: Social capital quantification with concrete items. *Social Networks*, *27*(1), 1–29.

VanderWeele, T. J., & Christakis, N. A. (2019). Network multipliers and public health. *International Journal of Epidemiology*, *48*(4), 1032–1037.

VanderWeele, T. J., & Ding, P. (2017). Sensitivity analysis in observational research: Introducing the E-value. *Annals of Internal Medicine*, *167*(4), 268–274.

Verbrugge, L. M., & Jette, A. M. (1994). The disablement process. *Social Science & Medicine (1982)*, *38*(1), 1–14. https://doi.org/10.1016/0277-9536(94)90294-1

Verhoef, G. (2001). Informal financial service institutions for survival: African women and stokvels in urban South Africa, 1930–1998. *Enterprise & Society*, *2*(2), 259–296.

Wahrendorf, M., Reinhardt, J. D., & Siegrist, J. (2013). Relationships of disability with age among adults aged 50 to 85: Evidence from the United States, England and continental europe. *PloS One*, *8*(8), e71893. https://doi.org/10.1371/journal.pone.0071893

Wittenberg, M., & Collinson, M. A. (2007). Household transitions in rural South Africa, 1996—2003. *Scandinavian Journal of Public Health*, *35*(69_suppl), 130–137. https://doi.org/10.1080/14034950701355429

Wrong, D. H. (1963). *The oversocialized conception of man in modern sociology.*

Xu, D., Arling, G., & Wang, K. (2019). A cross-sectional study of self-rated health among older adults: A comparison of China and the United States. *BMJ Open*, *9*(7), e027895. https://doi.org/10.1136/bmjopen-2018-027895

Xue, Q.-L., Guralnik, J. M., Beamer, B. A., Fried, L. P., & Chaves, P. H. M. (2015). Monitoring 6-month trajectory of grip strength improves the prediction of long-term change in grip strength in disabled older women. *The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences*, *70*(3), 367–373. https://doi.org/10.1093/gerona/glu149

Yamanis, T. J., Fisher, J. C., Moody, J. W., & Kajula, L. J. (2016). Young Men's Social Network Characteristics and Associations with Sexual Partnership Concurrency in Tanzania. *AIDS and Behavior*, *20*(6), 1244–1255. https://doi.org/10.1007/s10461-015-1152-5

Yan, J. (2002). geepack: Yet Another Package for Generalized Estimating Equations. *R-News*, *2/3*, 12–14.

Yan, J., & Fine, J. P. (2004). Estimating Equations for Association Structures. *Statistics in Medicine*, *23*, 859–880.

Yotebieng, M., Thirumurthy, H., Moracco, K. E., Edmonds, A., Tabala, M., Kawende, B., Wenzi, L. K., Okitolonda, E. W., & Behets, F. (2016). Conditional Cash Transfers to Increase Retention in PMTCT Care, Antiretroviral Adherence, and Postpartum Virological Suppression: A Randomized Controlled Trial. *Journal of Acquired Immune Deficiency Syndromes (1999)*, *72 Suppl 2*, S124-129. https://doi.org/10.1097/QAI.0000000000001062

Zeileis, A. (2004). *Econometric computing with HC and HAC covariance matrix estimators*.

Zeileis, A. (2006). Object-Oriented Computation of Sandwich Estimators. *Journal of Statistical Software*, *16*(9), 1–16. https://doi.org/10.18637/jss.v016.i09

Zhong, Y., Wang, J., & Nicholas, S. (2017). Gender, childhood and adult socioeconomic inequalities in functional disability among Chinese older adults. *International Journal for Equity in Health*, *16*(1), 165. https://doi.org/10.1186/s12939-017-0662-3

Appendices

A: Family network spillover of micro-incentives for HIV testing

We estimate the effect of having an additional family member offered a financial incentive for HIV testing using two modelling approaches: An outcome regression model estimated using general estimating equations (GEE), and a set of structural mean models estimated using the general method of moments (GMM).

Methods for GEE Model

Model and Estimation

Using a log-risk ratio model, we regress HIV testing consent (Y) on individual treatment (T) and family treatment (F), while adjusting for categorical network size (S):

$$logP[Y_i = 1] = \theta_0 + \theta_1 T_i + \theta_2 F_i + \theta_3 T_i \times F_i + \theta_4^T S_i$$

This model is estimated using General Estimating Equations (GEE). It is estimated under the assumption that observations are independent, and separately under the assumption that observations are clustered within communities.

Methods for GMM Models

Model and Estimation

We define a semiparametric structural mean model of potential HIV testing consent ($Y_i^{t,f}$) under the intervention of individual treatment (t) and family treatment (f), conditional on the size and spatial distribution of family members (X). i.e. We model the person's potential HIV testing consent had, possibly contrary to fact, individual and family treatment been set to values t and f:

$$log\frac{P[Y_i^{t,f} = 1|X]}{P\left[Y_i^{0,0} = 1|X\right]} = \theta_1 t + \theta_2 f + \theta_3 t \times f \qquad (1)$$

where $\theta = (\theta_1\ \theta_2\ \theta_3)^T$ is the target of inference.

We use the Generalized Method of Moments (GMM) to estimate the parameters of the model (Baum et al., 2003; Hansen, 1982). In GMM estimation we use the fact that treatment is randomized, and therefore independent of counterfactual outcomes, to construct a number of *moment conditions.* These are moments of the population distribution of the data which are equal to 0. We estimate these moments using observed data. The set of model parameters under which the moments are approximately equal to 0 are, by definition, the GMM estimates of the model.

In our case, using the structural model (1) we have:

$$\begin{aligned} E\left[Y_i^{0,0}|X\right] &= e^{-\theta_1 t_i - \theta_2 f_i - \theta_3 t_i f_i} E\left[Y_i^{t_i, f_i}|X\right] \\ &= e^{-\theta_1 t_i - \theta_2 f_i - \theta_3 t_i f_i} E[Y_i|T_i,\ F_i, X] \\ &= E\left[Y_i e^{-\theta_1 T_i - \theta_2 F_i - \theta_3 T_i F_i}\ \middle|T_i,\ F_i, X\right] \end{aligned} \qquad \begin{matrix}\text{(exchangeability} \\ \text{and consistency)} \\ \\ (2)\end{matrix}$$

$$\begin{aligned} E\left[Y_i^{0,0}|X\right] &= E\left[E\left[Y_i^{0,0}|X\right]|X\right] \\ &= E\left[E\left[Y_i e^{-\theta_1 T_i - \theta_2 F_i - \theta_3 T_i F_i}\ \middle|T_i,\ F_i, X\right]|X\right] \\ &= E\left[Y_i e^{-\theta_1 T_i - \theta_2 F_i - \theta_3 T_i F_i}|X\right] \end{aligned} \qquad \begin{matrix}\text{(law of iterated} \\ \text{expectations)} \\ \\ (3)\end{matrix}$$

Defining $L_i^{0,0}(\theta) = Y_i e^{-\theta_1 T_i - \theta_2 F_i - \theta_3 T_i F_i}$, the above equation implies that under the true value of θ, $L_i^{0,0}$ is mean-independent of $(T_i,\ F_i)$ given the network size and spatial distribution of family members. Therefore $L_i^{0,0}(\theta)$ is mean-independent of any function of T_i and F_i given X. Since the entire analysis is conditioned on X, we suppress the conditioning statement in the equations below.

For k-dimensional random variable $f_k(T_i,\ F_i)$ with $E[f_k(T_i,\ F_i)] = 0_{k\times 1}$, we have:

$$E\left[L_i^{0,0}(\theta) \times f_k(T_i,\ F_i)\right] = 0_{k\times 1} \qquad (4)$$

We estimate $\hat{\theta}$ as the value of θ for which:

$$\frac{1}{n}\sum_{i=1}^{n} L_i^{0,0}(\theta) \times f(T_i,\, F_i) \approx 0 \qquad (5)$$

The system of equations (4) is called the moment conditions. We find the value of θ which satisfies (5) by minimizing the scalar quantity:

$$\kappa_k(\theta) = \left(\sum_{i=1}^{n} L_i^{0,0}(\theta) \times f_k(T_i,\, F_i)\right)^{T}_{1\times k} W_{k\times k} \left(\sum_{i=1}^{n} L_i^{0,0}(\theta) \times f_k(T_i,\, F_i)\right)_{k\times 1}$$

i.e.

$$\hat{\theta}^{(k)} = \underset{\theta}{\operatorname{argmin}}\, \kappa_k(\theta).$$

where W is a positive-definite $k \times k$ matrix called the *weights matrix*, n is the total number of observations, and the superscript on $\hat{\theta}^{(k)}$ indicates that the estimator is based on k-dimensional random variable f_k. It can be shown that the most efficient weights matrix W is:

$$W_{\text{eff}} = cov\left[\frac{1}{\sqrt{n}}\sum_{i=1}^{n} L_i^{0,0}(\theta) \times f_k(T_i,\, F_i)\right]^{-1}.$$

The Moment Conditions

To minimize $\kappa_k(\theta)$, we construct moment conditions as defined above. We define a set of k-dimensional mean-0 functions $f_k(T_i,\, F_i)$ such that k is at least as large as the dimension of the parameter vector θ (i.e. $k \geq 3$). This is because a system of moment conditions with $k < 3$ restrictions does not yield a unique solution for θ in (4). Having more restrictions than parameters (i.e. over-identifying θ) sometimes improves the efficiency of the estimator. On the other hand, over-identifying the parameter risks creating a system of moment conditions with no solution. We estimate θ using systems of 3, 4, and 5, moment conditions.

