Title
The effects of mortality and health inequalities over disparities in political behavior

Author
Rodriguez, Javier M.

Publication Date
2012

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

The effects of mortality and health inequalities over disparities in political behavior

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Political Science

by

JAVIER MAURICIO RODRIGUEZ JR

2012

ABSTRACT OF THE DISSERTATION

The effects of mortality and health inequalities

over disparities in political behavior

by

Javier Mauricio Rodriguez Jr

Doctor of Philosophy in Political Science

University of California, Los Angeles, 2012

Professor David O. Sears, Chair

This dissertation project addresses three observations: the degree to which researchers routinely use data collected only on survivors, ignoring that who is included in samples is pre-determined by non-random mortality; that historical mortality gaps between blacks and whites have generated a cumulative demographic contraction among blacks that constitute a non-trivial vanishing electorate; and that differences in infant mortality rates (IMRs) between blacks and whites strongly correlate with the party that controls the presidency. I used data from several sources including the Centers for Disease Control and Prevention, the Multiple Causes of Death, MIDUS, and the U.S. Census Bureau among others.

Because researchers are always dealing with samples truncated by non-random mortality, paradigms in social research like the relationship between age and participation are partially spurious. I find that the effect that age exerts on participation is mediated by mortality rates linked to differences in SES because, as time progresses, low-SES individuals (i.e., low-level participants) die younger than high-SES individuals (i.e., high-level participants). I also set forth a model for the impact of racial mortality gaps on the total vote lost among blacks (1970-2004). Because the dead cannot vote or voice opinions, excess deaths exacerbate the imbalance of socio-political power, especially between blacks and whites, where health differences are most pronounced. I show that the total number of excess deaths among blacks is large enough to affect the outcome of presidential elections. And, finally, I investigate the fluctuations of national and black and white IMRs associated with the political party in power. Party differences in policy preferences lead to different results in many aspects of social development, of which health is no exception. Results show that a good portion of the comparative international underperformance of the U.S. for IMR as well as of persistent black-white disparities in IMR is related to variations imparted under Republican presidents between 1965 and 2010.

The dissertation of Javier Mauricio Rodriguez Jr. is approved.

Mark Q. Sawyer

Teresa E. Seeman

James DeNardo

Peter M. Bentler

David O. Sears, Committee Chair

University of California, Los Angeles

2012

DEDICATION

It would have been impossible to finish this work without an unpardonable level of selfishness. That is why I do not dedicate this effort to those I love *because I* love them. I dedicate it to those I love because they did not stop loving me in spite of everything: Sandra—the Daughter of Nature—my wife, *mi campesinita chiquinquireña*; Javier—Pipo—my dad, *el insobrevivible*; and Julián—Chupis Mupis—my brother, *el que me llenó las tardes de rubíes*.

This work is, in essence, about premature death. Graciela—Mima—my mother, *mi otro yo*, closed her eyes at age 42. Because this is not a coincidence, this work is also dedicated to the memory she imprinted in me of our few years together.

TABLE OF CONTENTS

LIST OF FIGURES

Figures for "Age, SES, Mortality, and Socio-political Participation"

Figures for "Differential Mortality and the Racial Composition of the Electorate

Figures for "Infant Mortality and the President's Party"

LIST OF TABLES

Tables for "Age, SES, Mortality, and Socio-political Participation"

Tables for "Differential Mortality and the Racial Composition of the Electorate

Tables for "Infant Mortality and the President's Party"

AKNOWLEDGMENTS

It was Mark Sawyer's idea that I should come to study at UCLA. Absolutely nothing of this would have been possible without him. Our conversations began in September 29, 2006, at UC-Irvine, and have not stopped ever since. Paradoxically, I never took a class of his on African American Thought or Afro-Caribbeans in the United States, and yet it was from him, during an afternoon of conversation, that I first heard the totally-unrelated terms "Allostatic Load" and "Disparities in Health". Such diverse and rich are our conversations. From start to finish, Mark's sharp eclecticism nourished this project without reservation.

It took me about nine months of readings to figure out the obvious: That Teresa Seeman, my very favorite author on Allostatic Load research, was a professor at the UCLA School of Medicine—in my own backyard. So I phoned her. The issue was, however, what does a political scientist have in common with an epidemiologist who works in geriatrics? But that is Teresa's genius: She makes everything easy… and interesting. This project would have never happened without MIDUS meetings, without Teresa's leadership, without her walking me through the methodological and theoretical palaces (and sometimes slums) of medical research and public health.

I recall that one of the first things I told Teresa was that I was working with Peter Bentler on a new application that would allow running sensitivity analyses on structural models. During those days Peter's seminars equipped me with a statistical jargon complex enough to pale the geekest biostatistician. No matter your methodological background, taking Peter's seminars you will be reborn; well, you have to. The methodological re-wiring that Peter helped assemble in my brain was crucial not only for the various methodological approaches that I applied in this project

but also for the interdisciplinary thinking required to maintain some coherence as one goes from one methodological approach to the next.

Looking back in time I can point out with certainty the moment and place when and where I became a social scientist. It was James DeNardo's seminar 200A. I have attended his classes on the Central Limit Theorem and on the shapes of distributions many times, and I still get goose bumps, every time. I paid part of my Engineering tuition writing poems, I self-published two books of poems and I won national prices with those poems, and I can assert without doubt that there are about six or seven of James' classes that are as close to poetry as you can get. The material that I feel I can teach the best for the rest of my days is the material that I learned from James. In the three articles of this dissertation I think I do not talk of anything else.

During my early years of college I often felt that attending college was interfering with my education. To learn was for me an ethical issue; it was a compromise I had to myself. And college was not helping. There were aspects of my person that got irreconcilably separated, and thus was the idea that I accepted about human nature. David Sears, with his sapience, put one by one, little by little, all the pieces back together. During my time at UCLA, David constructed and deconstructed the study of the human mind in its magnificent complexity, and it was through that procedure that he gave me the academic identity I have today. This dissertation, both as a document and as a process, celebrates that reconciliation.

This project is also the result of collective and individual actions, sometimes heroic, between or by persons, sometimes close to me sometimes that appeared and disappeared just like phantoms, that somehow led me to this point: Rafael Jimeno, Robert Hitchcock, Tania Pantoja, Joseph Brown, Elizabeth Stephenson, Arun Karlamangla, Elias Bareinboim, John Zaller, Ana González, Dana Martinez-Miller, Luz E. García, and César, Sebastián and Santiago Cifuentes.

VITA

Javier M. Rodriguez

EDUCATION

University of California, Los Angeles

Major Fields: American Politics, The Politics of Race and Ethnicity, and Quantitative Methods
Dissertation: The effects of mortality and health inequalities over disparities in political behavior
Committee: David Sears (Chair, Political Science/Psychology), Teresa Seeman (Medicine/ Epidemiology), Mark Sawyer (Political Science/African American Studies), Peter Bentler (Psychology/Statistics), and James DeNardo (Political Science)

University of California, Los Angeles	**M.A. in Political Science, 2009**
Arizona State University, Tempe	**M.A. in Political Science, 2007**
University of Nebraska, Lincoln	**B.A. in History, 2005**

OTHER EDUCATION

Universidad de La Sabana, Colombia	Completed 6 years of coursework in Engineering, 1993-1999
Stanford University, Palo Alto	Summer Institute in Political Psychology, 2008

RESEARCH INTERESTS

The Effects of Health Disparities and Psychosocial Factors on Socio-political Participation. The Consequences of Racial Health Disparities on Socio-Political Processes. Health Policy. Regression Analysis and Structural Equation Modeling

FELLOWSHIPS AND AWARDS

• Cotta Robles Graduate Research Mentorship	**2011**
• Graduate Research Mentorship, UCLA	**2010**
• Department of Political Science, UCLA, Travel Grant	**2009, 2010, 2011**
• Politics of Race, Immigration, and Ethnicity Consortium, Travel Grant	**2008, 2009, 2010**
• The Eugene and Cotta Robles Fellowship, UCLA	**2007-2009**
• Graduate Summer Research Mentorship, UCLA	**2009, 2010**
• Graduate Summer Research Mentorship, UCLA, Travel Grant	**2009, 2010**
• Center for the Study of Race, Ethnicity, and Politics, UCLA, Travel Grant	**2009, 2010**
• Center for the Study of Race, Ethnicity, and Politics, UCLA, Research Grant	**2009**
• National Science Foundation/University of California (DIGSSS), Travel Grant	**2009**
• Department of Psychology, Political Psychology Fund, UCLA, Travel Grant	**2009**
• National Science Foundation/University of California (DIGSSS)	**2008**
• Department of Psychology, Political Psychology Fund, UCLA, Summer Grant	**2008**
• Graduate and Professional Students Association, ASU, Travel Grant	**2006, 2007**
• Underrepresented Graduate Enrichment Match Fellowship (UGEM), ASU	**2005-2007**
• The Shades of Leadership Award, UNL	**2005**

PUBLICATIONS AND SELECTED PRESENTATIONS

Pantoja, Adrian D., Rafael Jimeno, and **Javier M. Rodriguez**. (Forthcoming.) "The Political Consequences of Latino Immigrant Transnational Ties." In David L. Leal and Jose E. Limon, Eds. *Immigration and The Border: Politics and Policy in the New Century.* South Bend: University of Notre Dame Press

Rodriguez, Javier M. "The Contribution of Mortality Gaps Between African Americans and Whites to Their Disparities in Voter Turnout." Presented at APSA, Annual Meeting, Washington D.C., September, 2010

Rodriguez, Javier M., Mark Sawyer, and David O. Sears. "Early and Later Political Socialization in Latin American Immigrants: Assessing the Impact of Immigration." Presented at MPSA, Annual Meeting, Chicago, 2009. *Under review, American Sociological Review*

Murugesan, Vani, **Javier M. Rodriguez**, and David O. Sears. "Impressionable Years: The Impact of Parental Socialization and College Education on Later Political Participation." Presented at ISPP, Annual Conference, Dublin, Ireland, July, 2009. *Under review, Journal of Politics*

OTHER WORKS IN PROGRESS

Rodriguez, Javier M. "Age, Socioeconomic Status, Mortality, and Socio-political Participation"

Rodriguez, Javier M. "Infant Mortality and the President's Party"

Rodriguez, Javier M., Teresa Seeman, Arun Karlamangla, Tara Gruenewald, Sharon Merkin, and Dana Martinez-Miller. "Social Stratification and Allostatic Load: Socioeconomic-demographic Structures of Health Differences in Society"

Rodriguez, Javier M., Teresa Seeman, and Arun Karlamangla. "Differences Between Chronological and Biological Aging in African Americans and Whites"

Rodriguez, Javier M., Eric Wu, and Peter M. Bentler. "Applying Sensitivity Analysis to Structural Equation Modeling: An Application of the R Interface to EQS"

Rodriguez, Javier M. "Correcting for Multicollinearity When Squared Terms are Included in Regression and SEM"

RELEVANT RESEARCH EXPERIENCE

Data Analyst, **03/2010–present**
Geriatrics Division, David Geffen School of Medicine, UCLA

- Collaborator and co-author in several papers that investigate the effects of socioeconomic status and psychosocial factors on biological markers of health.

Data Analyst, **10/2008–present**
UCLA Center for the Study of Race, Ethnicity, and Politics (CSREP)

- Ran and reported data and statistical analyses for data generated by the CSREP and other affiliated institutions. Contributed with survey design and questionnaire construction.