To construct functions f_k we define:

$\boldsymbol{A}$	$n \times n$ adjacency matrix for the family network
$\boldsymbol{A^w}$	$n \times n$ adjacency matrix setting entries of $\boldsymbol{A}$ to 0 if they represent between-community ties
$\boldsymbol{A^b}$	$n \times n$ adjacency matrix setting entries of $\boldsymbol{A}$ to 0 if they represent within-community ties
$\boldsymbol{T}$	n-dimensional vector of dichotomous treatments
$\boldsymbol{F}$	n-dimensional vector showing for each individual, the number of family members treated
$\mathbf{1_n}$	n-dimensional vector consisting of $1's$
$\mathbf{1_{n\times n}}$	$n \times n$ matrix consisting of $1's$
$\boldsymbol{D}$	$n \times n$ *bipartite projection of the community membership network*
$\odot$	element-wise matrix multiplication operator
$\boldsymbol{diag}\{\}$	For a given $n \times n$ matrix M, $diag\{M\}$ is an n −dimensional vector whose entries are the n entries on the main diagonal of M

Notes:

- $\boldsymbol{F = AT}$
- $\boldsymbol{A = A^b + A^w}$. This decomposition is useful since treatment is community-randomized. All alters within the same community have the same treatment status as the ego and all alters not in the same community have treatment status independent of the ego's.
- To obtain $\boldsymbol{D}$, we define a bipartite network connecting each individual to their community. If the total number of individuals is $\boldsymbol{n}$, the number of communities is $\boldsymbol{N}$, and the adjacency matrix for this bipartite network is $\boldsymbol{C_{n\times N}}$ then $\boldsymbol{D = CC^T}$. i.e. $\boldsymbol{D_{n\times n}}$ is an adjacency matrix such that if a pair of individuals are connected to the same community in the bipartite network represented by $\boldsymbol{C_{n\times N}}$ then the pair is connected to one another in the network represented by $\boldsymbol{D_{n\times n}}$. We set entries on the main diagonal of $\boldsymbol{D}$ to 1.
- For $\boldsymbol{a, b} \in \mathbb{R}^{\boldsymbol{n}}$, $\boldsymbol{a \odot\ b = b \odot\ a = diag\{ab^t\} = diag\{ba^t\}}$

The functions f_k are constructed as follows:

$$f_3(T_i,\ F_i) = \begin{pmatrix} T_i - \mu_{T_i} \\ F_i - \mu_{F_i} \\ T_i \times F_i - \mu_{T_i \times F_i} \end{pmatrix} \qquad f_4(T_i,\ F_i) = \begin{pmatrix} T_i - \mu_{T_i} \\ F_i - \mu_{F_i} \\ T_i \times F_i - \mu_{T_i \times F_i} \\ T_i \times F_i^2 - \mu_{T_i \times F_i^2} \end{pmatrix}$$

$$f_5(T_i,\ F_i) = \begin{pmatrix} T_i - \mu_{T_i} \\ F_i - \mu_{F_i} \\ T_i \times F_i - \mu_{T_i \times F_i} \\ T_i \times F_i^2 - \mu_{T_i \times F_i^2} \\ F_i^2 - \mu_{F_i^2} \end{pmatrix}$$

Where (see Calculations for Moment Conditions below):

$\mu_T = E[T]$

$$\mu_F = AE[\,T\,]$$

$$\mu_{T\times F} = E[T] \odot (A^b E[T] + A^w 1_n)$$

$$\mu_{T\times F^2} = E[T] \odot (diag\{A^b(E[T]E[T]' \odot (1_{n\times n} - D) + E[T]1_n' \odot D)A^{b'}\} + 2A^w 1_n \odot A^b E[T] + A^w 1_n \odot A^w 1_n)$$

$$\mu_{F^2} = diag\{A^b(E[T]E[T]' \odot (1_{n\times n} - D) + E[T]1_n' \odot D)A^{b'}\} + 2E[T] \odot A^w 1_n \odot A^b E[T] + E[T] \odot A^w 1_n \odot A^w 1_n$$

To check whether there is a solution to the moment conditions associated with f_3, f_4 and f_5 we conducted a J-test for over-specification (Baum et al., 2003). Under the null hypothesis for this test, $\kappa_k(\hat{\theta}^{(k)}) = 0$ and under the alternate hypothesis $\kappa_k(\hat{\theta}^{(k)}) > 0$.

We carried out estimation using 3, 4, and 5 moment conditions, choosing the most efficient estimator among those produced by these moment conditions. i.e.: we chose as our estimator $\hat{\theta}^{(k)}$ such that $var[a^t\hat{\theta}^{(k)}] \leq var[a^t\hat{\theta}^{(l)}]$ for $l \in \{3,4,5\}$ and $a^t \in \mathbb{R}^3$. i.e. we chose $k \in \{3,4,5\}$ such that $cov[\hat{\theta}^{(k)}] - cov[\hat{\theta}^{(l)}]$ was negative semi-definite[7] for $l \in \{3,4,5\}$.

The Weights Matrix

To consistently estimate the efficient weights matrix W_{eff}, it is necessary to make assumptions about the data-generating process. We made two sets of assumptions separately. Under the assumption that observations are independent and identically distributed, we estimate W_{eff} using:

$$\widehat{W}_{\text{eff}}^{indep} = \left(\frac{1}{n}\sum_{i=1}^{n}\left(f_{k,i}L_i^{0,0}\right)\left(f_{k,i}L_i^{0,0}\right)^T\right)^{-1} \quad (6)$$

[7] i.e. the matrix determinant of $cov[\hat{\theta}^{(k)}] - cov[\hat{\theta}^{(l)}]$ is non-positive

whereas under the assumption that observations are dependent within clusters but independent across clusters, we use:

$$\widehat{W}_{\text{eff}}^{clust} = \left(\frac{1}{n} \sum_{c=1}^{N} \left(f_{k,c} L_c^{0,0} \right) \left(f_{k,c} L_c^{0,0} \right)^T \right)^{-1} \quad (7)$$

Where

$$f_{k,c}(\boldsymbol{T}_c, \boldsymbol{F}_c) = \left[f_k(T_{c,1}, F_{c,1}) \quad f_k(T_{c,2}, F_{c,2}) \quad \dots \quad f_k(T_{c,n_c}, F_{c,n_c}) \right]_{k \times n_c}$$

$$L_c^{0,0}(\theta, \boldsymbol{T}_c, \boldsymbol{F}_c) = \begin{pmatrix} L_1(\theta) \\ L_2(\theta) \\ \vdots \\ L_{n_c}(\theta) \end{pmatrix}_{n_c \times 1}$$

n_c is the number of observations in cluster c, $\boldsymbol{T}_c$ is the n_c-dimensional sub-vector of $\boldsymbol{T}$ consisting of entries from individuals in cluster c and $T_{c,i}$ is the i^{th} entry in that sub-vector. $\boldsymbol{F}_c$ is similarly defined. N is the total number of clusters. Clusters, here, represent communities in the HITS study.

Asymptotic Properties

Under some regularity conditions, estimates from a GMM estimator are consistent and asymptotically normal. i.e.

$$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} N\left(0, \left(\mathrm{G}^{\mathrm{T}} W_{\text{eff}} G\right)^{-1}\right)$$

where $\xrightarrow{d}$ indicates convergence in distribution and

$$\mathrm{G} = \mathrm{E}\left[\sum_{i=1}^{n} \frac{d}{d\theta} \left(L_i^{0,0}(\theta) \times f_k(T_i,\ F_i) \right) \right].$$

We use these asymptotic properties to construct confidence intervals and hypothesis tests, estimating G using

$$\widehat{\mathrm{G}} = \frac{1}{\mathrm{n}}\sum_{i=1}^{n} \frac{d}{d\theta}\left(L_i^{0,0}(\theta) \times f_k(T_i,\, F_i)\right)$$

and estimating W_{eff} using $\widehat{W}_{\mathrm{eff}}^{indep}$ and $\widehat{W}_{\mathrm{eff}}^{clust}$ separately (see (6) and (7)).

Remark on Independence Assumptions

We note that both $\widehat{W}_{\mathrm{eff}}^{indep}$ and $\widehat{W}_{\mathrm{eff}}^{clust}$ are not consistent with our understanding of the data generating process. We believe that observations might be correlated across family ties and possibly across space. If this assumption is correct $\widehat{W}_{\mathrm{eff}}^{indep}$ will tend to grossly under-estimate standard errors for model estimates. On the other hand, it is likely that observations are more strongly correlated among household members and between individuals and their spatial neighbors, but weakly correlated between people in the same community who are not socially or spatially proximate. Therefore $\widehat{W}_{\mathrm{eff}}^{clust}$ might overstate correlation between observations and inflate standard errors. Since network connections are much denser within communities than between communities, $\widehat{W}_{\mathrm{eff}}^{clust}$ likely accounts for most of the correlation that arises as a result of spatial and social proximity. In summary, $\widehat{W}_{\mathrm{eff}}^{indep}$ is likely too anti-conservative, and $\widehat{W}_{\mathrm{eff}}^{clust}$ is likely too conservative. Hence, we conduct the analysis using both, and present in the main results estimates using the more conservative of the two.

Results

The effect of individual treatment in the absence of family treatment is positive. This estimated effect is stronger according to GEE models (Risk Ratio = 1.21) as compared to the GMM estimates (RR ≈ 1.13). Confidence intervals do not include the null for models estimated using

the independence assumption but do for models estimated using the clustering assumption. A similar pattern is observed for the effect of family treatment in the absence of individual treatment, except the risk ratios indicate that increased family treatment in the absence of individual treatment leads to a lower probability of consenting for HIV testing.

The interaction term shows that in the presence of individual treatment, family treatment leads to a higher propensity to consent for HIV testing than in the absence of individual treatment. The interaction effect is higher when estimated using GMM models (RR ≈ 1.19) as compared to GEE models (RR $= 1.08$). Confidence intervals do not include the null when we assume independence, and marginally include or exclude the null when we assume clustering. High p-values for the J-test for over-specification indicate that models have unique solutions.