RELEVANT TEACHING EXPERIENCE

Teaching Associate
Department of Political Science, University of California, Los Angeles

Michael Ross. *The Political Economy of Development*	**Spring quarter, 2010**
James DeNardo. *Data Analysis and Statistics*	**Winter quarters, 2009, 2010**
James Honaker. *Introduction to Statistics*	**Spring quarter, 2009**
John Zaller. *Introduction to Data Analysis and Statistics*	**Fall quarter, 2008**
John Zaller and Daniel Posner. *Comparative Politics and Statistics*	**Fall quarter, 2009**

- Taught three 50- or two 120-minute sections per week. Assisted instructor on duties such as grading homework and final papers, keeping grades and holding office hours

PROFESSIONAL AFFILIATIONS AND RESEARCH GROUP MEMBERSHIP

American Political Science Association
American Public Health Association
International Society of Political Psychology
UCLA MIDUS Research Group, Teresa Seeman, Director
Causal Inference Group, Judea Pearl, Director
Political Psychology Laboratory, David Sears, Director
Advanced Quantitative Methods in Education Research (AQM), Noreen Webb, Director

AGE, SOCIOECONOMIC STATUS, MORTALITY, AND SOCIO-POLITICAL PARTICIPATION

Javier M. Rodriguez
Department of Political Science
University of California, Los Angeles

Abstract

Who is included in samples is at minimum conditional on survival. Samples, therefore, contain a bias induced by age-specific socioeconomic status (SES) differences in mortality rates, thus affecting relationships fundamental to the study of politics. This study examines how the relationship between age and socio-political participation is inflated by the fact that, as time progresses, a higher proportion of low-SES, low-level participants die and are therefore excluded from samples faster than high-SES, high-level participants. I use data from MIDUS I (*Midlife in the United States: A National Study of Health and Well-being*). Mortality was assessed over 10 years. Results show that selective survival affects the age-participation relationship (1) by exclusion of non-survivors from samples, (2) by a heterogeneous distribution of health factors between survivors and non-survivors, (3) by individual differences in health, (4) because age-specific mortality shapes the age composition of the age distribution of survivors, and (5) through mediation mechanisms.

Introduction

Aging is an inevitable and irreversible process that, while typically associated with the monotonic advancement of time, exhibits enormous variation across diverse human populations. Some nations, and some communities within nations, age faster than others. Mortality outcomes are, therefore, incessantly shaping and re-shaping the age distribution of survivors, thus impacting the numerous socio-political processes that we study on survivors.

One of the empirical and theoretical pillars in the social sciences is the relationship between age and socio-political participation. Socio-political participation increases as individuals grow older, peaking in middle age, and declines slightly thereafter (Tingsten 1937). Another powerful relationship is that of socioeconomic status (SES) and participation. Participation is higher among individuals of higher SES and lower among the disadvantaged (Verba and Nie 1972; Verba et al. 1978). A third relationship of interest—although commonly overlooked in the study of socio-political processes—is that between SES and mortality. Mortality rates are higher among individuals of lower SES; in fact, individuals of low SES die sooner compared to their high SES counterparts (Kitagawa and Hauser 1968, 1973; Pappas et al. 1993; Seeman et al. 2004, 2010).

Considering these three fundamental relationships, I hypothesize that the relationship between age and socio-political participation is partially spurious. Those individuals with lower participation rates die younger for reasons related to their low SES. Participation rates thus appear to increase as individuals grow older because, as time progresses, samples are being "distilled" by the absence of those low-SES individuals who die sooner and whose participation is lower while a higher proportion of high-SES, high-participation individuals survive and are included in samples.

This article illustrates how older-years survivors are higher-level participants (due to their higher SES), while those who die younger are lower-level participants (due to their lower SES). It further demonstrates that samples contain a natural bias induced by SES differences in mortality rates, an important relationship often ignored in the analysis of socio-political processes. Researchers are measuring participation (and, indeed, everything) on survivors alone. Current research analyses are thus capturing a relationship between age and participation artificially inflated by the fact that low-level participants die out and are therefore excluded from samples faster than high-level participants. Such a situation generates a favorable accumulation of participants *vis-à-vis* non-participants in samples. This is especially true in the middle-age years during which mortality rates increase due to SES disadvantage and the association between SES and participation becomes more notable. This pattern coincides with the period of the human lifespan in which participation is highest.

Age and Participation

The bivariate relationship between age and participation follows a pattern that has been noted across time and space for at least the last century. Tingsten (1937) summarized this pattern from a series of elections carried out between 1911 and 1935 in Switzerland, Germany, Denmark, Holland, Sweden, and the United States: "[…] lowest in the youngest age groups, rising successively, and reaching a maximum in the age groups around fifty; with increasing age the political interest once again tends to decrease" (79). Subsequent research found a similar pattern in other nations as well as in various American contexts for both electoral and non-electoral participation (Almond and Verba 1963; Campbell et al. 1960; Lipset 1960; Milbrath 1965; Nie, Verba and Kim 1974; Wolfinger and Rosenston 1980). After controlling for relevant covariates, multivariate analyses have also reproduced this age-participation trend. There is some

uncertainty, however, about the true participation trend of the elderly (Glenn and Grimes 1968; Hout and Knoke 1975). Some studies find that after controlling for some relevant covariates, participation levels continue to increase with age even among the elderly (Campbell et al. 1960; Rosenstone and Hansen 1993; Verba and Nie 1972).

A variety of reasons exists as to why participation is a function of age. Group relations, social roles, and statuses vary with age. Age is a primary force of social organization and performance (Neugarten 1968; Taylor et al. 1997). Resources, psychological engagement traits, and being the target of mobilization networks also vary with age, and are representative of the most common models of participation (Brady, Verba, and Schlozman 1995). An analysis of the specific causes of participation related to age, and of the specific causes of death, however, is beyond the scope of this article. My focus is narrowly on how the relationship between age and participation is altered by mortality rates that are linked to socioeconomic status.

Socioeconomic Status and Participation

Studies showing the association between measures of socioeconomic status and participation have proliferated throughout decades of empirical research (Arneson 1925; Campbell and Kahn 1952; Connelly and Field 1944; Gosnell 1927; Verba and Nie 1972). Coefficients of socioeconomic indicators are usually found to be the most substantive and robust in models of participation; a great portion of our current knowledge about participation stems from the SES model (Junn 2010; Leighley 1995). Contextual and personal factors, as well as aggregate and individual level trends that are related to participation, strongly correlate with both indices of, and independent, socioeconomic indicators (Almond and Verba 1963; Cho, Gimpel and Dyck 2006; Erbe 1964; Milbrath 1965; Olsen and Tully 1972; Welch 1990). Indeed, the centrality of the SES model may lie in that, common or similar components of the psychological

engagement model, the mobilization model, and the rational choice model of participation (Blais 2007) typically correlate with one or more indicators of socioeconomic status.

High-SES individuals are precisely those who are more psychologically engaged, who are the target of mobilization by organizations, and who have higher levels and quality of information to pursue more effectively their personal interests; the contrary is the case for low-SES individuals (Leighley 2001; Lijphart 1997; Plutzer 2002). Socioeconomic indicators, therefore, embody the conjunction of numerous complex mechanisms that prompt individual action, or inhibit it. From living in high-quality neighborhoods and having access to key services and institutions, to the acquisition and utilization of certain skills that advance social standing and wealth, SES is a fundamental variable for participation. Socioeconomic resources epitomize a spatial domain in which contextual variables meet individual-level forces for development and socio-political performance.

Age, Socioeconomic Status, Mortality, and Participation

The connections between health outcomes and the social environment can be traced as far back as Hippocrates (Seeman and Crimmins 2001). Research has consistently found that socioeconomic status is a robust predictor of disease, disability, and premature death, independent from biomedical predictors of health (Adler et al. 1993, 1994; Antonovsky 1967; Durkheim 1955). Of equal importance is the relationship between age and health outcomes. Age is related to patterns of physiological decline that converge with disability, comorbidity, and frailty (Fried et al. 2001), thus explaining why age-specific mortality rates are critical in providing an informative index of health (Collins 1935; Sydenstricker 1928). Age and SES are important determinants of the physiological deterioration of the major systems in the human body therefore influencing mortality (Seeman et al. 2010).

There is a clear pattern of associations between age, SES, and mortality rates that suggests a connection to socio-political participation trends in the lifespan. During the younger years, the vigorous biological condition of individuals helps them to endure the adversities related to low SES hence showing a low correlation between SES and mortality. But then, as multi-system physiological dysfunction increases with aging, the individual's vulnerability to environmental factors increases, and SES begins to correlate with mortality. As time progresses, low SES individuals die because of their poorer health while high SES individuals continue to survive because of their better health. As the majority of individuals with lower SES die off, the correlation between SES and mortality begins to decline at older ages once again (Crimmins et al. 2009). In sum, the relationship between SES and health outcomes is subtle among the younger-age groups, higher among the middle-age groups, and weakens again during the older years.

This pattern is very similar to the trend depicted by the age-participation relationship. This similarity suggests that part of the effects of age over participation may be mediated by mortality rates due to differences in SES. Considering that researchers are assessing socio-political participation (and, indeed, everything) on survivors alone, and that age-specific mortality does not happen at random, researchers are studying patterns of interest by comparing different groups of survivors at different stages of the lifespan (Crimmins, Kim and Seeman 2009); that is, while violating the norm of statistical comparability.

The associations described above imply that age and SES effects over participation should vary significantly between survivors and non-survivors. Younger-age groups show lower levels of participation because this subpopulation is a mix of high-level participants (high-SES, future survivors) and lower-level participants (low-SES, future non-survivors). With low

mortality rates among the young, participation is partially independent from age and SES. Alternatively, middle-age groups show peak levels of participation, because it is in this subpopulation where mortality differences increase as a result of physiological dysregulation related to aging and the effects of SES inequality. With low mortality rates among high-SES, high-level participants and high mortality rates among low-SES, low-level participants, participation shows a strong association with age and SES. Socioeconomic differences in participation among the elderly attenuate as there is an increase in disability, comorbidity, and/or frailty related to biological aging.

Research Design and Framework

As a general rule, who is and who is not included in research samples is at minimum conditional on survival; survival is *conditio sine qua non* for observational social research. The problem, therefore, is that many covariates that determine mortality are correlated with the processes that researchers study on samples composed of survivors alone. Being selected or not into a sample is pre-determined by many of the same covariates that researchers use to explain the processes they attempt to understand. To survive is therefore endogenous with respect to the outcomes researchers study using samples. Consequently, as sampling is selective on survival, inferences about the relationships extant in the world and that researchers assess using survivors (e.g., age and socio-political participation) are biased.

This line of reasoning suggests that an individual's probability of being selected into a sample is conditional on his/her probability of survival (e.g., if a person dies, the probability of being part of a study is exactly zero). Considering that randomization assigns equal probabilities of selection, then the net probability at time t_1 of being selected into a randomized study that will take place t years into the future is proportional to (1) the probability of being selected (in theory

equal for everyone) and to (2) the probability of survival t years into the future (proper of each individual at time t_1). This has critical implications for social research, because this means that processes of sample selection, even if they are randomized, are being compromised by non-random mortality. Thus, if for a given sample we are able to estimate at t_1 an individuals' probability of survival t years into the future, we could simulate how such a sample would look like in a future time t_2 should we run a longitudinal study on the sample t years later. Said differently, just for purposes of illustration, the selection of individuals into a sample at t_1 was, in the same manner, already pre-determined by their probability of survival at t_0, t years before the sample was even selected. Neither those individuals who died between t_0 and t_1 made it to the sample frame at t_1 nor those individuals who died between t_1 and t_2 to the sample frame at t_2, thus inducing a selective survival bias in the samples.

This depiction indicates that researchers are *always* dealing with samples truncated by death. Or, in plain words, all samples are biased *ab initio*. The scenario differs from cross-sectional to longitudinal data. Cross-sectional data are always truncated by pre-existent mortality whereas longitudinal data are not only always truncated at the first wave by pre-existing mortality, but are further truncated in subsequent waves by mortality that occurs between the waves. This situation is particularly troublesome because truncation by death is neither a problem of censored nor of missing data. It is a situation where observations are simply inexistent (Zhang and Rubin 2003).

An alternative to handle some of these intricacies is to use cross-sectional data (or say, the least-biased, most representative first wave of a longitudinal study) on which an epidemiological/mortality follow up was run. The epidemiological follow up affords researchers the ability to use all available data of both future survivors and future non-survivors before death

truncates the observations. This strategy has the advantage of reducing some of the complexities related to future sample heterogeneity, non-random missing data, and future truncation by death.

Inferences from this approach will be valid insofar as those with a low probability of survival were already manifesting lower levels of participation at the moment of measurement. If this assumption is true, the probability of survival should affect the relationship between age and participation extant in the cross-sections of the data before individuals die. This approach is therefore a conservative yet simple test of the influence that mortality exerts over the relationship between age and participation. This is especially true as one considers that we are attempting to capture the effects of mortality years before individuals die, and that samples already contain the bias induced by non-random pre-existent mortality.

Hypotheses

This study addresses the following testable hypotheses:

(1) Health differences between future survivors and non-survivors should have an impact on actual participation. Accordingly, by matching on the probability of survival between future survivors and non-survivors, their difference in participation should diminish.

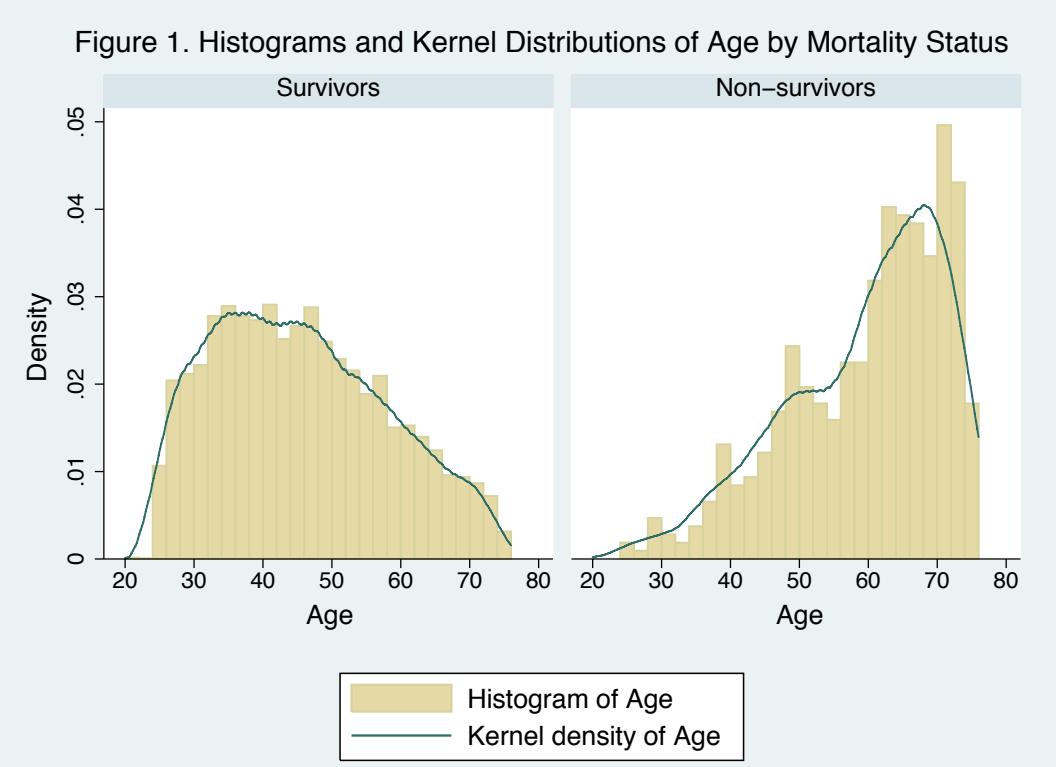
(2) The effect of the linear component of age over participation should differ depending on the sub-sample used in the analysis. It should be (a) Positive and substantive for a sub-sample of future survivors (because of the exclusion of low-level participation, future non-survivors); (b) Positive, although not as substantive as for the sub-sample of survivors, for the complete sample (a mix of high-participation future survivors and low-participation future non-survivors); and (c) A close-to-zero effect for a sub-sample of future non-survivors.

(3) The probability of survival should have a positive, independent effect over participation after controlling for age and socioeconomic status.

(4) Considering that the relationship between mortality and SES is weak in the younger-years, strong during the middle-age years, and weakens again in the older-years, the relationship between age and participation should be reverse between future survivors and non-survivors.

(5) Average participation should be higher for future survivors in comparison to future non-survivors across all age groups, but more notably between the ages of 40 to 65 years where mortality differentials by SES are most notable (Seeman and Crimmins 2001). Future survivors should also show a higher mean level of participation across all SES levels.

(6) The relationship between age and participation is not only altered by differences in participation levels between survivors and non-survivors. It is also altered by differences in the compositional structure of their age distributions. The age distribution of survivors is skewed to the right whereas the age distribution of non-survivors is skewed to the left (Figure 1). In participation terms, non-survivors—who show low participation levels—disappear at higher rates from the mid- to older-age groups in which participation is higher, leaving behind an artificial accumulation of low-level participants in the younger-age groups among survivors. Explicitly, we are underestimating the average level of participation of survivors because the proportion of survivors is much higher in the age groups in which participation is lower. At the same time, we are overestimating the average participation level of non-survivors because there is a high proportion of people dying during the years in which participation is the highest. Death occurs throughout the entire human lifespan, but especially toward the right tail of the age distribution in which participation is highest. Differently, individuals who join the age-participant distribution exclusively enter it through the left tail of the distribution where participation is lowest. Succinctly, mean differences in participation between survivors and non-survivors should be larger than what they appear to be.



(7) Considering that selective mortality due to SES differences between survivors and non-survivors affects the relationship between age and participation, the effects of both age and SES over participation should be mediated by the probability of survival. The total effect (i.e., the direct plus the indirect effects) of age over participation should be higher (i.e., overestimated) when the probability of survival is not considered as a mediator and lower when it is considered.

These hypotheses exemplify how non-random mortality affects the age-participation relationship in five different ways: (1) By exclusion of non-survivors from samples, (2) by a heterogeneous distribution of health factors between survivors and non-survivors, (3) by individual differences in health, (4) because age-specific mortality rates affect the composition of the age distribution of survivors, and (5) through mediation effects.

Methods

I first generate a propensity score on survival (i.e., the probability of survival) as a first step to remove the bias related to pre-treatment covariates between future survivors and non-

survivors. Subsequently, future survivors are matched to non-survivors on their propensity scores in order to balance on the covariates related to survival. In theory, participation (the outcome of interest) will be orthogonal on propensity score between future survivors and non-survivors. Third, I use a series of linear regressions implementing three samples (i.e., the complete sample, the sub-sample of survivors, and the sub-sample of non-survivors) to test for differences in the effect of age over participation. I then run an estimate of how much of the difference in participation rates between survivors and non-survivors is attributable to differences in the compositions of their age distributions. I do this by using customary demographic methods of standardization and decomposition. In the final step, I test for survival mediation effects using two structural equation models.[1]

Propensity Scores. I estimate the probability of survival (i.e., the "probability of being treated" (Mattei 2009; Rosenbaum and Rubin 1983; Rosenbaum 2010)) using an informative set of pre-treatment covariates related to health. Future survivors and non-survivors are included thus avoiding loss of information related to mortality. The "treatment" is to survive the 10-year period of the mortality follow up. This is a propensity score on survival—a summary variable I now can control for, and which represents a relevant number of covariates related to survival selectivity at the baseline. Explicitly, this propensity score is the predicted probability of survival using a vector of pre-treatment covariates X in a logit regression (Austin 2011).[2]

Matching on the Propensity Scores. I use Automated Coarsened Exact (ACE) Matching (Blackwell et al. 2010).[3] In theory, by matching on the probability of survival, to survive (or not)

[1] All statistical analyses are run using STATA 11 and EQS6.2.

[2] For a detailed description of variables, models, and other criteria please refer to the Propensity Scores Estimation Appendix.

[3] In order to avoid matching on propensity scores clustered around one (or zero), and to obtain a more normally distributed covariate, I also matched on the linear predictor and not on the predicted probabilities (Sekhon 2007). All results were virtually identical to the ones obtained by matching on the

should not be confounded with participation, conditional on the vector of pre-treatment covariates *X*. I am hence reducing the bias related to the differences in *X* between survivors and non-survivors; namely, the bias related to the probability of being or not included in the sample. In essence, I am matching survivors to non-survivors on propensity score intervals optimized by the ACE algorithm.[4]

***Linear Regressions*.** The polynomial models reproduce what is commonly reported in current research. These models were run using the complete sample (i.e., future survivors plus non-survivors), the sub-sample of future survivors, and the sub-sample of future non-survivors to illustrate how the relationship of age with participation changes depending on the sample that is analyzed. The matching output was used to correct for underlying health differences between future survivors and non-survivors. The probability of survival was used to illustrate how individual differences in health affect the relationship between age and participation. The four explanatory variables in the models are age, age-squared residuals, SES, and the probability of survival. An additional interactive model includes age group and SES level by survival status. Since very few individuals die before the age of 40, I truncated the sample to individuals 40 years or older. About 6% of non-survivors in the analyzed sample died before the age of 40 (that

---

predicted probability. For the sake of interpretation simplicity, I am therefore reporting results on the matched predicted probabilities. Worth noting is that ACE uses survivors as the reference group for balancing. Non-survivors are balanced according to the covariate distributions of survivors. Accordingly, non-survivors are assigned a case-weight whereas the case-weight for survivors equals 1.

[4] I used Sturge's automatic rule as the binning algorithm. I used different matching techniques including Exact, Coarsened Exact, Genetic, and Mahalanobis Nearest Neighbor. I ran these matching techniques on both the propensity scores and directly on the covariates extant in *X*. Matching directly on the covariates in *X* drastically diminished the number of successful matches; thus, matching on the propensity scores was a second-best but most efficient option for the purposes of this study, especially because the loss of information due to un-matched cases was much lower. Among all matching techniques ran on the propensity scores, I chose ACE because of its simplicity and easiness to understand, its adequate resulting balance and overlapping, the high final number of matched cases (i.e., low loss of information), it requires no normality assumption, and because it automatically confines the matching procedure to areas of the data with empirical support thus diminishing extrapolations dependent on the model (Iacus, King and Proo 2008; King and Zeng 2006; Blackwell et al. 2010).

is, 26 individuals for a 15-year range). By truncating the sample, inferences are supported by the data and not founded in model extrapolations. Robust estimation corrects the standard errors in all models for non-normality. Reported standard errors are also corrected for possible non-independence due to family membership.

***Age-standardization and Decomposition of Participation*.** The regression models do not capture entirely the age group variation extant in participation rates of future survivors and non-survivors. By controlling for age in regression models I am simply capturing the individual differences in participation due to age, not the true differences between age subgroups related to the composition of their age distributions. The composition of the age distributions is a variable by itself, and it needs to be taken into account (Bourgeois-Pichat and Taleb 1970; Bongaarts 1994). I therefore adjust for these differences between future survivors and non-survivors by making them have the same proportional age composition. Since survivors have an overwhelmingly younger age composition compared to the much older age composition of non-survivors, outcomes of age-standardization would differ depending on the distribution I choose as standard. For this reason, I used the average age compositions of the age distributions of survivors and non-survivors (Preston, Heuveline and Guilott 2001). By running a decomposition of the differences in the age-specific rates of participation between survivors and non-survivors, I estimate how much of their participation differences are due to the different age-compositions of their age distributions (Kitagawa 1955, 1964; Preston, Heuveline and Guilott 2001).

***Structural Equation Models*.** One of the advantages of structural equation modeling is its clear illustration of mediation effects. I ran two structural models; the first uses only future survivors and the second applies the matching output to the complete sample. The first model is run without the probability of survival, and the second is run including the probability of survival

as a mediator between the linear and the quadratic components of age and participation, and between SES and participation (Figure 4). These models afford the ability to conveniently estimate the direct and indirect effects on participation not possible to recover from the regression analyses. By doing so, I am able to assess how much of the effect that has been customarily attributed to age belongs to mortality when mediation is taken into account.

Both models were fitted using normal distribution theory maximum likelihood. Considering the possibility of non-independence between family members included in the sample, both models use Satorra's (1992) fourth-order weight matrix for clustered data. Lagrange Multiplier tests were also implemented to examine for the statistical necessity of parameter associations that might be added to the model to increase the quality of the fit. Wald tests were carried out to examine for redundancy in parameter associations that might be dropped out of the model for the sake of parsimony (Bentler and Mooijaart 1989). Comparisons of alternative nested models were done, and parameter associations were added or dropped depending on the improvement of the Chi-square statistic (χ^2) and on their theoretical feasibility (Bentler 2006; Steiger, Shapiro and Browne 1985). Final models show that no significant improvement of model fit could be achieved by either dropping or adding alternative paths between the latent factor, the variables and errors.