Table 0.1: Model Results from GMM Estimation

	Individual Treatment $\exp(\theta_1)$		**Network Treatment** $\exp(\theta_2)$		**Interaction** $\exp(\theta_3)$		**Over-specification Test**
	RR	**95% CI**	**RR**	**95% CI**	**RR**	**95% CI**	**p**
Independence Assumption							
GEE	1.21	1.16-1.26	0.94	0.89-0.99	1.08	1.03-1.14	•
GMM 3	1.13	1.06-1.20	0.90	0.84-0.97	1.19	1.11-1.28	•
GMM 4	1.13	1.06-1.20	0.90	0.84-0.97	1.19	1.11-1.28	0.879
GMM 5	1.13	1.06-1.20	0.90	0.84-0.97	1.20	1.11-1.29	0.820
Clustering Assumption							
GEE	1.21	0.85-1.73	0.94	0.87-1.01	1.08	0.99-1.19	•
GMM 3	1.13	0.60-2.11	0.90	0.77-1.06	1.19	0.99-1.44	•
GMM 4	1.13	0.61-2.11	0.90	0.77-1.05	1.19	1.00-1.43	0.938
GMM 5	1.15	0.61-2.15	0.92	0.80-1.07	1.16	0.99-1.37	0.808

Table 0.2 uses results from Table 0.1 to show the effect on HIV Testing of two hypothetical joint interventions:

- Intervention 1: the individual incentive is withheld from the ego and an incentive is offered to one family member vs. the individual incentive is withheld from the ego and the incentive is withheld from family members.
- Intervention 2: the individual incentive is offered to the ego and offered to a family member vs. the individual incentive is offered to the ego and withheld from the family member

Table 0.2: Joint Intervention Effects for Individual and Family treatment. Risk Ratio for having and Additional Family Member be offered the Incentive

	1. Withhold Individual Incentive			2. Offer Individual Incentive		
	RR	**95% CI**	**p**	**RR**	**95% CI**	**p**
Independence Assumption						
GEE	0.94	0.87-1.02	0.127	1.02	1.01-1.03	0.001
GMM 3	0.90	0.84-0.97	0.006	1.08	1.05-1.10	0.000
GMM 4	0.90	0.84-0.97	0.006	1.08	1.05-1.10	0.000
GMM 5	0.90	0.84-0.97	0.004	1.08	1.05-1.10	0.000
Clustering Assumption						
GEE	0.94	0.87-1.02	0.127	1.02	1.01-1.03	0.001
GMM 3	0.90	0.77-1.06	0.211	1.08	1.00-1.16	0.046
GMM 4	0.90	0.77-1.05	0.194	1.08	1.01-1.15	0.035
GMM 5	0.92	0.80-1.07	0.269	1.07	1.00-1.14	0.041

Intervention 1 is shown to reduce the propensity to consent for HIV testing (RR ≈ 0.90). The confidence intervals for this risk ratio do not include the null when we assume data are independent but do include the null when we assume that data are clustered.

Intervention 2 is shown to increase the propensity to consent for HIV testing. Risk ratios are RR = 1.02 for the GEE model and RR ≈ 1.08 for GMM models. All the confidence intervals do not include the null, though confidence intervals based on the assumption that observations are clustered have lower bounds that are close to the null.

Table 0.3 compares all the estimators based on their variance. The symbol ++ (–) indicates that the estimator in the row has variance higher (lower) than the estimator in the column. In general, GMM estimators have higher variance than GEE estimators, and GMM estimators increase in variance as we increase the number of moment conditions.

Table 0.3: Relative variance of GEE and GMM estimators

	GEE	GMM 3	GMM 4
Independence Assumption			
GMM 3	++		
GMM 4	++	++	
GMM 5	++	++	++
Clustering Assumption			
GMM 3	++		
GMM 4	++	++	
GMM 5	++	++	++

Finally, Table 0.4 compares the variance of estimators based on the assumption of clustering vs. independence. As expected, variance is higher when we assume the data are clustered.

Table 0.4: Relative variance of Clustered and Independence Estimators

	Clustered vs. Independence Assumption
GEE	++
GMM 3	++
GMM 4	++
GMM 5	++

Since it has lowest variance among the GMM models, and since it makes fewer assumptions than the GEE models, we use model GMM 3 for our final estimate of spillover. We use the estimator based on the assumption that observations are clustered. Not being offered the

incentive and having an additional family member offered the incentive did not lead to a significantly different probability of consenting for an HIV test compared with not being offered the incentive and having no additional family member offered the incentive (RR=0.90 95% CI: 0.77-1.06). On the other hand, being offered the incentive and having an additional family member be offered the incentive led to an improvement in the probability of consenting for an HIV test (RR=1.08 95% CI: 1.00-1.16) compared to being offered the incentive and not having an additional member that was offered the incentive.

Calculations for Moment Conditions

$$\mu_T = E[\,T\,]$$

$$\mu_F = E[\,AT\,] = AE[T]$$

$$\begin{aligned}\mu_{T\times F} &= E[\,T \odot F\,] \\ &= E[\,T \odot AT\,] \\ &= E[\,T \odot A^bT] + E[\,T \odot A^wT] \\ &= E[T] \odot A^bE[T] + E[\,T \odot T \odot A^w1_n] \\ &= E[T] \odot A^bE[T] + E[T \odot T] \odot A^w1_n \\ &= E[T] \odot A^bE[T] + E[T] \odot A^w1_n \\ &= E[T] \odot (A^bE[T] + A^w1_n)\end{aligned}$$

$$\begin{aligned}\mu_{T\times F^2} &= E[T \odot F \odot F] \\ &= E[T \odot AT \odot AT] \\ &= E[T \odot (A^bT + A^wT) \odot (A^bT + A^wT)] \\ &= E[T \odot A^bT \odot A^bT] + 2E[T \odot A^wT \odot A^bT] + E[T \odot A^wT \odot A^wT] \\ &= E[T] \odot E[A^bT \odot A^bT] + 2E[T \odot A^wT] \odot E[A^bT] + E[T] \odot A^w1_n \odot A^w1_n \\ &= E[T] \odot diag\{A^bE[TT']A^{b'}\} + 2E[T \odot A^w1_n] \odot A^bE[T] + A^w1_n \odot A^w1_n \odot E[T] \\ &= E[T] \odot \left(diag\{A^b(E[T]E[T]' \odot (1_{n\times n} - D) + E[T]1_n' \odot D)A^{b'}\} + 2A^w1_n \odot A^bE[T] + A^w1_n \odot A^w1_n\right)\end{aligned}$$

$$\begin{aligned}\mu_{F^2} &= E[\,F \odot F] \\ &= E[AT \odot AT] \\ &= E[\,(A^bT + A^wT) \odot (A^bT + A^wT)] \\ &= E[A^bT \odot A^bT] + 2E[A^wT \odot A^bT] + E[A^wT \odot A^wT] \\ &= E[A^bT \odot A^bT] + 2E[A^wT] \odot E[A^bT] + E[T \odot A^w1_n \odot A^w1_n] \\ &= diag\{A^bE[TT']A^{b'}\} + 2E[T] \odot A^w1_n \odot A^bE[T] + A^w1_n \odot A^w1_n \odot E[T] \\ &= diag\{A^b(E[T]E[T]' \odot (1_{n\times n} - D) + E[T]1_n' \odot D)A^{b'}\} + 2E[T] \odot A^w1_n \odot A^bE[T] + A^w1_n \odot A^w1_n \odot E[T]\end{aligned}$$

The following are implied by community-randomization:

Green Text: The treatment status of the ego is independent of the treatment status of alters who live in communities other than the ego's.

Orange Text: The treatment status of the ego and her within-community alters is independent of the treatment status of alters in communities other than the ego's.

Blue Text: The treatment status of alters within the same community is identical to the treatment status of the ego.

Purple Text: The matrix $E[TT']$ has entries $E[T_i]E[T_j]$ for i,j in different communities and $E[T_i^2] = E[T_i]$ for i,j in the same community. This is because $T_i \perp T_j$ for i,j in different communities, $T_i = T_j$ for i,j in the same community, and $T_i = T_i^2$ since $T_i \in \{0,1\}$

B: Household wealth and physical function

Causal Estimands and Causal Identification

Given a graph $\mathcal{G}$ with n nodes, we define a $n \times 1$ vector of dichotomous potential outcomes $\boldsymbol{Y}(\boldsymbol{a})$ under $h \times 1$ treatment vector $\boldsymbol{a}$. That is, the i^{th} entry of vector $\boldsymbol{Y}(\boldsymbol{a})$ represents the physical function score individual i would obtain if households were forced to take treatment as dictated by $\boldsymbol{a}$**.** The i^{th} entry of $\boldsymbol{a}$ represents the value of the wealth index imposed on household i.

We assume conditional exchangeability holds:

$$\boldsymbol{Y}(\boldsymbol{a}) \perp \boldsymbol{A}|\boldsymbol{L}$$

where $\boldsymbol{A}$ is a $h \times 1$ vector of observed wealth and $\boldsymbol{L}$ is a $n \times k$ matrix of covariates.

We also assume consistency holds:

$$\boldsymbol{Y} = \boldsymbol{Y}(\boldsymbol{A})$$

We are interested in $E[\boldsymbol{Y}(\boldsymbol{a})|\boldsymbol{L}]$ **—** the expected potential outcome for physical function under treatment $\boldsymbol{a}$ for a set of individuals with exactly the same values of covariates $\boldsymbol{L}$ as were observed.

Using conditional exchangeability and consistency:

$$\begin{aligned} E[\boldsymbol{Y}(\boldsymbol{a})|\boldsymbol{L}] &= E[\boldsymbol{Y}(\boldsymbol{a})|\boldsymbol{L}, \boldsymbol{A} = \boldsymbol{a}] \\ &= E[\boldsymbol{Y}|\boldsymbol{L}, \boldsymbol{A} = \boldsymbol{a}] \end{aligned}$$

We are interested in the change in entries of $E[\boldsymbol{Y}(\boldsymbol{a})|\boldsymbol{L}]$ produced by a unit increase in each entry of $\boldsymbol{a}$:

$$\frac{d}{d\mathbf{a}}E[\boldsymbol{Y}(\boldsymbol{a})|\boldsymbol{L}] = \frac{d}{d\mathbf{a}}E[\boldsymbol{Y}|\boldsymbol{L}, \boldsymbol{A} = \boldsymbol{a}]$$

with i, j^{th}entry of the form:

$$\frac{d}{d\mathrm{a}_j}E[Y_i(\boldsymbol{a})|\boldsymbol{L}].$$

We calculate three causal estimands derived from this $n \times h$ matrix — the average direct effect (ADE), average spillover effect (ASE), and average total effect (ATE). These are defined as follows:

$$ADE \equiv \frac{1}{n}\sum_{j=1}^{h}\sum_{i=1}^{n}\frac{d}{da_j}\mathrm{E}[Y_i|\boldsymbol{L}, \boldsymbol{A} = \boldsymbol{a}]\,\big|_{\mathbf{a}=\mathbf{A}} \times I_{i\in\mathcal{H}_j}(i)$$

$$ASE \equiv \frac{1}{\mathrm{n}}\sum_{j=1}^{h}\sum_{i=1}^{n}\frac{d}{da_j}\mathrm{E}[Y_i|\boldsymbol{L}, \boldsymbol{A} = \boldsymbol{a}]\,\big|_{\mathbf{a}=\mathbf{A}} \times I_{i\notin\mathcal{H}_j}(i)$$

$$ATE \equiv \frac{1}{n}\sum_{j=1}^{h}\sum_{i=1}^{n}\frac{d}{da_j}\mathrm{E}[Y_i|\boldsymbol{L}, \boldsymbol{A} = \boldsymbol{a}]\,\big|_{\mathbf{a}=\mathbf{A}}$$

Where $\mathcal{H}_j$ is the set of indices labelling individuals who are members of household j, $I_{i\in\mathcal{H}_j}(i)$ is an indicator function that obtains when $i \in \mathcal{H}_j$, and $I_{i\notin\mathcal{H}_j}(i)$ is an indicator function that obtains when $i \notin \mathcal{H}_j$.