Sample and Variables

The data are from the MIDUS I study (*Midlife in the United States: A National Study of Health and Well-being*) (c.f., Brim et al. 1999). MIDUS I integrates a wide range of psychosocial and physical aspects of well-being related to social gradients in health status. It is a U.S. national representative sample of non-institutionalized English-speaking individuals aged 25 to 75 years living in a household with telephone service. This piece makes use of the main RDD, the sibling,

Table 1. Summary statistics for participation, survival status, demographic, SES, and health variables

	Analyzed Sample						After Matching	
	Complete Sample		Survivors		Non-survivors		Survivors	Non-survivors
	% (n)	Mean (SD)	% (n)	Mean (SD)	% (n)	Mean (SD)	Mean (SD)	Mean (SD)
Main Analyses Variables								
Participation index (0-4)		1.74 (1.24)		1.76 (1.24)		1.51 (1.30)	1.76 (1.24)	1.65 (1.28)
Socioeconomic status index (1-3)		1.959 (.79)		1.965 (.78)		1.897 (.79)	1.965 (.78)	1.979 (.77)
Age (25-75 years)		47 (12.9)		45.9 (12.5)		59.9 (11.0)	45.9 (12.5)	58.4 (11.1)
Propensity Scores Variables								
Survivor	92.6 (5,121)							
Black	4.57 (253)		4.53 (232)		5.12 (21)			
Female	52.1 (2,881)		52.6 (2,692)		46.1 (189)			
Neighborhood quality (1-4)		3.43 (.53)		3.43 (.53)		3.44 (.55)	3.43 (.53)	3.46 (.53)
Self-rated physical health (1-5)		3.56 (.97)		3.61 (.95)		2.91 (1.07)	3.61 (.95)	3.51 (.94)
Mental or emotional health (1-5)		3.81 (.93)		3.82 (.93)		3.64 (.95)	3.82 (.93)	3.74 (.94)
Self-rated general health (0-10)		7.48 (1.59)		7.54 (1.54)		6.71 (1.96)	7.54 (1.54)	7.48 (1.55)
Sum of chronic conditions (0-10)		2.36 (2.33)		2.28 (2.27)		3.34 (2.83)	2.28 (2.27)	2.46 (2.37)
Waist-to-hip ratio (0.46-1.39)		.88 (.10)		.88 (.10)		.91 (.09)	.88 (.10)	.88 (.09)
Body mass index (14.4-64)		26.7 (5.25)		26.6 (5.20)		27.0 (5.86)	26.6 (5.20)	27.1 (6.01)
Hospital/physic. visit index (0-4)		.51 (.55)		.49 (.51)		.79 (.86)	.49 (.51)	.49 (.54)
Alcohol consumption (0-5)		.22 (.67)		.22 (.67)		.21 (.72)	.22 (.67)	.22 (.71)
Smoking cigarettes (0-4)		.56 (1.19)		.53 (1.16)		.95 (1.46)	.53 (1.16)	.68 (1.29)

and the twin samples totaling 7,108 completed interviews. It also makes use of the MIDUS I 10-year epidemiological/mortality follow up. Mortality status of the baseline sample was recorded in 2005, 10 years after the survey was fielded in 1995/1996. The status of non-survivors was confirmed via phone, and by submitting names and social security numbers of participants to be

matched by the National Death Index (NDI). In total, 539 participants (7.6% of the baseline sample) were confirmed dead by 2005.

The present analyses use complete-data cases, with no imputation of missing data. Of the total 7,108 cases, 5,531 had enough data to estimate a propensity score.[5] Out of these 5,531 cases, 5,121 are future survivors and 410 are non-survivors. Of the 410 non-survivors, 407 were matched on propensity score to 5,119 survivors (total n = 5,526).[6] Variables used in the estimation of the propensity scores were chosen from a pool of variables related to health commonly implemented in the prediction of mortality (Table 1).

The variables included in the estimation of the propensity scores are: Survival status (dependent variable), race/ethnicity, physical health, mental or emotional health, present general health, sum of chronic conditions, alcohol consumption, waist-to-hip ratio (WHR), body mass index (BMI), illness-related hospitalization and physician visits, neighborhood quality, smoking (also interacted by gender), and gender (see Propensity Scores Estimation Appendix).

A description of the variables included in the main analyses follows.[7]

Socio-political participation. This is the outcome of interest. It is a summary index of participation constructed from 9 types of non-electoral participation.[8] These types of

[5] Variations of the propensity score models were tested including and excluding variables, recodes of the variables of interest, with and without interactions and quadratic terms where appropriate, with latent and/or summation indexes or with the composite covariates included independently, and dichotomization of categorical variables to correct for effect modifications or possible non-linearity, always considering the amount of missing data extant in the covariates and the best possible resulting balance using different matching procedures (Gelman and Hill 2007).

[6] The maximum number of cases lost to missing data on covariates used in multivariate analyses was very low (n = 22). All analyses were re-run using simple mode and/or mean imputation on these 22 missing-data cases. All results were virtually identical to the ones reported herein.

[7] Variables were included using the same format in the regression analyses and the structural equation models unless specified.

[8] Non-electoral forms of participation require more initiative than voting and are accessible to most of the population for they do not have residency requirements. It is also expected that by considering various forms of non-electoral participation a more comprehensive portrait of the age-participation relationship will be achieved (Nie, Verba and Kim 1974).

participation are classified in 3 different subscales, with each subscale contributing one third to the final index score, and each of the items in each subscale contributes one fourth, one third, or one half to the respective subscale score depending on the number of items in each subscale. The three participation subscales are: (1) volunteering, (2) attending meetings, and (3) giving money. The first subscale includes 4 items assessing the number of hours per month dedicated to volunteering work in hospitals, care-giving, charity, school, youth-related organizations, political organizations, and other organizations with a different cause. The second subscale includes 3 items assessing the number of times per month the respondent attended group meetings related to unions, professional and social organizations, or other groups. The third subscale includes 2 items that assess the average amount of dollars per month given to political organizations, charities, and other organizations. The final 5-category index of socio-political participation ranges from 0 to 4, with 0 meaning no participation across all 9 participation items, and 1 to 4 indicating quartile participation levels.

Age. This is the independent variable of interest. It represents the linear component of age; the quadratic component of age is represented by the age-squared residuals (see below). This variable is originally recorded as a continuous variable ranging from 25 to 74 years, and it was used in this format in the structural equation models. For illustrative purposes, and to simplify the regression analyses and interpretation of results, I censored the variable into 4 categories (less than 40 years of age = 1, 40 to 54 years = 2, 55 to 64 years = 3, and 65 to 75 years = 4). In the first set of regressions this variable is treated as a 4-category variable. In the interactive model the first category is eliminated to avoid model-dependent extrapolations, leaving thus a 3-category variable (40 to 54 years = 2, 55 to 64 = 3, and 65 to 75 = 4).[9]

[9] Misleading model-dependent extrapolations are common during the younger years. Mortality rates are very low under the age of 40. In the sample only 26 individuals died under the age of 40. The

Age-squared residuals. The inclusion of this term models the curvilinear trend commonly attributed to older-age groups and reduces bias due to a purely linear misspecification. Due to very high level of correlation between age and age-squared (ρ = 0.98, 0.98, and 0.99 for the complete sample, and survivors and non-survivors sub-samples, respectively), I regressed age-squared on age and recovered the residuals for each of the three samples. These residuals represent the quadratic aspect of age that is not correlated with the linear aspect of age. By including the age-squared residuals in the models, I avoid the instability that originates from high multicollinearity while capturing the quadratic relationship between participation and age.[10]

Socioeconomic status. This is an index constructed from 5 indicators. In the regression analyses, this index is the summation of the 5 indicators, all coded in values from 0 to 2, so each has the same range hence contributing equally to the final index score. In the structural equation models, this index constitutes a latent factor that represents the correlation of the 5 indicators.[11] Information on the 5 indicators follows.

A. *Education*. For the regression analyses (Less than high school = 0, High school/GED = 1, and Some college or higher = 2). For the structural equation models a 5-category variable was preferred (Less than high school = 0, High school/GED = 1, Some college/AA = 2, Bachelor's degree = 3, and Graduate college = 4).[12]

---

outcome of interest is manifested over the age of 40 years. Multiple ways of coding age were carried out for robustness check; all analyses showed virtually identical patterns to those reported herein.

[10] In order to reduce the range of variances and facilitate computation this variable was divided by 100 in the structural models.

[11] SES is modeled as a "reflective" factor, different from a "formative" approach (c.f. Bollen and Bauldry 2011). Considering the early stage of development of formative measurement theory (Hardin and Marcoulides 2011) and that research has not yet provided neither a mathematical nor statistical principle to decide which approach is most appropriate to operationalize a construct (Treiblmaier, Bentler, and Mair 2011), the SES latent variable here represents the correlation of the indicators, as it is used in structural models in the standard way.

[12] By treating this variable as a 5-category variable in the structural equation models (instead of a 3-category variable as in the regression models), multivariate kurtosis due to non-normality was

B. *Family-size adjusted poverty-to-income ratio*. This variable is adjusted for inflation using the Consumer Price Index.[13] This variable accounts for the respondent's, the spouse's, and other family members' personal earnings. It also accounts for social security, government assistance, and other sources of income. All dollar values were adjusted to 2006 dollars. After estimating the total family income, the size of the family was estimated by considering the number of children in the family aged 17 years or under, if the respondent has or does not have a spouse or a partner, plus other family members living in the household. Subsequently, poverty thresholds were obtained from the U.S. Census Bureau, and participants were assigned poverty thresholds according to where they live, family size, number of children under 18 years of age, and the year in which the data was collected.[14] Then the total participant's family income was divided by his or her respective poverty threshold. Finally, for the regression analyses, this variable was recoded as follows: Less than 300% = 0, 300 to 599% = 1, and ≥600% = 2. For the structural equation models, this variable was censored into quintiles.[15]

C. *Current financial situation*. This measurement of financial strain asks respondents about their perception of their current financial situation. It is coded as follows: Worst possible financial situation = 0, Average = 1, and Best possible financial situation = 2.

D. *Enough money to meet needs*. This is another measurement of financial strain that asks respondents if their income is enough to cover their basic needs. This variable is coded as follows: Not enough money to meet needs = 0, Just enough = 1, and More than enough = 2.

---

diminished without loss of information. Other ways of coding this variable were used; results were practically identical to the ones reported herein.

[13] Data from the U.S. Department of Labor is available to the public at ftp://ftp.bls.gov/pub/special.requests/cpi/cpiai.txt

[14] Data available to the public at http://www.census.gov/hhes/www/poverty/threshld.html

[15] By censoring the variable into quintiles, multivariate kurtosis related to non-normality was diminished without loss of information. The models were also run using other codifications of this variable; results were practically identical in all instances.

E. *Difficulty paying bills*. This variable measures financial strain and taps on difficulties respondents endured to fulfill their financial responsibilities. This variable is coded as follows: Very or somewhat difficult to pay bills = 0, Not very difficult = 1, and Not difficult at all = 2.

For the regression analyses, the final summation of these 5 indicators varies from 0 to 10. In order to simplify the interpretation of results, the variable was censored into 3 categories (values from 0 to 3 = 1 (meaning low SES), from 4 to 6 = 2 (meaning middle SES), and from 7 to 10 = 3 (meaning high SES)).[16] For the structural equation models, a single-factor model showed to be the most adequate (Eigenvalue = 2.442).[17]

Results

There is a difference in participation of about 0.25 (SE = 0.06) participation units between survivors and non-survivors, meaning that future survivors participate 17% more than future non-survivors at the baseline.[18] This difference in participation is high considering, for example, that the difference in participation between young adults (25 to 39 years of age) and older adults (65 to 75 years of age) is 0.21 participation units. The age distribution differs drastically between future survivors and non-survivors; 26% of future survivors are 55 years or older compared to 70% of non-survivors. Future non-survivors are 14 years older than future survivors, on average.[19] Similar important differences conditional on survival status can be

[16] Censoring the variable into 3 categories is convenient because the number of non-survivors in some cells of the original 10-category SES index is low, and this would be problematic in an interactive model. By pooling participants into 3 levels of SES a less model-dependent, more intuitive interpretation of the results can be made without much loss of information.

[17] Eigenvalue for a second latent factor = 1.079. Due to its closeness to 1.0, the structural models were also run using a two-factor SES indicator. These alternative models showed a poorer fit to the data, more instability, and added no extra explanatory power to the models. In general, however, results were very similar to the ones reported herein.

[18] An independent samples T-test shows that the difference in participation between future survivors and non-survivors is statistically significant (T = 3.8376, $p = 0.0001$).

[19] An independent samples T-test shows that the difference in age between survivors and the deceased is statistically significant (T=21.9182, $p = 0.0000$).

detected between blacks and non-blacks as well as between males and females. For instance, future survivors consists of 53% females while non-survivors consists of 46% females. These three demographic indicators (race, gender, and age) encapsulate substantial differences in mortality schedules.

The mean SES level of future survivors was 1.97 (SD = 0.78) and 1.90 (SD = 0.79) for non-survivors (i.e., an apparently subtle difference of 4% in SES favoring future survivors).[20] But this small difference is not representative of other SES differences related to other indicators. For instance, 55-to-64-year-old future survivors have a SES that is 16% higher than non-survivors of the same age group, and future black survivors have a SES 26% higher than future black non-survivors. Future high-SES survivors participate about 60% more than low-SES survivors and 85% more than low-SES non-survivors, on average. Differences in SES between future survivors and non-survivors are radically higher once other factors are taken into account.

Health and well-being indicators show that the sample was fairly healthy. Participants lived in mid-high quality neighborhoods, on average. They also showed mid-to-high self-rated physical, mental/emotional, and current general health scores. Participants also showed low mean levels of hospitalization as well as low alcohol and cigarette consumption. Their average score in the number of chronic conditions was 2.36 (SD = 2.33) out of a range between 0 and 10, and their average WHR and BMI were out of the high-risk clinical range. Important differences, however, were detected between future survivors and non-survivors. For instance, 35% of future non-survivors reported a "fair" or "poor" physical health compared to 11% of survivors, 47% of non-survivors had two or fewer chronic conditions compared to 63% of survivors, and 66% of non-survivors were non-smokers compared to 80% of survivors.

[20] An independent samples T-test shows that the difference in SES between future survivors and non-survivors is marginally statistically significant at conventional levels (T = 1.6741, $p = 0.0942$).

Table 2. Matching Summary and Balance Statistics

Variables	$\mathcal{L}_1^{j}$ *		$\mathcal{I}_1$ **		25%		50%		75%	
	Before	**After**	**Before**	**After**	**Before**	**After**	**Before**	**After**	**Before**	**After**
Propensity score	*.40*	*.12*	*.09*	*.00*	*.13*	*.01*	*.06*	*.01*	*.03*	*.00*
Race	.01	.01	-.01	-.01	.00	.00	.00	.00	.00	.00
Perceived neighborhood quality	.09	.11	-.01	-.03	.00	.00	.00	.00	-.25	-.25
Physical health rate	.27	.05	.71	.11	1.00	.00	1.00	.00	.00	.00
Mental or emotional health	.10	.05	.19	.08	.00	.00	.00	.00	1.00	1.00
Self-rated general health	.18	.05	.84	.06	2.00	.00	1.00	.00	1.00	.00
Sum chronic conditions	.17	.04	-1.06	-.18	.00	.00	-1.00	.00	-2.00	-1.00
Alcohol consumption	.04	.02	.01	.00	.00	.00	.00	.00	.00	.00
Waist-to-hip ratio	.20	.15	-.04	-.01	-.05	.00	-.04	-.01	-.03	.00
Body mass index	.11	.13	-.40	-.46	-.13	.00	-.17	-.17	-.47	-.08
Hospital/physician visit index	.18	.05	-.30	-.00	.00	.00	.00	.00	-.33	.00
Smoking cigarettes	.14	.06	-.42	-.15	.00	.00	.00	.00	-2.00	-1.00
Gender	.06	.01	.07	.01	.00	.00	1.00	.00	.00	.00

Number of strata = 13
Number of matched strata = 11

	Non-survivors	Survivors
All =	410	5121
Matched =	407	5119
Non-matched =	3	2

* A perfect balance would be indicated by $\mathcal{L}_1 = 0$, a perfect imbalance by $\mathcal{L}_1 = 1$.

** Difference of the means between survivors and non-survivors.

Levels of participation between future survivors and non-survivors differed drastically across health indicators. For instance, future survivors with "good" or "better" physical health at the baseline participated 23% more than survivors with "fair" or "poor" physical health and 37% more than non-survivors with "fair" or "poor" physical health. Future survivors with a BMI<30 participated 8% more than survivors with BMI ≥30 and 23% more than non-survivors with BMI ≥30. Non-smoking future survivors participated 25% more than smoking survivors and 37% more than smoking non-survivors. There are strong differences in participation not only across demographic indicators but equally so as these indicators are linked to health.

Table 2 reports a summary of the matching procedure and a set of imbalance statistics comparing the level of imbalance of the propensity score and the covariates extant in X before and after matching. These imbalance statistics are the variable-by-variable distributional difference $\mathcal{L}_1^j$, the difference of the means $\mathcal{I}_1$, and the differences for the 25th, 50th, and 75th percentiles in the distributions.[21] Considering that the goal is to augment the level of independence between participation of future survivors and non-survivors on the propensity score, the statistic of main interest is the change in $\mathcal{L}_1$ due to matching.

Table 2 shows that most of the covariates in X diminished their imbalance, especially those with higher levels of imbalance before matching. Low imbalanced variables (race, perceived neighborhood quality, and BMI) increased their imbalance in trivial amounts. Most importantly, the imbalance on the propensity score got significantly diminished (imbalance after

[21] A better balance is achieved should the differences between global means and within the percentiles diminish. A perfect balance would be indicated by $\mathcal{L}_1 = 0$, a perfect imbalance by $\mathcal{L}_1 = 1$ (Blackwell et al., 2010). A full description of these statistics as well as their mathematical formulae are developed somewhere else (Iacus, King and Porro 2008).

Table 3. Participation OLS Polynomial Models Parameter Estimates

	(1)	(2)	(3)	(4)
	Complete Sample	**Complete Sample**	**Future Survivors**	**Future Non-survivors**
Corrections				
***Exclusion*[1]**	Yes	Yes	No	NA‡
***Individual health diffs.*[2]**	No	Yes	No	No
Covariates				
Age	.05**	.07**	.08**	-.01
	(.017)	(.018)	(.018)	(.07)
Age-squared residuals	-.05**	-.05**	-.06**	.23**
	(.02)	(.02)	(.02)	(.07)
SES	.40**	.38**	.40**	.40**
	(.02)	(.02)	(.02)	(.08)
Probability of Survival		.98**		
		(.23)		
Constant	.84**	-.06	.81**	.78**
	(.05)	(.22)	(.05)	(.24)
Observations	5,509	5,509	5,101	408
R-squared	.070	.074	.073	.075

[1] "Yes" means that the analysis *includes* future non-survivors in the sample.
[2] "Yes" means that I am accounting for differences in participation due to individual differences in their probability of survival.
‡ Not Applicable.
Note: Dependent variable is socio-political participation. Robust standard errors correcting for non-normality and family membership clusters in parentheses.
Statistical significance code: ** p<0.01, * p<0.05.

matching was contracted by 69%).[22] Importantly, Table 1 shows that the difference of the means in participation between future survivors and non-survivors was diminished from 0.25 to 0.11 participation units, a contraction of 56% after matching. This contraction of the difference of the means states that the combination of the health-related factors included in *X* used to estimate the probability of survival explains more than half of the original difference in participation between future survivors and non-survivors. This finding favors the expectations of the first hypothesis.

[22] Computed as follows: [((0.396 – 0.123) / 0.396) • 100] = 69% (Iacus, King and Porro 2008).

Table 3 shows the polynomial models output. Two main corrections are applied in these models: (1) Exclusion of non-survivors due to selective survival,[23] and (2) individual health differences related to the probability of survival.[24] Model 1 captures the age-participation relationship as is in the complete sample at the baseline (i.e., it includes both future survivors and non-survivors but it does not control for individual health differences related to survival). Model 2 reproduces Model 1 but controls for individual differences in health. Model 3 is an approximation of what current research reports: Samples of only survivors without controlling for health factors. Model 4 models the participation pattern of future non-survivors.

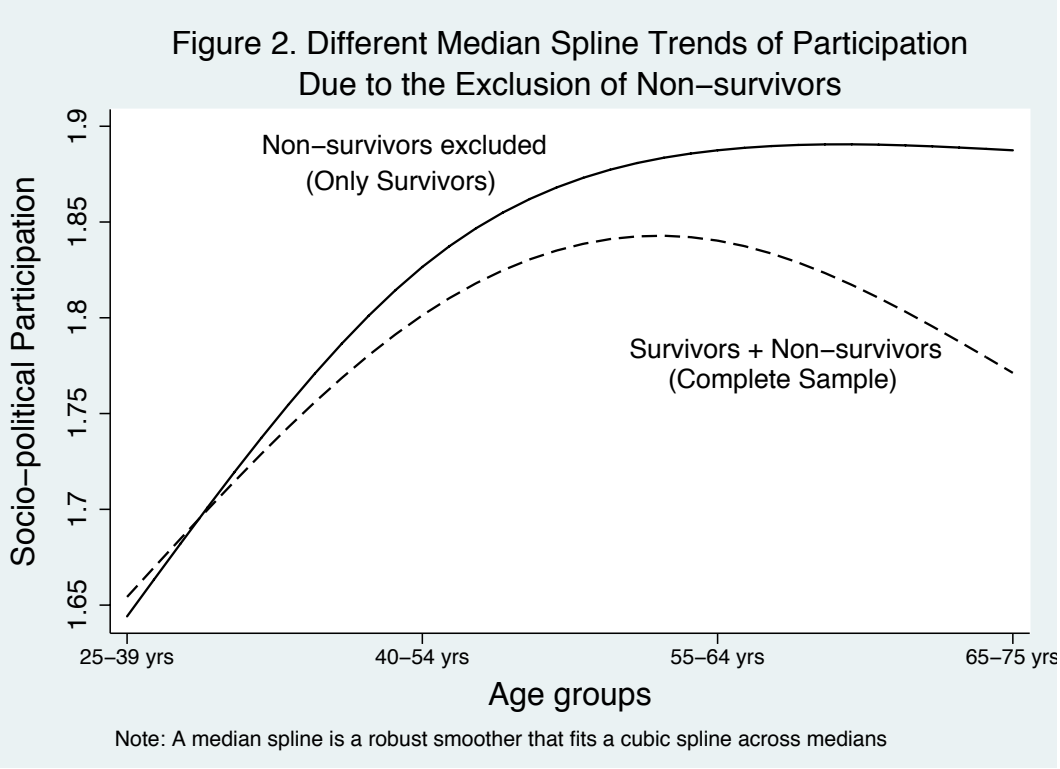
Models 1, 3, and 4 evince that the coefficients of both age terms vary both in size and statistical significance depending on the analyzed sample. The effect of the linear component of age is higher for survivors (Model 3), intermediate for the complete sample (Model 1), and much lower (and non-statistically significant) for non-survivors (Model 4), as expected by the second hypothesis. By disaggregating the complete sample into future survivors and non-survivors, it becomes evident that the strong effect of age over participation, as we know it, is, in fact, artificially inflated by selective survival. A comparison between Models 1 and 3 shows an overestimation of the effects of the linear and quadratic components of age on participation of 55% and 13%, respectively, due to the exclusion of non-survivors (Figure 2).[25]

Model 2 shows that the probability of survival has a positive, independent effect over participation after controlling for age and socioeconomic status. This corroborates the third hypothesis. Healthy individuals with the highest probability of survival during the next 10 years

[23] I correct for this bias by using the complete sample at the baseline, namely including both future survivors *and* non-survivors, thus there is no loss of information due to *exclusion*.

[24] I correct for this bias by including the probability of survival in the models.

[25] Table 3 shows rounded figures. Model 1 exact coefficients are 0.0542 and -0.0540 for the linear and quadratic components of age, respectively. Model 3 exact coefficients are 0.0840 and -0.0608 for the linear and quadratic components of age, respectively.



from the baseline participate about 1 participation units (25%) more than individuals with the lowest probability of survival. Notice that only the coefficient of the linear component of age is mostly affected, meaning that differences in participation due to health factors mostly occur before the older years, when the effects of SES differences on survival are most pronounced. Should there be no detrimental differences in health, individuals would participate 28% more as they age;[26] according to democratic theory, a target of high priority for democratic systems.

Models 3 and 4 show that the patterns of association between age and participation are reverse between survivors and non-survivors, confirming the fourth hypothesis. Coefficients of the linear component of age evince that, on average, only survivors increase their participation as they grow older. The coefficient for non-survivors is negative and non-statistically significant.