In words, ADE is the average change in an individual's physical function that results from a 1-standard deviation increase in that person's household wealth index. ASE is the average change in an individual's physical function that results from a 1-standard deviation increase in every other household's wealth index. The ATE is the sum of ADE and ASE — it is the total average change in individuals' physical function caused by a 1-standard deviation increase in every household's wealth index.

Statistical Model

To calculate the matrix $\frac{d}{d\mathbf{a}}E[\boldsymbol{Y}|\boldsymbol{L},\boldsymbol{A}=\boldsymbol{a}]$, it is necessary to make assumptions about the distribution of $(\boldsymbol{L},\boldsymbol{A},\boldsymbol{Y})$. We assume that $\boldsymbol{Y}|\boldsymbol{L},\boldsymbol{A}$ arises from a conditional Markov random field defined by the following local Markov property [8]:

$$Y_i \perp Y_j |\, \boldsymbol{L}_{\mathcal{N}_i}, \boldsymbol{A}_{\Gamma_i}\boldsymbol{Y}_{\mathcal{N}_i} \quad \text{where } j \notin \mathcal{N}_i.$$

$\mathcal{N}_i$ is a set containing the indices of individuals connected to individual i in the graph $\mathcal{G}$. $\boldsymbol{L}_{\mathcal{N}_i}$ is the sub-vector of $\boldsymbol{L}$ corresponding to entries whose indices are contained in $\mathcal{N}_i$. $\boldsymbol{Y}_{\mathcal{N}_i}$ is similarly defined. $\boldsymbol{A}_{\Gamma_i}$ is the sub-vector of $\boldsymbol{A}$ corresponding to entries whose indices are contained in Γ_i. Γ_i is a set containing the indices of all households j such that $i \notin \mathcal{H}_j$ and $\mathcal{H}_j \cap \mathcal{N}_i \neq \emptyset$. In words, Γ_i contains all the households that i's alters are members of and excludes the household that i is a member of.

In addition, we make an assumption about the conditional mean of $\boldsymbol{Y}$. For ease of exposition, we show the case where $\boldsymbol{L}$ is a $n \times 1$ matrix (as opposed to a $n \times k$ matrix):

$$E[Y_i|\boldsymbol{L}=\boldsymbol{l},\boldsymbol{A}=\boldsymbol{a},\boldsymbol{Y}_{-i}] = \beta_0 + \beta_1 a_{j:i\in\mathcal{H}_j} + \beta_2 l_i + \beta_3 \sum_{j\in\Gamma_i} a_j + \beta_4 \sum_{j\in\mathcal{N}_i} l_j + \phi \sum_{j\in\mathcal{N}_i^{40+}} y_j$$

$\mathcal{N}_i^{40+} \subseteq \mathcal{N}_i$ is the set of family members of person i who are aged 40 or older. For the remainder of this appendix, we name the above equation the *conditional mean model*. Through a slight abuse of notation, we can re-write this as:

$$E[\boldsymbol{Y}|\boldsymbol{L},\boldsymbol{A},\boldsymbol{Y}] = \beta_0\mathbf{1} + \beta_1\boldsymbol{HA} + \beta_2\boldsymbol{L} + \beta_3\boldsymbol{HNA} + \beta_4\boldsymbol{ML} + \phi\boldsymbol{M}_{\boldsymbol{40+}}\boldsymbol{Y}$$

Where $\boldsymbol{M}$ is the $n \times n$ adjacency matrix associated with graph $\mathcal{G}$, $\boldsymbol{M}_{\boldsymbol{40+}}$ is the adjacency matrix for the induced subnetwork of the family network $\mathcal{G}$ consisting only of individuals over 40 years of age, $\boldsymbol{H}$ is an $n \times h$ household membership matrix such that the ij^{th} entry is 1 if individual i is a member of household j. $\boldsymbol{N}$ is the $h \times h$ matrix associated with graph $\mathcal{G}^H$ — the coarsening of graph $\mathcal{G}$ using household membership matrix $\boldsymbol{H}$, and $\odot$ represents element-wise multiplication. Note that $\boldsymbol{M}$ and $\boldsymbol{N}$ have zeros on the main diagonal. $\mathbf{1}$ is a $n \times 1$ vector of $1's$ and $\boldsymbol{I}$ is the $n \times n$ identity matrix.

To coarsen graph $\mathcal{G}$ using household membership matrix $\boldsymbol{H}$, we create a new graph $\mathcal{G}^H$ such that a pair of nodes (i,j) in $\mathcal{G}^H$ has a tie if and only if there is a pair of nodes (a,b) in $\mathcal{G}$ such that a is a member of household i and b is a member of household j and (a,b) has a tie in $\mathcal{G}$.

Now:

$$E[E[\boldsymbol{Y}|\boldsymbol{L},\boldsymbol{A},\boldsymbol{Y}]|\boldsymbol{A},\boldsymbol{L}] = \beta_0\mathbf{1} + \beta_1\boldsymbol{HA} + \beta_2\boldsymbol{L} + \beta_3\boldsymbol{HNA} + \beta_4\boldsymbol{ML} + \phi\boldsymbol{M}_{\boldsymbol{40+}}\boldsymbol{E}[\boldsymbol{Y}|\boldsymbol{A},\boldsymbol{L}]$$

$$E[\boldsymbol{Y}|\boldsymbol{A},\boldsymbol{L}] = (\boldsymbol{I} - \phi\boldsymbol{M}_{\boldsymbol{40+}})^{-1}(\beta_1\boldsymbol{H} + \beta_3\boldsymbol{HN})\boldsymbol{A} + \cdots$$

We can write an $n \times h$ Jacobian matrix:

$$\frac{d}{d\boldsymbol{a}} E[\boldsymbol{Y}|\boldsymbol{A} = \boldsymbol{a}, \boldsymbol{L}] = (\boldsymbol{I} - \phi \boldsymbol{M_{40+}})^{-1}(\beta_1 \boldsymbol{H} + \beta_3 \boldsymbol{HN})$$

and using this matrix, calculate ADE, ASE and ATE as:

$$ATE = \frac{1}{n} \text{grandsum}\{(\boldsymbol{I} - \phi \boldsymbol{M_{40+}})^{-1}(\beta_1 \boldsymbol{H} + \beta_3 \boldsymbol{HN})\}$$

$$ADE = \frac{1}{\text{n}} \text{trace}\{\boldsymbol{H}'(\boldsymbol{I} - \phi \boldsymbol{M_{40+}})^{-1}(\beta_1 \boldsymbol{H} + \beta_3 \boldsymbol{HN})\}$$

$$ASE = ATE - \text{ADE}$$

Where $\text{grandsum}\{\boldsymbol{A}\}$ is the sum of all the entries in matrix $\boldsymbol{A}$.

Multiple Imputation using Chained Equations

We imputed missing values of household wealth and physical function using a custom-built R package based on the multiple imputation using chained equations (MICE) algorithm (Buuren & Groothuis-Oudshoorn, 2011).

In each iteration of this algorithm, we imputed missing household wealth values using household-level measures as predictors and calculated network wealth values using these imputed values.

Table 0.5 shows the level of missingness in household and network wealth for two sub-networks: the HAALSI network consists of individuals in the HAALSI cohort. The Community network consists of individuals in HAALSI as well as individuals who are connected to those in HAALSI. In addition to household and network wealth, the following Level 2 variables were used in imputation model (these variables were fully observed):

- For each year:
    - Number of People in Household,
    - Number of Males > 60 years of age,
    - Number of Females > 60 years of age,
    - Number of Males 18-60 years of age,
    - Number of Children
    - Network Degree

Table 0.5: Level-2 Imputation: Households

	Household Attributes		Network Attributes	
	% Missing HAALSI	% Missing Community	% Missing HAALSI	% Missing Community
Wealth 2001	6.6	9.2	31.3	20.7
Wealth 2003	3.9	6.0	29.9	19.1
Wealth 2005	7.0	10.2	30.0	18.8
Wealth 2007	9.6	13.0	31.5	20.1
Wealth 2009	4.4	7.1	28.6	16.8
Wealth 2011	8.3	14.8	29.6	17.1
Wealth 2013	9.0	18.6	30.2	17.2
Wealth 2014	5.7	13.9	28.0	15.2
Wealth 2015	7.7	16.5	27.8	15.4
Wealth 2016	8.9	17.2	27.9	15.4
Wealth 2017	18.5	30.0	33.5	20.5

In the same iteration of the algorithm, we then imputed individual level physical function using individual level measures in addition to household-level measures as predictors, and calculated network physical function. Table 0.6 shows the level of missingness for in household and network physical function for the same sub-networks described above. In addition to Level-2 variables, individual physical function, and network physical function, we used individual network degree to impute physical function.

We iterated these steps 60 times for each imputed dataset, creating 64 such datasets. Dichotomous variables were imputed using a random draw from the Bernoulli distribution with probability of success given by predicted values from a logistic regression. Continuous variables were imputed using predictive mean matching, choosing randomly from the 5 closest predicted values.