[26] Model 2 exact age coefficient is 0.0692 (compared to 0.0542, Model 1). No significant change is detected for age-squared residuals coefficients (-0.0544 and -0.0540 for Models 2 and 1, respectively).

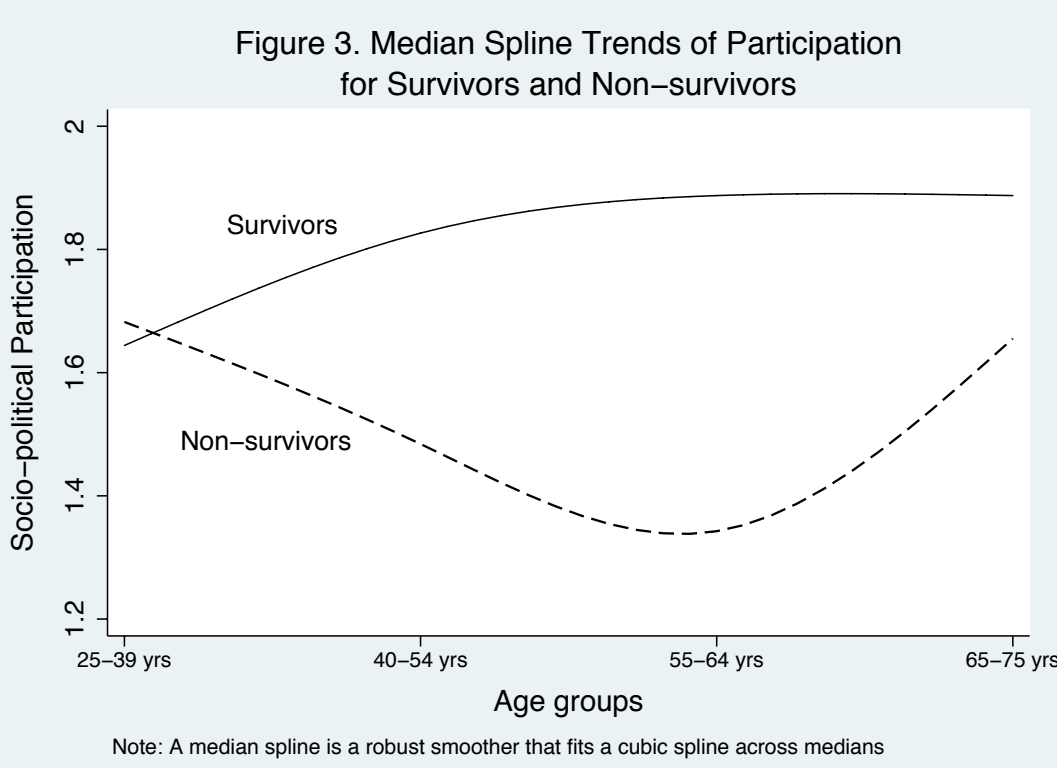


Figure 3 illustrates that survivors increase their participation as they age while non-survivors' trend of participation declines reaching a minimum located in the same age range where survivors reach their peak, then followed by a strong "come-back" in the older years. Maximum differences in participation between the groups are detected in their 50s and 60s, precisely where maximum differences in mortality due to SES differences manifest (Crimmins, Kim and Seeman 2009).

Non-survivors' reverse trend of participation has an explanation: Participation levels between survivors and non-survivors are similar among the young because the relationship between mortality and SES and age during this period is weak. As non-survivors age, their health vulnerability to SES disadvantages intensify generating a decline in their average participation. But if non-survivors survive to the older years, it is because their characteristics were more alike to those of survivors. Simply put, if a non-survivor died at age 65-75, in many ways he or she was already a "survivor". Average participation differences in the older years due to SES

Table 4. OLS Interactive Model Parameter Estimates

Covariates	Coeff. (SE)
Survival status (deceased=0, survivor=1)	.16
	(.17)
Age (=2)	.07
	(.20)
Age (=3)	.34*
	(.20)
Age (=2) x Survival status (=1)	-.19
	(.21)
Age (=3) x Survival status (=1)	-.32
	(.21)
SES (=2)	-.01
	(.19)
SES (=3)	.58***
	(.22)
SES (=2) x Survival status (=1)	.40**
	(.20)
SES (=3) x Survival status (=1)	.28
	(.22)
Constant	1.31***
	(.17)
Observations	3,681
R-squared	.075

Statistical significance code: *** p<0.01, ** p<0.05, * p<0.1
Note 1: Dependent variable is socio-political participation.
Note 2: Robust standard errors correcting for non-normality and family membership clusters in parentheses. Samples are weighted according to matching output.

attenuate, and non-survivors' participation begins to resemble that of survivors as they reach the right tail of the human lifespan. These findings evince that, to survive, is a condition linked to some of the same factors that make individuals participate because of their SES as they age.

Table 4 reports the output of the interactive model. Table 4.1 shows the predicted average participation of survivors and non-survivors by age group and SES level. Results show that survivors participate at higher levels than non-survivors across all age groups and SES-levels, on average. This finding supports the fifth hypothesis. If non-survivors participate less at all SES-

Table 4.1. Predicted Age-specific and SES-specific Participation by Survival Status (from Interactive Model)

Interactive terms	Predicted Participation	S.E.*	[95% C.I.]
Survival status by age-group			
Deceased 40 to 54 years	1.49	.14	1.22 1.76
Survivor 40 to 54 years	1.90	.03	1.84 1.95
Deceased 55 to 64 years	1.55	.15	1.26 1.84
Survivor 55 to 64 years	1.77	.04	1.68 1.85
Deceased 65 to 75 years	1.83	.14	1.54 2.11
Survivor 65 to 75 years	1.92	.06	1.81 2.03
Survival status by SES level			
Deceased low SES	1.38	.14	1.11 1.66
Survivor low SES	1.44	.04	1.36 1.52
Deceased mid SES	1.37	.13	1.11 1.63
Survivor mid SES	1.82	.03	1.76 1.89
Deceased high SES	1.96	.17	1.62 2.31
Survivor high SES	2.30	.04	2.23 2.38

* Standard errors estimated using the Delta-method

levels than survivors, and these are the same individuals who die off from samples and participate less across all age groups, then samples are progressively composed of high-level participants who did not die because of their SES comparative advantage.

Table 4.1 illustrates that differences in participation between survivors and non-survivors decrease as age progresses and that this difference is higher for middle- and high-SES individuals. Considering that the size of the bias in samples due to selective mortality is proportional to the size of the differences in participation between survivors and non-survivors, an important portion of the contribution to the bias comes from middle- to high-SES, 40-to-64 years individuals. But the size of the bias is also proportional to the ratio of the number of individuals who die off from, to the number of survivors in, each age group. Accordingly,

Table 4.2 Percent of Non-survivors After 10 Years by SES Level and Age Group

	40 to 54 years	55 to 64 years	65 to 75 years	Total
Low SES	6.3	14.3	33.8	134
Middle SES	4.2	13.5	21.9	144
High SES	3.6	8.8	26.4	103
Total	96	119	166	381

another important portion of the bias comes from low- to middle-SES individuals who comprise a bigger share of non-survivors compared to high-SES individuals across age groups (Table 4.2).

This situation suggests that estimations of participation differences between survivors and non-survivors may still be underestimated due to the different age composition of their age distributions. As stated above, mortality rates differ across age groups, thus the pace at which lower-level participants die off shapes the composition of the age distribution of higher-level participation survivors. Correspondingly, Table 5 shows the age-standardization of age-specific rates of participation.[27] The difference of crude participation rates derived from the predicted age-specific means of participation between survivors and non-survivors is 0.20, but should survivors and non-survivors have the same proportion of individuals across age groups, the difference in participation would be 0.26. The age-standardization of the difference of crude participation rates between survivors and non-survivors confirms an approximate 30% underestimation of the difference in participation due to age compositional differences in the age distributions of survivors and non-survivors (Table 5.1). This finding favors the sixth hypothesis.

[27] For our purposes, the age-specific participation rates are the age-specific predicted means estimated with the interactive model (see Table 4.1), meaning that these figures control for SES individual differences in participation between survivors and non-survivors.

Table 5. Standardization of Age-specific Rates of Socio-political Participation (from Interactive Model)

Age group	Crude Part. Rate (non-survivors)	Crude Part. Rate (survivors)	Standardized Crude Part. Rate (non-survivors)	Standardized Crude Part. Rate (survivors)
i	CPR^N	CPR^S	$ASCPR^N$	$ASCPR^S$
40-54 years	0.37	1.13	0.63	0.80
55-64 years	0.48	0.47	0.45	0.51
65-75 years	0.80	0.27	0.53	0.55
	1.66	1.86	1.60	1.86
			Difference of crude participation rates ($CPR^S - CPR^N$) = 0.20	
			($ASCPR^S - ASCPR^N$) = 0.26*	

* This is our difference of interest: The difference in participation should survivors and non-survivors had the same age distribution (averaged from their age-distribution compositions)

Note: ASCPR = Age-Standardized Crude Participation Rate, where

C_i^N = Fraction of non-survivors in i

C_i^S = Fraction of survivors in i

$(C_i^N + C_i^S) / 2$ = Fraction of individuals in i averaged from survivors' and non-survivors's age com

P_i^N = Age-specific participation rate of non-survivors (from interactive model (Table 4.

P_i^S = Age-specific participation rate of survivors (from interactive model (Table 4.1))

$CPR^N = \Sigma (P_i^N \bullet C_i^N)$

$CPR^S = \Sigma (P_i^S \bullet C_i^S)$

$ASCPR^N = \Sigma \{P_i^N \bullet [(C_i^N + C_i^S) / 2]\}$

$ASCPR^S = \Sigma \{P_i^S \bullet [(C_i^N + C_i^S) / 2]\}$

Table 5.1 depicts the decomposition of the difference of crude rates of participation into the contributions of age-specific compositional differences and of age-specific participation differences. Results evince that most of age-specific differences of crude participation rates between survivors and non-survivors arise from their age compositional differences, much more so than the contribution from the differences of their age-specific participation rates (Table 5.1).[28] Between the ages of 40 and 54 years, the contribution to age-specific differences of crude

[28] Please note that the difference of crude participation rates is composed of two terms: (1) The contribution of age compositional differences (θ), and (2) the contribution of age-specific participation rate differences (ψ). The first one represents the contribution that emanates from compositional differences; the second represents the contribution that emanates from participation differences.

Table 5.1. Decomposition of Differences in Crude Rates of Socio-political Participation Between Survivors and Non-survivors[29]

Age group	Contribution of age compositional differences to crude participation rate differences (θ)	Contribution of age-specific participation rate differences to crude participation rate differences (ψ)	Difference of crude participation rates
i	$\theta = (C_i^s - C_i^d) \bullet [(P_i^s + P_i^d) / 2]$	$\psi = (P_i^s - P_i^d) \bullet [(C_i^s + C_i^d) / 2]$	$\theta + \psi$
40-54 years	0.579 (77%)	0.172 (23%)	0.751
55-64 years	−0.079 (56%)	0.062 (44%)	−0.017
65-75 years	−0.551 (95%)	0.026 (5%)	−0.525
Sum	−0.052 (17%)	0.261 (83%)	0.209

Difference of crude rates of participation between survivors and non-survivors = $CPR^s - CPR^d$

$$= \underbrace{\sum_i (C_i^s - C_i^d) \bullet \left[\frac{P_i^s + P_i^d}{2}\right]}_{\theta} + \underbrace{\sum_i (P_i^s - P_i^d) \bullet \left[\frac{C_i^s + C_i^d}{2}\right]}_{\psi}$$

participation rates that result from their age compositional differences is 77%, between the ages of 55 and 64 years is 56%, and between the ages of 65 to 75 years is 95%.

These findings mean that, as we study the relationship between age and participation (and, indeed, many others) through the analysis of survivors alone, we incur in biased estimations of the true age-participation relationship because we are attributing to proportions of survivors, located in different age groups, properties that arise because of the exclusion of individuals who hold an important fraction of the variation in participation in such age groups. As time progresses and low-level participants die off, the variation in participation of survivors contracts and gets concentrated in the ranges proper of middle- to high-SES strata. Mortality

[29] This decomposition is based on the work of Evelyn M. Kitagawa (1955 and 1964). The deduction of the formulae has been developed somewhere else (Preston et al., 2001).

Figure 4. Two Structural Equation Models (Model 1 (Solid Paths) and Model 2 (Model 1 + Dash Paths)) and the Mediation Effects of the Probability of Survival

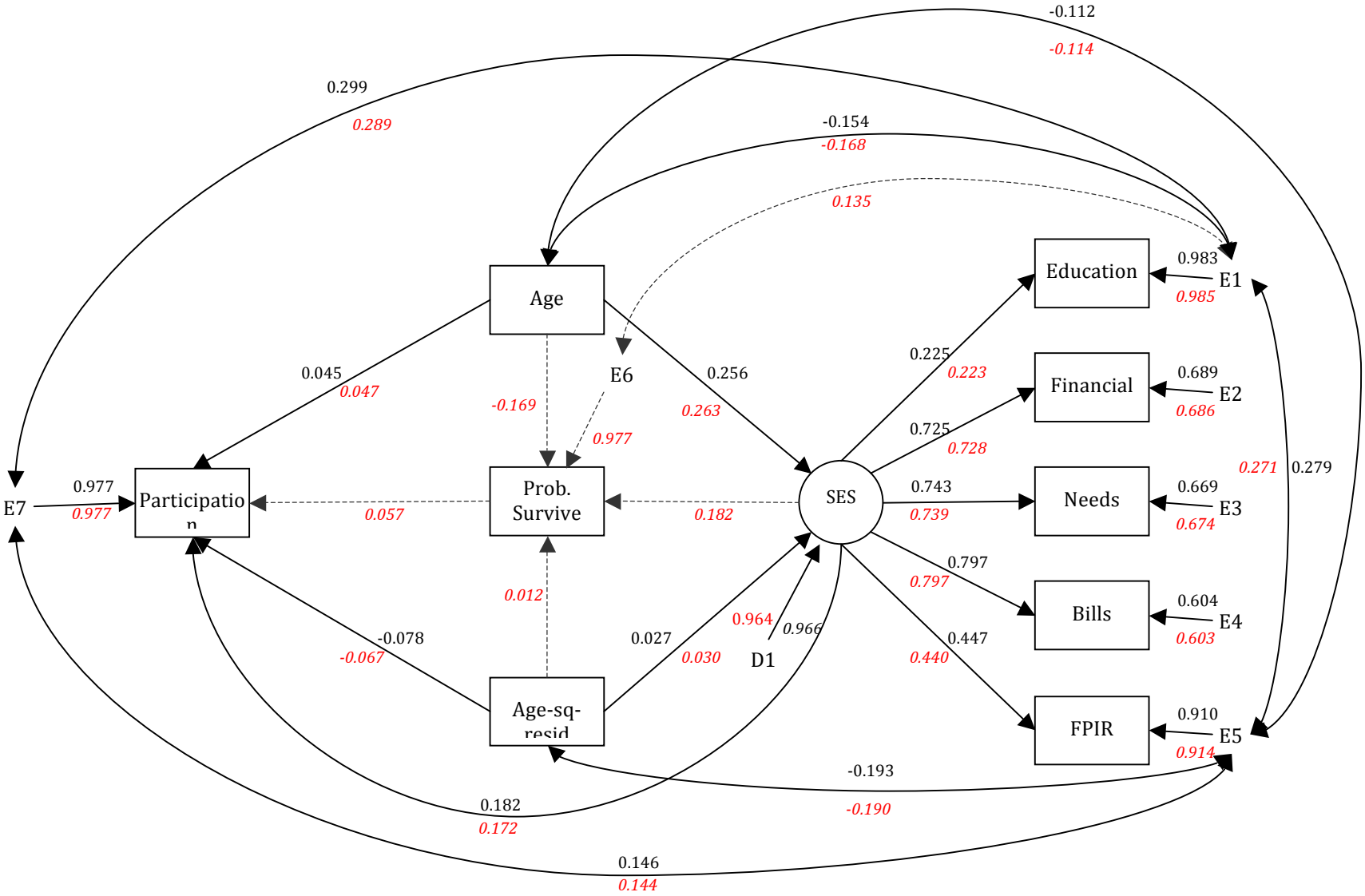


Table 6. Correlations of Variables Included in Structural Equation Models 1 and 2

	1	2	3	4	5	6	7	8	9
1. Participation	----	.10*	-.08*	.32*	.17*	.18*	.11*	.22*	.08*
2. Age	.07*	----	.00	-.10*	.21*	.14*	.21*	.03*	-.12*
3. Age-sq. Resid.	-.08*	.00	----	-.02	-.03*	.02	.05*	-.17*	.02
4. Education	.31*	-.11*	-.04*	----	.11*	.18*	.13*	.33*	.20*
5. Financial	.16*	.21*	-.03*	.10*	----	.53*	.58*	.33*	.11*
6. Needs	.18*	.14*	.02	.18*	.53*	----	.60*	.33*	.12*
7. Bills	.10*	.22*	.06*	.12*	.58*	.60*	----	.31*	.10*
8. FPIR	.23*	.00	-.17*	.34*	.32*	.33*	.29*	----	.14*
9. Prob. Survival	.09*	-.18*	-.02	.19*	.10*	.10*	.08*	.15*	----

* Correlation is significant at the .01 level

Note: Correlations between variables in complete sample in the lower triangle. Correlations between variables in sample of survivors in the upper triangle.

eliminates part of the variation extant in the world and so fabricates relationships that appear genuine to us because we study relationships using only survivors.

How much are researchers, then, overestimating the effect of age over participation due to selective mortality? Figure 4 shows the diagram of the two structural equation models. For comparison purposes, Model 1 estimates what is commonly reported in the literature while Model 2 corrects for health differences between survivors and non-survivors and for the probability of survival.[30] Standardized parameters are reported for comparability purposes, for Model 1 above the paths (in black) and for Model 2 under the paths (in red, *italic*). Table 6 reports the correlation matrices of the variables implemented in the analyses (Model 1 upper triangle, Model 2 lower triangle). Since Model 2 shows a somewhat large multivariate kurtosis, as indicated by Mardia's (1970) coefficient (= 16.4), robust methods are implemented to correct for non-normality using Satorra-Bentler test statistics (Satorra and Bentler 1994).[31]

[30] Be reminded that, by applying the matching output, differences in participation due to health differences between survivors and non-survivors already contracted by 56% (see above). Mediation effects are thus estimated strictly for individual differences in the probability of survival.

[31] Model 1 shows a low Mardia's coefficient of multivariate kurtosis of -1.6.

Both models show a good fit to the data. Model 1 produced a standardized root mean-squared residual (SRMR) of 0.021, a comparative fit index (CFI) of 0.980, and a root mean-squared error of approximation (RMSEA) of 0.050 [90% CI: 0.043 - 0.057]. Model 2 produced a SRMR of 0.024, a CFI of 0.974, and a RMSEA of 0.047 [90% CI: 0.041 - 0.053].[32] All parameters of interest are statistically significant ($p < 0.05$), with the exception of the direct effect of age-squared residuals over the probability of survival (Model 2), meaning that the variance of the probability of survival is mostly explained by the direct effect of the linear component of age, and by the indirect effects that both the linear and the quadratic components of age exert over it through SES. This suggests that SES helps to explain survival throughout the lifespan, but more exclusively survival onto the older years, one of our theoretical expectations. All total effects over participation in both models are statistically significant ($p < 0.05$).

Model 2 captures three mediation processes between the age terms and participation: (1) Mediation strictly through SES, (2) mediation through SES and through the probability of survival, and (3) mediation strictly through the probability of survival. Comparing the models, the indirect effect of both age terms over participation that go through SES remained unaltered (0.005 for age-squared and 0.045 for linear-age). This means that changes of the total effects of the age terms between the models are closely related to mediation in which the probability of survival is involved. Further, the positive sign and the size of these indirect effects suggest that, participation that emanates from SES variations related to age, continues to increase in the older years. Interestingly, the direct effect of linear-age and age-squared over participation went from

[32] Values of SRMR < 0.05, CFI > 0.95, and RMSEA < 0.06 indicate a good fit of the model (Bentler 2006; Hu and Bentler 1999). The SRMR is an informative index of model misspecification, very similar to the average absolute standardized residual (AASR), indicating how close, on average, is the model reproducing the correlation matrix. The CFI ranges from 0 to 1 and is an assessment of how good the hypothesized model fits the data compared to one of complete independence. Finally, the RMSEA is a fit statistic per degree of freedom that accounts for the number of cases used in the analysis (Hu and Bentler 1998; Ogasawara 2001a, 2001b).

0.047 to 0.054 and from -0.067 to -0.077, respectively, an overestimation of 15% in both instances. The reverse sign of these coefficients means that participation increases with age until the older years when it declines. The true pattern of participation of the elderly, therefore, depends on the mediation role of SES and not on aging alone.

Model 2 shows that both linear-age and SES have strong direct effects on the probability of survival (-0.169 and 0.182, respectively). Path coefficients show that, as SES increases with aging, both age terms have a positive impact on the probability of survival. The indirect effect of age-squared residuals on participation that goes through the probability of survival is practically zero, and of linear-age is 0.018. These results reaffirm the notion that selective mortality is mostly affecting the relation between age and participation during the years in which mortality differentials by SES are more salient.

The total effect (i.e., the sum of direct and indirect effects) of the linear component of age over participation estimated using Model 1 is 0.100 and using Model 2 is 0.085. This suggests that current research overestimates the increase of participation as people grow older by about 17% due to unaccounted mediation effects of health factors. The total effect of the quadratic component of age over participation using Model 1 is -0.072, and using Model 2 is -0.061. This suggests that current research overestimates the decline in participation in the older years by about 18%. These overestimations suggest that the true relationship between age and participation is less sloped than it appears to be. Taken together, these findings support the seventh hypothesis.

Conclusions

The analyses developed here show that mortality, and health factors related to mortality, affect the age-participation relationship in five different ways: (1) By excluding non-survivors

from samples, (2) by a heterogeneous distribution of health factors between survivors and non-survivors, (3) by individual differences in health, (4) because age-specific mortality influences the age composition of the age distribution of survivors, and (5) by intervening the effects of age on participation through three mediation mechanisms: (a) Mediation strictly through SES, (b) mediation through SES and through the probability of survival, and (c) mediation strictly through the probability of survival.

The findings of this article resonate with research that shows the selectivity of the older population with respect to their health status and SES (Crimmins, Kim and Seeman 2009; Liang et al. 2002; Ross and Wu 1996). Some of this research shows that high mortality associated with physiological dysfunction has precisely helped to maintain a relatively low level of physiological risk scores in the older population (Crimmins et al. 2003). This article shows that SES-related mortality operates in a similar way with respect to aging and participation. High mortality of low-level participants due to their low SES between middle to early old age artificially inflates the correlation between participation and age.

These findings also replicate what was expected from other lines of research. SES differences in mortality and health are low during young adulthood, are at a maximum level between middle to early old age, and then decline at the right tail of the human lifespan (Robert and House 2000; Seeman and Crimmins 2001). This demographic/epidemiologic process explains why selectivity on survival generates convergence between SES levels and health in the older years, when survivors become increasingly homogeneous due to mortality schedules. As individuals advance into late old age, comorbidity, frailty, and disability increase irrespective of socioeconomic factors. We all will eventually die regardless of our SES (Fried et al. 2001; House et al. 1990; Liang et al. 2002).

The analyses developed here also demonstrate how non-survivors have a reverse age-participation relationship compared to survivors, thus the bias due to selectivity. As non-survivors enter old age, differences in participation between survivors and non-survivors contract. The bigger the gaps between survivors and non-survivors in their early adulthood, the more radical the process of homogenization in the older-years. Initial differences are the cause of posterior similarities (c.f., Crimmins, Kim and Seeman 2009).

Consistent with the hypotheses, the main results of this article are as follows: (1) Results from propensity scores and matching procedures evinced that 56% of the difference in participation levels between survivors and non-survivors are related to health differences between these groups; (2) Regression analyses showed that the exclusion of non-survivors generates an overestimation of the effects of the linear and quadratic components of age on participation of 55% and 13%, respectively; (3) The probability of survival has a positive, independent effect over participation after controlling for age and socioeconomic status; should there be no detrimental individual differences in health, individuals would participate 28% more as they age; (4) Patterns of association between age and participation are reverse between survivors and non-survivors; maximum differences in participation between these groups are detected in their middle age years, where maximum differences in mortality due to SES differences manifest; (5) Survivors participate significantly more than future non-survivors across age groups and SES levels; (6) Differences in participation between survivors and non-survivors are underestimated by about 30% due to age compositional differences in the age distributions of these groups; the contribution to age-specific differences of crude participation rates that result from their age compositional differences between the ages of 40 and 54 years is 77%, between 55 and 64 years is 56%, and between 65 to 75 years is 95%; and (7) Total

mediation effects of the probability of survival evinced 17% and 18% overestimations of the linear and quadratic aspects of age effects on participation.