Table 0.6: Level-1 Imputation: Individuals

	Individual Attributes		Network Attributes	
	% Missing HAALSI	% Missing Community	% Missing HAALSI	% Missing Community
ADL (W1)	0.2	60.6	46.2	34.5
ADL (W2)	17.8	67.6	54.3	45.2
Grip Strength (W1)	9.6	64.4	48.6	39.2
Grip Strength (W2)	36.4	74.9	58.9	54.2
Gait Speed (W1)	8.1	63.8	48.0	38.5
Gait Speed (W2)	39.3	76.0	59.4	55.9
Education at Baseline	0.3	60.7	•	•
Employment at Baseline	0.3	60.6	•	•
Married at Baseline	0.1	60.6	•	•
Has Children at Baseline	0.1	60.6	•	•
Recieves Pension Income at Baseline	0.0	60.5	•	•
Wealth 2001	6.3	6.8	30.8	27.7
Wealth 2003	3.7	3.8	29.3	26.6
Wealth 2005	6.6	7.0	29.4	26.3
Wealth 2007	9.0	9.5	30.9	28.5
Wealth 2009	4.1	4.3	28.2	25.8
Wealth 2011	7.9	8.9	29.3	26.5
Wealth 2013	8.4	10.8	29.8	27.0
Wealth 2014	5.3	7.6	27.8	25.0
Wealth 2015	7.3	9.3	27.5	25.0
Wealth 2016	8.2	10.1	27.7	25.1
Wealth 2017	17.6	20.7	33.5	30.7

Obtaining a Stable Set

To find a sub-graph $\mathcal{G}_{\text{stable}}$ of $\mathcal{G}$ such that no pair of nodes in $\mathcal{G}_{\text{stable}}$ is connected by a tie in $\mathcal{G}$, and such that each node in $\mathcal{G}_{\text{stable}}$ represents a member of the HAALSI cohort, we follow this algorithm:

a) Start with a list of all the nodes in $\mathcal{G}$ that are also members of the HAALSI cohort, and an empty graph $\mathcal{G}_{\text{stable}}$
b) Remove all singletons (degree 0 nodes) from this list and place them in $\mathcal{G}_{\text{stable}}$
c) Randomly remove a node i from the list and place this node in $\mathcal{G}_{\text{stable}}$
d) Discard from the list all nodes all $j \in \mathcal{G}$ such that $j \in \mathcal{N}_i$
e) Repeat c) and d) until there are no more nodes in the list

Additional figures: autocorrelation of wealth over time

Figure 0.1: Autocorrelation of household wealth index over time

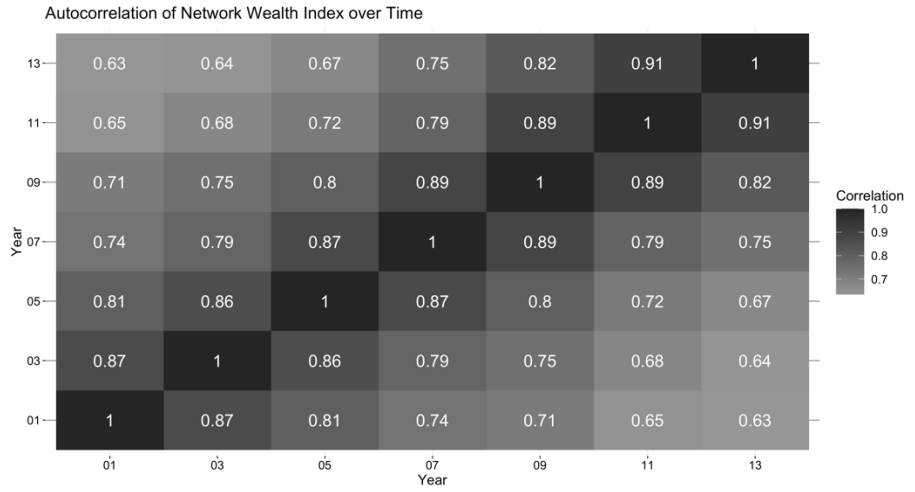

Figure 0.2: Autocorrelation of network wealth index over time

Robustness Check for Household Wealth vs. Gait Speed

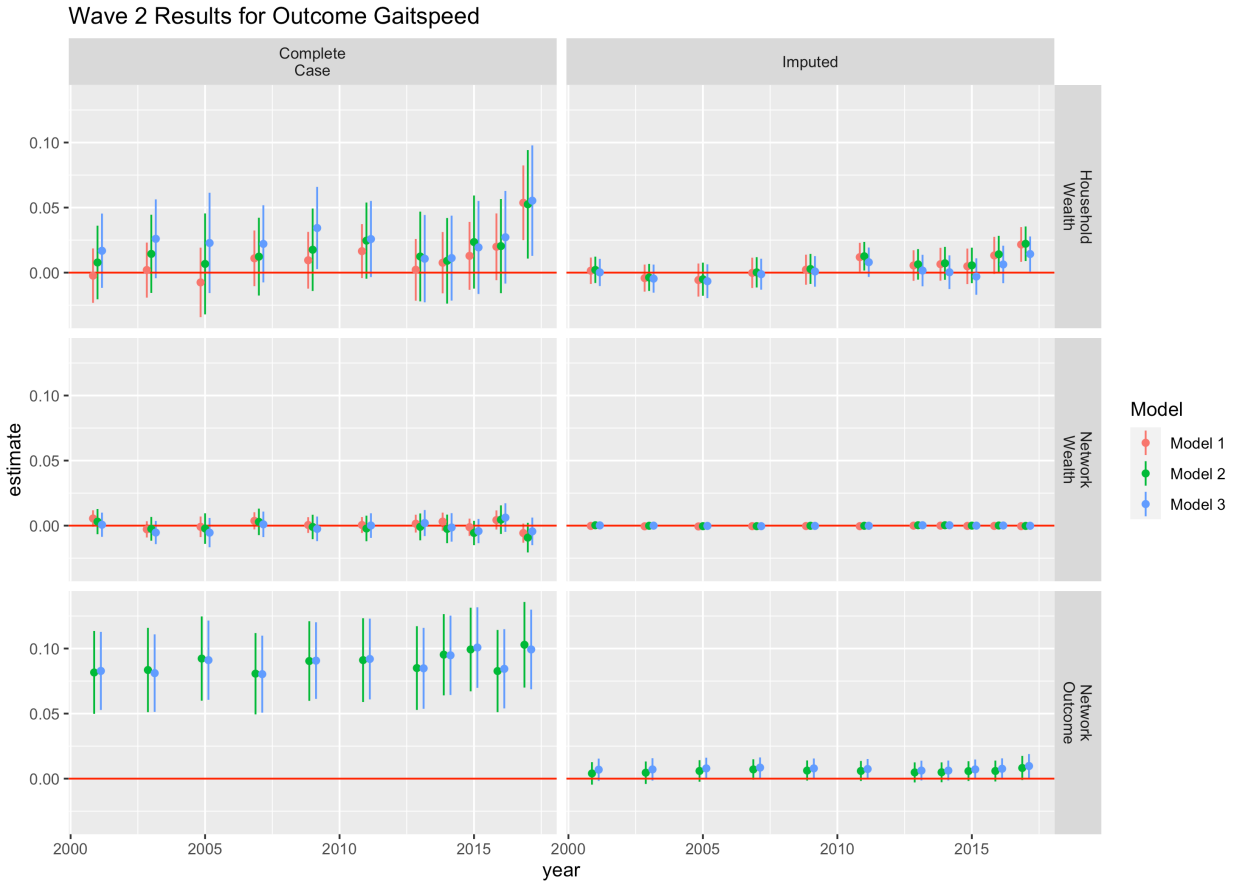

Figure 0.3: Conditional Mean Model for Wave 2 Gait Speed

Model 1 includes Network Wealth and Household Wealth as predictors. Model 2 includes these as well as Network ADL as predictors. Model 3 includes these as predictors and adjusts age and gender as potential confounders.

C: Household wealth and mortality

Notation

Unless otherwise stated, for a given vector $\boldsymbol{V}$:

- $\boldsymbol{V}_{-i}$ is the subvector of $\boldsymbol{V}$ containing all entries except the i^{th} one
- For a given set A, $\boldsymbol{V}_A$ is the subvector of $\boldsymbol{V}$ containing all entries which correspond with indices contained in A
- V_i is the i^{th} entry of $\boldsymbol{V}$
- $\bar{V}$ is the mean of the entries of $\boldsymbol{V}$

$\boldsymbol{\mathcal{G}}$	The graph / network of households.
$\boldsymbol{n}$	The number of nodes in network $\mathcal{G}$
$\boldsymbol{\mathcal{N}_i}$	The set of nodes connected to node i in network $\mathcal{G}$
$\boldsymbol{Y, A, L}$	The n-dimensional vectors of observed deaths (outcome), observed household wealth (exposure), and observed covariates
$\boldsymbol{a}$	The n-dimensional vector of assigned household wealth
$\boldsymbol{Y(a)}$	The n-dimensional vector of counterfactual household deaths under the intervention of assigning household wealth according to vector $\boldsymbol{a}$
$\boldsymbol{M}$	The $n \times n$ symmetric adjacency matrix associated with network $\mathcal{G}$
$\mathbf{1}$	An n-dimensional vector whose entries are all 1
$\boldsymbol{I}$	The $n \times n$ identity matrix
$\boldsymbol{d}$	The n-dimensional vector whose i^{th} entry is the degree of node i in network $\mathcal{G}$
$\odot$	The element-wise matrix multiplication operator
$\mathbf{diag}\{\}$	For a given n-dimensional vector $\boldsymbol{V}$, $\mathrm{diag}\{\boldsymbol{V}\}$ is an $n \times n$ matrix with entries $\boldsymbol{V}$ on the main diagonal and 0's elsewhere.
$\mathbf{grandsum}\{\}$	$\mathrm{grandsum}\{\mathbf{D}\}$ is the sum of the entries of matrix $\boldsymbol{D}$
$\mathbf{trace}\{\}$	$\mathrm{trace}\{\mathbf{D}\}$ is the sum of the entries on the main diagonal of matrix $\boldsymbol{D}$
$\lvert * \rvert$	For a given set S, $\lvert S \rvert$ is the cardinality of S

Counterfactual Outcome

Given a graph $\mathcal{G}$ with n nodes, we define a $n \times 1$ vector of dichotomous potential outcomes $\boldsymbol{Y(a)}$ under $n \times 1$ treatment vector $\boldsymbol{a}$. That is, the i^{th} entry of vector $\boldsymbol{Y(a)}$ represents whether or not household i would experience the outcome of household death if every node in the network

were forced to have household wealth as dictated by $\boldsymbol{a}$**.** The i^{th} entry of $\boldsymbol{a}$ represents the value of household wealth imposed on household i.

We assume conditional exchangeability holds:

$$\boldsymbol{Y}(\boldsymbol{a}) \perp \boldsymbol{A}|\boldsymbol{L}$$

where $\boldsymbol{A}$ is a $n \times 1$ vector of observed wealth and $\boldsymbol{L}$ is a $n \times k$ matrix of covariates. We also assume consistency holds:

$$\boldsymbol{Y} = \boldsymbol{Y}(\boldsymbol{A})$$

We are interested in $E[\boldsymbol{Y}(\boldsymbol{a})|\boldsymbol{L}]$ **—** the expected potential outcome for household death under treatment $\boldsymbol{a}$ for a set of households with exactly the same values of covariates $\boldsymbol{L}$ as were observed. i.e. We are not interested in $E[\boldsymbol{Y}(\boldsymbol{a})]$ - the expected potential outcome of an intervention among a notional set of networks of households whose covariates are drawn randomly from the support of $\boldsymbol{L}$.

Using conditional exchangeability and consistency:

$$E[\boldsymbol{Y}(\boldsymbol{a})|\boldsymbol{L}] = E[\boldsymbol{Y}(\boldsymbol{a})|\boldsymbol{L}, \boldsymbol{A} = \boldsymbol{a}]$$
$$= E[\boldsymbol{Y}|\boldsymbol{L}, \boldsymbol{A} = \boldsymbol{a}]$$

Causal Estimands

We are interested in the change in entries of $E[\boldsymbol{Y}(\boldsymbol{a})|\boldsymbol{L}]$ produced by a unit increase in each entry of $\boldsymbol{a}$:

$$\frac{d}{d\mathbf{a}} E[\boldsymbol{Y}(\boldsymbol{a})|\boldsymbol{L}] = \frac{d}{d\mathbf{a}} E[\boldsymbol{Y}|\boldsymbol{L}, \boldsymbol{A} = \boldsymbol{a}]$$

with i, j^{th}entry of the form:

$$\frac{d}{d\mathrm{a}_j} E[Y_i(\boldsymbol{a})|\boldsymbol{L}].$$

We calculate three causal estimands derived from this $n \times n$ matrix — the average direct effect (ADE), average spillover effect (ASE), and average total effect (ATE). These are defined as follows:

$$ADE \equiv \frac{1}{n}\sum_i \frac{d}{da_i} \mathrm{E}[Y_i|\boldsymbol{L}, \boldsymbol{A} = \boldsymbol{a}] \Big|_{\mathbf{a=A}}$$

$$ASE \equiv \frac{1}{n}\sum_i \sum_{j \neq i} \frac{d}{da_j} \mathrm{E}[Y_i|\boldsymbol{L}, \boldsymbol{A} = \boldsymbol{a}] \Big|_{\mathbf{a=A}}$$

$$ATE \equiv \frac{1}{n}\sum_i \sum_j \frac{d}{da_j} \mathrm{E}[Y_i|\boldsymbol{L}, \boldsymbol{A} = \boldsymbol{a}] \Big|_{\mathbf{a=A}}$$

In words, ADE is the average change in a household's likelihood of experiencing death as a result of a 1-standard deviation increase in that household's wealth. ASE is the average change in a household's likelihood of experiencing death as a result of a 1-standard deviation increase in every other household's wealth. The ATE is the sum of ADE and ASE — it is the total average change in household death caused by a 1-standard deviation increase in every household's wealth.

Statistical Model

To calculate the matrix $\frac{d}{d\mathbf{a}} E[\boldsymbol{Y}|\boldsymbol{L}, \boldsymbol{A} = \boldsymbol{a}]$, it is necessary to make assumptions about the joint distribution of $(\boldsymbol{L}, \boldsymbol{A}, \boldsymbol{Y})$. We assume that $\boldsymbol{Y}|\boldsymbol{L}, \boldsymbol{A}$ arises from a conditional Markov random field defined by the following local Markov property (Besag, 1974; Tchetgen Tchetgen et al., 2020):

$$Y_i \perp Y_j | \, \boldsymbol{L}_{\mathcal{N}_i}, \boldsymbol{A}_{\mathcal{N}_i} \boldsymbol{Y}_{\mathcal{N}_i} \quad \text{where } j \notin \mathcal{N}_i.$$

$\mathcal{N}_i$ is a set containing the indices of households connected to household i in the network $\mathcal{G}$. $\boldsymbol{L}_{\mathcal{N}_i}$ is the sub-vector of $\boldsymbol{L}$ corresponding to entries whose indices are contained in $\mathcal{N}_i$. $\boldsymbol{A}_{\mathcal{N}_i}$ and $\boldsymbol{Y}_{\mathcal{N}_i}$ are similarly defined.

In addition, we make an assumption about the conditional mean of $\boldsymbol{Y}$. For ease of exposition, we show the case where $\boldsymbol{L}$ is a $n \times 1$ matrix (as opposed to a $n \times k$ matrix):

$$\begin{aligned} &E[Y_i|\boldsymbol{L} = \boldsymbol{l}, \boldsymbol{A} = \boldsymbol{a}, \boldsymbol{Y}_{-i}] \\ &\quad = \beta_0 + \beta_1 a_i + \beta_2 l_i + \beta_3 \sum_{j \in \mathcal{N}_i} a_j + \beta_4 \sum_{j \in \mathcal{N}_i} l_j + \beta_5 a_i l_i + \beta_6 l_i \sum_{j \in \mathcal{N}_i} a_j \\ &\quad + \beta_7 l_i \sum_{j \in \mathcal{N}_i} l_j + \phi \sum_{j \in \mathcal{N}_i} y_j \end{aligned}$$

For the remainder of this appendix, we name the above equation the *conditional mean model*.

Through a slight abuse of notation, we can re-write this as:

$$E[\boldsymbol{Y}|\boldsymbol{L}, \boldsymbol{A}, \boldsymbol{Y}] = \beta_0 \mathbf{1} + \beta_1 \boldsymbol{A} + \beta_2 \boldsymbol{L} + \beta_3 \boldsymbol{MA} + \beta_4 \boldsymbol{ML} + \beta_5 \boldsymbol{A} \odot \boldsymbol{L} + \beta_6 \boldsymbol{L} \odot \boldsymbol{MA} + \beta_7 \boldsymbol{L} \odot \boldsymbol{ML} + \phi \boldsymbol{MY}$$

Where $\boldsymbol{M}$ is the $n \times n$ adjacency matrix associated with graph $\mathcal{G}$, and $\odot$ represents element-wise multiplication. Note that $\boldsymbol{M}$ has zeros on the main diagonal. $\mathbf{1}$ is a $n \times 1$ vector of $1's$ and $\boldsymbol{I}$ is the $n \times n$ identity matrix.

Now:

$$E[E[\boldsymbol{Y}|\boldsymbol{L},\boldsymbol{A},\boldsymbol{Y}]|\boldsymbol{A},\boldsymbol{L}]$$
$$= \beta_0\mathbf{1} + \beta_1\boldsymbol{A} + \beta_2\boldsymbol{L} + \beta_3\boldsymbol{M}\boldsymbol{A} + \beta_4\boldsymbol{M}\boldsymbol{L} + \beta_5\boldsymbol{L} \odot \boldsymbol{A} + \beta_6\boldsymbol{L} \odot \boldsymbol{M}\boldsymbol{A} + \beta_7\boldsymbol{L} \odot \boldsymbol{M}\boldsymbol{L}$$
$$+ \phi\boldsymbol{M}\boldsymbol{E}[\boldsymbol{Y}|\boldsymbol{A},\boldsymbol{L}]$$

$$E[\boldsymbol{Y}|\boldsymbol{A},\boldsymbol{L}] = (\boldsymbol{I} - \phi\boldsymbol{M})^{-1}(\beta_1\boldsymbol{I} + \beta_3\boldsymbol{M} + \beta_5\mathrm{diag}\{\boldsymbol{L}\} + \beta_6\mathrm{diag}\{\boldsymbol{L}\}\boldsymbol{M})\boldsymbol{A} + \cdots$$

where $\mathrm{diag}\{\boldsymbol{L}\}$ is an $n \times n$ matrix with $\boldsymbol{L}$ on the main diagonal and zeros elsewhere**.** We can write an $n \times n$ Jacobian matrix:

$$\frac{d}{d\boldsymbol{a}}E[\boldsymbol{Y}|\boldsymbol{A} = \boldsymbol{a},\boldsymbol{L}] = (\boldsymbol{I} - \phi\boldsymbol{M})^{-1}(\beta_1\boldsymbol{I} + \beta_3\boldsymbol{M} + \beta_5\mathrm{diag}\{\boldsymbol{L}\} + \beta_6\mathrm{diag}\{\boldsymbol{L}\}\boldsymbol{M})$$

and using this matrix, calculate ADE and ASE as:

$$ATE = \frac{1}{n}\mathrm{grandsum}\left\{\frac{d}{d\boldsymbol{a}}E[\boldsymbol{Y}|\boldsymbol{A} = \boldsymbol{a},\boldsymbol{L}]\right\}$$

$$ADE = \frac{1}{\mathrm{n}}\mathrm{trace}\left\{\frac{d}{d\boldsymbol{a}}E[\boldsymbol{Y}|\boldsymbol{A} = \boldsymbol{a},\boldsymbol{L}]\right\}$$

$$ASE = ATE - \text{ADE}$$

Where $\text{grandsum}\{\boldsymbol{A}\}$ is the sum of all the entries in matrix $\boldsymbol{A}$.

We note that when

$$Y_i \perp Y_j | \, \boldsymbol{L}_{\mathcal{N}_i}, \boldsymbol{A}_{\mathcal{N}_i} \text{ for all nodes } i, j \in \mathcal{G}$$

we can use a regression model to estimate the parameters of the conditional mean model.

Furthermore,

$$\begin{aligned} ADE &= \frac{1}{n}\text{trace}\{\beta_1 \boldsymbol{I} + \beta_3 \boldsymbol{M} + \beta_5 \text{diag}\{\boldsymbol{l}\} + \beta_6 \text{diag}\{\boldsymbol{l}\}\boldsymbol{M}\} \\ &= \frac{1}{n}\left(\beta_1 n + \beta_5 \sum_i l_i\right) \\ &= \beta_1 + \beta_5 \bar{l} \end{aligned}$$

$$\begin{aligned} ATE &= \frac{1}{n}\text{grandsum}\{\beta_1 \boldsymbol{I} + \beta_3 \boldsymbol{M} + \beta_5 \text{diag}\{\boldsymbol{l}\} + \beta_6 \text{diag}\{\boldsymbol{l}\}\boldsymbol{M}\} \\ &= \frac{1}{n}\left(\beta_1 n + \beta_3 n\bar{d} + \beta_5 n\bar{l} + \beta_6 \sum_i \sum_{j \in N_i} l_i\right) \\ &= \beta_1 + \beta_3 \bar{d} + \beta_5 \bar{l} + \beta_6 \frac{1}{n}\sum_i l_i d_i \end{aligned}$$

$$ASE = \beta_3 \bar{d} + \beta_6 \frac{1}{n}\sum_i l_i d_i$$

where d_i is the degree of node i, $\bar{d}$ is the average degree, and $\bar{l}$ is the average of l_i.

Additional Tables and Figures

Table 0.7: Descriptive statistics for Households in AHDSS

	2009 (N=16275)	2010 (N=16629)	2011 (N=17047)	2012 (N=17539)	2013 (N=21130)
# Connected Households					
0	4769 (29.3%)	4807 (28.9%)	4884 (28.7%)	5097 (29.1%)	7868 (37.2%)
1	3372 (20.7%)	3381 (20.3%)	3455 (20.3%)	3521 (20.1%)	3989 (18.9%)
2	2734 (16.8%)	2832 (17.0%)	2917 (17.1%)	2908 (16.6%)	3055 (14.5%)
3	1943 (11.9%)	2038 (12.3%)	2066 (12.1%)	2164 (12.3%)	2193 (10.4%)
4	1332 (8.2%)	1361 (8.2%)	1416 (8.3%)	1441 (8.2%)	1489 (7.0%)
5	799 (4.9%)	851 (5.1%)	906 (5.3%)	928 (5.3%)	971 (4.6%)
6+	1326 (8.1%)	1359 (8.2%)	1403 (8.2%)	1480 (8.4%)	1565 (7.4%)
# Household Members					
0	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
1	1647 (10.1%)	1697 (10.2%)	1788 (10.5%)	1952 (11.1%)	2336 (11.1%)
2-5	7380 (45.3%)	7655 (46.0%)	7958 (46.7%)	8162 (46.5%)	9844 (46.6%)
6-10	5751 (35.3%)	5790 (34.8%)	5805 (34.1%)	5933 (33.8%)	7113 (33.7%)
11-20	1458 (9.0%)	1442 (8.7%)	1461 (8.6%)	1461 (8.3%)	1794 (8.5%)
21+	39 (0.2%)	45 (0.3%)	35 (0.2%)	31 (0.2%)	43 (0.2%)
# Children (<5)					
0	3487 (21.4%)	3630 (21.8%)	3785 (22.2%)	4056 (23.1%)	4915 (23.3%)
1	2751 (16.9%)	2874 (17.3%)	3017 (17.7%)	3167 (18.1%)	3871 (18.3%)
2-5	8938 (54.9%)	9042 (54.4%)	9238 (54.2%)	9310 (53.1%)	11146 (52.7%)
6-10	1041 (6.4%)	1020 (6.1%)	961 (5.6%)	964 (5.5%)	1152 (5.5%)
11-20	57 (0.4%)	62 (0.4%)	45 (0.3%)	40 (0.2%)	45 (0.2%)
21+	1 (0.0%)	1 (0.0%)	1 (0.0%)	2 (0.0%)	1 (0.0%)
# Adults (15-59)					
0	551 (3.4%)	547 (3.3%)	565 (3.3%)	573 (3.3%)	691 (3.3%)
1	2993 (18.4%)	3050 (18.3%)	3151 (18.5%)	3276 (18.7%)	3869 (18.3%)
2-5	10815 (66.5%)	11091 (66.7%)	11329 (66.5%)	11583 (66.0%)	13943 (66.0%)
6-10	1851 (11.4%)	1878 (11.3%)	1932 (11.3%)	2049 (11.7%)	2538 (12.0%)
11-20	65 (0.4%)	62 (0.4%)	68 (0.4%)	56 (0.3%)	88 (0.4%)
21+	0 (0%)	1 (0.0%)	2 (0.0%)	2 (0.0%)	1 (0.0%)
# Elders (60+)					
0	12546 (77.1%)	12901 (77.6%)	13270 (77.8%)	13675 (78.0%)	16496 (78.1%)
1	3115 (19.1%)	3095 (18.6%)	3162 (18.5%)	3251 (18.5%)	3916 (18.5%)
2	590 (3.6%)	609 (3.7%)	591 (3.5%)	584 (3.3%)	684 (3.2%)
3+	24 (0.1%)	24 (0.1%)	24 (0.1%)	29 (0.2%)	34 (0.2%)
Household Wealth Index					
Mean (SD)	2.51 (0.464)	2.51 (0.460)	2.61 (0.470)	2.61 (0.466)	2.64 (0.447)
Median [Min, Max]	2.56 [0.902, 4.04]	2.56 [0.902, 4.04]	2.65 [0.825, 4.10]	2.66 [0.905, 4.10]	2.68 [0.925, 4.17]
Missing	1498 (9.2%)	1517 (9.1%)	3557 (20.9%)	3539 (20.2%)	5497 (26.0%)

Descriptive statistics for Households in AHDSS (continued)

	2014 (N=21527)	2015 (N=21808)	2016 (N=21948)	2017 (N=21788)	2018 (N=21291)
# Connected Households					
0	7840 (36.4%)	7803 (35.8%)	7725 (35.2%)	7649 (35.1%)	7645 (35.9%)
1	4119 (19.1%)	4215 (19.3%)	4258 (19.4%)	4343 (19.9%)	4364 (20.5%)
2	3107 (14.4%)	3192 (14.6%)	3301 (15.0%)	3283 (15.1%)	3190 (15.0%)
3	2245 (10.4%)	2299 (10.5%)	2368 (10.8%)	2333 (10.7%)	2254 (10.6%)
4	1560 (7.2%)	1625 (7.5%)	1599 (7.3%)	1577 (7.2%)	1491 (7.0%)
5	1045 (4.9%)	1023 (4.7%)	1035 (4.7%)	1007 (4.6%)	947 (4.4%)
6+	1611 (7.5%)	1651 (7.6%)	1662 (7.6%)	1596 (7.3%)	1400 (6.6%)
# Household Members					
0	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
1	2384 (11.1%)	2424 (11.1%)	2504 (11.4%)	2547 (11.7%)	2513 (11.8%)
2-5	10075 (46.8%)	10278 (47.1%)	10383 (47.3%)	10295 (47.3%)	10053 (47.2%)
6-10	7200 (33.4%)	7247 (33.2%)	7213 (32.9%)	7155 (32.8%)	7038 (33.1%)
11-20	1834 (8.5%)	1823 (8.4%)	1815 (8.3%)	1753 (8.0%)	1657 (7.8%)
21+	34 (0.2%)	36 (0.2%)	33 (0.2%)	38 (0.2%)	30 (0.1%)
# Children (<5)					
0	5012 (23.3%)	5162 (23.7%)	5321 (24.2%)	5498 (25.2%)	5651 (26.5%)
1	3959 (18.4%)	4015 (18.4%)	4079 (18.6%)	4011 (18.4%)	4033 (18.9%)
2-5	11375 (52.8%)	11496 (52.7%)	11470 (52.3%)	11278 (51.8%)	10744 (50.5%)
6-10	1136 (5.3%)	1096 (5.0%)	1045 (4.8%)	968 (4.4%)	838 (3.9%)
11-20	44 (0.2%)	38 (0.2%)	32 (0.1%)	32 (0.1%)	25 (0.1%)
21+	1 (0.0%)	1 (0.0%)	1 (0.0%)	1 (0.0%)	0 (0%)
# Adults (15-59)					
0	702 (3.3%)	713 (3.3%)	715 (3.3%)	696 (3.2%)	680 (3.2%)
1	3907 (18.1%)	3929 (18.0%)	3958 (18.0%)	3923 (18.0%)	3701 (17.4%)
2-5	14245 (66.2%)	14554 (66.7%)	14626 (66.6%)	14535 (66.7%)	14318 (67.2%)
6-10	2585 (12.0%)	2524 (11.6%)	2549 (11.6%)	2527 (11.6%)	2506 (11.8%)
11-20	88 (0.4%)	88 (0.4%)	99 (0.5%)	105 (0.5%)	84 (0.4%)
21+	0 (0%)	0 (0%)	1 (0.0%)	2 (0.0%)	2 (0.0%)
# Elders (60+)					
0	16795 (78.0%)	16892 (77.5%)	17014 (77.5%)	16708 (76.7%)	16122 (75.7%)
1	3999 (18.6%)	4163 (19.1%)	4221 (19.2%)	4349 (20.0%)	4413 (20.7%)
2	705 (3.3%)	724 (3.3%)	689 (3.1%)	710 (3.3%)	735 (3.5%)
3+	28 (0.1%)	29 (0.1%)	24 (0.1%)	21 (0.1%)	21 (0.1%)
Household Wealth Index					
Mean (SD)	2.65 (0.417)	2.67 (0.398)	2.68 (0.381)	2.67 (0.397)	2.68 (0.387)
Median [Min, Max]	2.69 [0.902, 4.02]	2.70 [0.921, 4.00]	2.70 [0.902, 4.31]	2.69 [0.861, 3.98]	2.71 [0.921, 4.18]
Missing	4340 (20.2%)	4839 (22.2%)	5089 (23.2%)	7335 (33.7%)	4761 (22.4%)

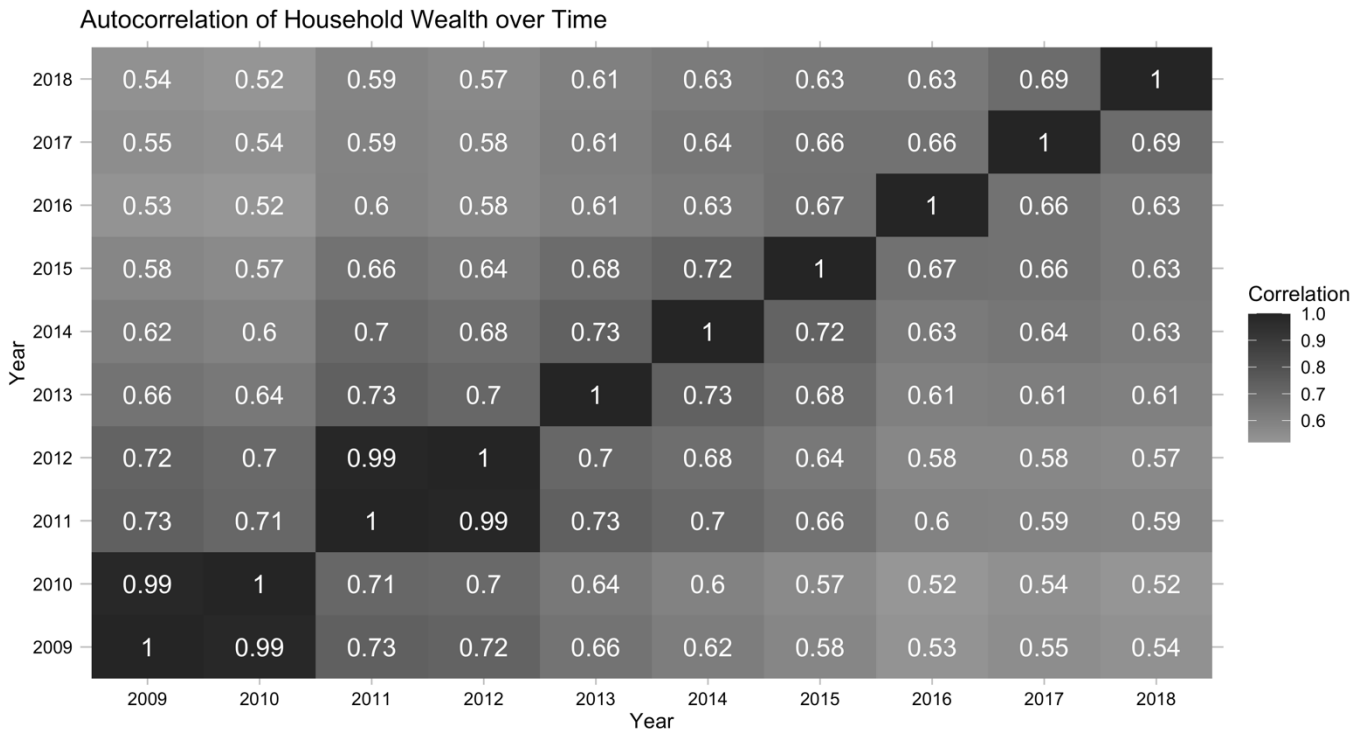

Figure 0.4: Autocorrelation of Household Wealth over Time

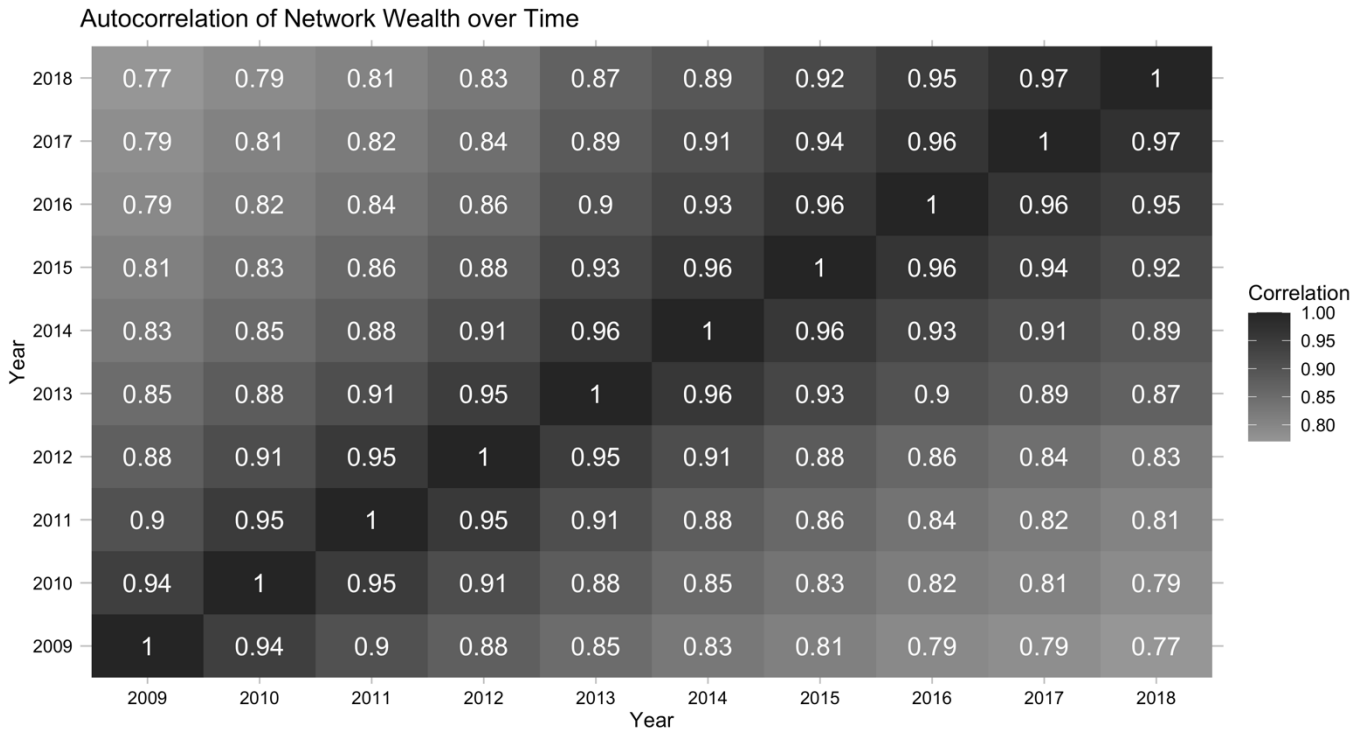

Figure 0.5: Autocorrelation of Network Wealth over Time

Table 0.8: Outcome Regression for Household Deaths 2009-2018 (Complete Case Analysis)

Outcome Regression for Death 2009-2018 (Complete Case Analysis)

	Estimate	95% CI	p	Estimate	95% CI	p	Estimate	95% CI	p	Estimate	95% CI	p
0-year Lag												
SES	-0.34%	-0.44% : -0.23%	0									
Network SES	-0.04%	-0.09% : 0.02%	0.212									
Network Deaths				0.02%	-0.42% : 0.46%	0.922	-0.03%	-0.49% : 0.44%	0.915	-0.03%	-0.51% : 0.45%	0.915
1-Year Lag												
SES				-0.53%	-0.68% : -0.38%	0	-0.30%	-0.53% : -0.07%	0.011	-0.25%	-0.50% : -0.01%	0.045
Network SES				-0.06%	-0.10% : -0.03%	0.001	-0.04%	-0.10% : 0.02%	0.21	-0.03%	-0.10% : 0.03%	0.28
Network Deaths				0.15%	-0.28% : 0.58%	0.501	0.13%	-0.33% : 0.58%	0.583	0.16%	-0.32% : 0.63%	0.515
Deaths				-0.74%	-1.43% : -0.05%	0.037	-0.72%	-1.44% : 0.00%	0.052	-0.67%	-1.42% : 0.08%	0.081
2-Year Lag												
SES							-0.29%	-0.51% : -0.06%	0.013	-0.01%	-0.30% : 0.29%	0.958
Network SES							-0.04%	-0.09% : 0.02%	0.166	-0.03%	-0.10% : 0.04%	0.45
Network Deaths							0.39%	-0.06% : 0.83%	0.087	0.45%	-0.01% : 0.91%	0.057
Deaths				0.77%	0.10% : 1.43%	0.024	0.61%	-0.16% : 1.39%	0.122	0.54%	-0.26% : 1.33%	0.189
3-Year Lag												
SES										-0.35%	-0.61% : -0.09%	0.008
Network SES										-0.01%	-0.06% : 0.05%	0.799
Network Deaths										0.36%	-0.09% : 0.82%	0.118
Deaths				0.36%	-0.27% : 1.00%	0.262	0.33%	-0.34% : 0.99%	0.337	0.08%	-0.67% : 0.84%	0.825
4-Year Lag												
SES												
Network SES												
Network Deaths												
Deaths				0.23%	-0.39% : 0.86%	0.461	0.36%	-0.30% : 1.02%	0.286	0.28%	-0.40% : 0.95%	0.425

**adjusting for village, # members in household , # women in household, # children in household, # elders in household, # members of connected households, # women in connected households, # children in connected households, # elders in connected households, number of connected household

We show results from four regression models of household death on lagged values of wealth, network wealth, deaths and network deaths. We found that contemporaneous (i.e. 0-year Lag) values of network deaths are independent of household deaths. As a result, we fitted regression models under the assumption that outcomes are conditionally independent in any given year.

We also show that household deaths is not independent of lagged values of household deaths after controlling for covariates. i.e. There is within-household correlation of household death over time. To account for this dependence, we assumed that repeated measures of household death were correlated over time. We conducted the cluster bootstrap in order to calculate confidence intervals.

Table 0.9: Causal Estimands calculated using Multiple Imputation

Causal Estimands calculated at Actual Value of Covariates (Multiple Imputation, Full Graph)

	Average Direct Effect		Average Spillover Effect		Average Total Effect	
	Estimate	95% CI	Estimate	95% CI	Estimate	95% CI
Children	-0.15%	-0.19% : -0.11%	-0.02%	-0.05% : 0.02%	-0.17%	-0.22% : -0.12%
Adults	-0.29%	-0.37% : -0.21%	-0.08%	-0.15% : 0.00%	-0.37%	-0.47% : -0.26%
Elders	-0.09%	-0.14% : -0.03%	0.01%	-0.04% : 0.06%	-0.08%	-0.15% : -0.01%
Overall	-0.53%	-0.63% : -0.42%	-0.08%	-0.18% : 0.01%	-0.61%	-0.74% : -0.48%