These findings signal the need to control for sample heterogeneity. It is customary to leave key intervening factors like mortality and health out of models of participation (and, indeed, out of many other models that attempt to explain socio-political processes). This article shows that recording the health status of participants at the baseline can help control for health factors. Another way to control for some of these confounding effects is by running mortality follow ups on social science surveys. Many pivotal surveys in social science, such as the American National Election Studies, The General Social Survey, and the Current Population Survey, among many others, do not include mortality follow ups as part of their research design. These efforts would increase the quality of the inferences derived from survey data analysis and contribute for the development of multidisciplinary research.

It is worth mentioning that the evidence developed here is indirect because of the nature of cross-sectional data and because of the methodological challenges that truncation by death brings to data analysis. Many socioeconomic differentials in health and mortality that are related to the socio-political world were left unexplored. Processes of social stratification, which are related to the world of politics in general, and to policy implementation in particular, suggest that causal directionality maybe subjected to endogeneity. Nevertheless, the presented models were parsimonious and yet fitted well theoretical expectations. The validities of the different pieces of evidence offered in this article favoring the hypotheses should be better interpreted in conjunction and not independently.

The analyses in this article were carried out under conservative circumstances. First, estimations ignore the selectivity effects of all pre-existent selective mortality already present in

the sample. Second, the health-related variation affecting the age-participation relationship in the models is years away from actual mortality. Third, findings are subjected to a short 10-year mortality follow up, which is a period much shorter than the actual range of time in which mortality influences the age-participation association. And finally, the analyzed sample was mostly composed of white individuals, had a higher SES than the average population, and was not representative of disadvantaged communities where mortality rates are higher. Taking these factors into consideration, we can gain confidence in that the depictions of the influence of selective mortality and health over the age-participation relationship as exemplified by the different methods and functional forms are in the right direction, and that evidence that emanates from them bring support to the hypotheses proposed in this study.

References

Adler, Nancy E., Thomas W. Boyce, Margaret A. Chesney, Sheldon Cohen, Susan Folkman, Robert L. Kahn, and S. Leonard Syme. 1994. "Socioeconomic Status and Health: The Challenge of the Gradient." *American Psychologist* 49 (1): 15-24.

Adler, Nancy E., Thomas W. Boyce, Margaret A. Chesney, Susan Folkman, and S. Leonard Syme. 1993. "Socioeconomic Inequalities in Health: No Easy Solution." *Journal of the American Medical Association* 269: 3140-3145.

Almond, Gabriel A., and Sidney Verba. 1963. *The Civic Culture.* Princeton, NJ: Princeton University Press.

Antonovsky, Aaron. 1967. "Social Class, Life Expectancy and Overall Mortality." *The Milbank Memorial Fund Quarterly* 45 (2): 31-73.

Arneson, Ben A. 1925. "Non-voting in a Typical Ohio Community." *American Political Science Review* 19 (November): 816-25.

Austin, Peter C. 2011. "An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies." *Multivariate Behavioral Research* 46: 399-424.

Bentler, Peter M. 2006. *EQS Structural Equations Program Manual.* Encino, CA: Multivariate Software.

Bentler, Peter M., and Ab Mooijaart. 1989. "Choice of Structural Model via Parsimony: A Rationale Based on Precision." *Quantitative Methods in Psychology* 106 (2): 315-17.

Blackwell, Matthew, Stefano Iacus, Gary King, and Giuseppe Porro. 2010. CEM: Coarsened Exact Matching in STATA. *Stata Journal* 9 (4): 524-546.

Blais, André. 2007. "Turnout in Elections." In *The Oxford Handbook of Political Behavior*, eds. Russell J. Dalton and Hans-Dieter Klingemann. New York: Oxford University Press.

Bollen, Kenneth A. and Shawn Bauldry. 2011. "Three Cs in Measurement Models: Causal Indicators, Composite Indicators, and Covariates." *Psychological Methods*, Vol. 16, No. 3: 265-284.

Bongaarts, John. 1994. "Population Policy Options in the Developing World." *Science* 263: 771-76.

Borugian, Marylin J., Samuel B. Sheps, Charmine Kim-Sing, Ivo A. Olivotto, Cheri Van Patten, Bruce P. Dunn, Andrew J. Coldman, John D. Potter, Richard P. Gallagher, and T. Gregory Hislop. 2003. "Waist-to-hip Ratio and Breast Cancer Mortality." *American Journal of Epidemiology* 158 (10): 963-68.

Bourgeois-Pichat, J. and Si-Ahmed Taleb. 1970. "Un Taux D'accroissement Nul Pour Les Pays en Voie de Développement en L'an 2000. Rêve ou Réalité?" *Population*, 25 (5): 951-74.

Brady, Henry E., Sidney Verba, Kay Lehman Schlozman. 1995. "Beyond SES: A Resource Model of Political Participation." *The American Political Science Review* 89 (2): 271-294.

Brim, Orville G., Paul B. Baltes, Larry L. Bumpass, Paul D. Cleary, David L. Featherman, William R. Hazzard, Ronald C. Kessler, Margie E. Lachman, Hazel Rose Markus, Michael G. Marmot, Alice S. Rossi, Carol D. Ryff, and Richard A. Shweder. (1999). *National Survey of Midlife Development in the United States (MIDUS), 1995-1996* [Computer file]. ICPSR02760-v7. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2011-02-18. doi:10.3886/ICPSR02760

Campbell, Angus, Philip E. Converse, Warren E. Miller, and Donald Stokes. 1960. *The American Voter*. New York: Wiley.

Campbell, Angus, and Robert L. Kahn. 1952. *The People Elect a President.* Ann Arbor, MI: Survey Research Center, Institute for Social Research.

Cho, Wendy K. Tam, James G. Gimpel, and Joshua J. Dyck. 2006. "Residential Concentration, Political Participation, and Voter Turnout". *The Journal of Politics* 68 (1): 156-67.

Cochran, William G. 1968. "The Effectiveness of Adjustment by Subclassification in Removing Bias in Observational Studies." *Biometrics* 24 (2): 295-313.

Collins, Selwyn D. 1935. "The Incidence and Causes of Illness at Specific Ages." *The Milbank Memorial Fund Quarterly* 13 (4): 320-38.

Connelly, Gordon M., and Harry H. Field. 1944. "The Non-Voter—Who He Is, What He Thinks." *Public Opinion Quarterly* 8 (2): 175-87.

Crimmins, Eileen M., Melanie Johnston, Mark Hayward, and Teresa Seeman. 2003. "Age Differences in Allostatic Load: An Index of Physiological Dysregulation." *Experimental Gerontology* 38: 731-34.

Crimmins, Eileen M., Jung K. Kim, and Teresa Seeman. 2009. "Poverty and Biological Risk: The Earlier "Aging" of the Poor." *Journal of Gerontology* 64A (2): 286-92.

Durkheim, E. 1955. *Suicide*. New York: Free Press.

Erbe, William. 1964. "Social Involvement and Political Activity: A Replication and Elaboration." *American Sociological Review* 29 (2): 198-215.

Fried, Linda P., Catherine M. Tangen, Jeremy Walston, Anne B. Newman, Calvin Hirsch, John Gottdiener, Teresa Seeman, Russell Tracy, Willem J. Kop, Gregory Burke, and Mary

Ann McBurnie. 2001. “Frailty in Older Adults: Evidence for a Phenotype.” *Journal of Gerontology* 56A (3): 146-156.

Gelman, Andrew, and Jennifer Hill. 2007. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. New York, New York: Cambridge University Press.

Glenn, Norval D., and Michael Grimes. 1968. “Aging, Voting, and Political Interest.” *American Sociological Review* 33 (4): 563-575.

Gosnell, Harold F. 1927. *Getting Out the Vote: An Experiment in the Stimulation of Voting.* Chicago: University of Chicago Press.

Hardin, Andrew and George A. Marcoulides. 2011. “A Commentary on the Use of Formative Measurement.” *Educational and Psychological Measurement*, Vol. 71, No. 5: 753-764.

House, James S. Ronald C. Kessler, and A. Regula Herzog. 1990. “Age, Socioeconomic Status, and Health. *The Milbank Quarterly* 68 (3): 383-411.

Hout, Michael, and David Knoke. 1975. “Change in Voting Turnout, 1952-1972.” *The Public Opinion Quarterly* 39 (1): 52-68.

Hu, Li-tze, and Peter M. Bentler. 1998. “Fit Indices in Covariance Structural Modeling: Sensitivity to Underparameterized Model Misspecification.” *Psychological Methods* 3 (4): 424-53.

———. 1999. “Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria Versus New Alternatives.” *Structural Equation Modeling* 6 (1): 1-55.

Iacus, Stefano M., Gary King, and Giuseppe Porro. 2008. “Matching for Causal Inference without Balance Checking.” Harvard University. Typescript.

Junn, Jane. 2010. “On Participation.” In *The Oxford Handbook of American Elections and Political Behavior,* ed. Jan E. Leighly. Oxford, UK: University Press. Online: http://dx.doi.org/10.1093/oxfordhb/9780199235476.001.0001

King, Gary, and Langche Zeng. 2006. “The Dangers of Extreme Counterfactuals.” *Political Analysis* 14: 131-159.

Kitagawa, Evelyn M. 1955. “Components of a Difference between Two Rates.” *Journal of the American Statistical Association* 50 (272): 1168-94.

———. 1964. “Standardized Comparisons in Population Research.” *Demography* 1: 296-315.

Kitagawa, Evelyn M., and Philip Hauser. 1968. “Education Differentials in Mortality by Cause of Death: United States, 1960.” *Demography* 5 (1): 318-353.

———. 1973. *Differential Mortality in the United States: A Study in Socioeconomic Epidemiology*. Cambridge: Harvard University Press.

Leighley, Jan E. 1995. "Attitudes, Opportunities and Incentives: A Field Essay on Political Participation." *Political Research Quarterly* 48 (1): 181-209.

Liang, Jersey, Joan Bennett, Neal Krause, Erika Kobayashi, Hyekyung Kim, J. Winchester Brown, Hiroko Akiyama, Hidehiro Sugisawa, and Arvind Jain. 2002. "Old Age Mortality in Japan: Does the Socioeconomic Gradient Interact with Gender and Age?" *Journal of Gerontology* 57B (5): 294-307.

Lijphart, Arend. 1997. "Unequal Participation: Democracy's Unresolved Dilemma." *American Political Science Review* 91 (1): 1-14.

Lipset, Seymour Martin. 1960. *Political Man: The Social Bases of Politics.* Garden City, NY: Doubleday.

Mardia, Kanti V. 1970. "Measures of Multivariate Skewness and Kurtosis with Applications." *Biometrika* 57 (3): 519-30.

Mattei, Alessandra. 2009. "Estimating and Using Propensity Score in Presence of Missing Background Data: An Application to Assess the Impact of Childbearing on Wellbeing." *Statistical Methods and Applications* 18: 257-273.

Milbrath, Lester W. 1965. *Political Participation; How and Why do People Get Involved in Politics?* Chicago: Rand McNally

Neugarten, Bernice Levin. 1968. *Middle Age and Aging: A Reader in Social Psychology*. Chicago: University Press.

Nie, Norman H., Sidney Verba, and Jae-on Kim. 1974. "Political Participation and the Life Cycle." *Comparative Politics* 6 (3): 319-40.

Ogasawara, Haruhiko. 2001a. "Standard Errors of Fit Indices Using Residuals in Structural Equation Modeling." *Psychometrika* 66 (3): 421-36.

———. 2001b. "Approximations to the Distributions of Fit Indices for Misspecified Structural Equation Models." *Structural Equation Modeling* 8: 556-74.

Olsen, Marvin E., and Judy Corder Tully. 1972. "Socioeconomic-Ethnic Status Inconsistency and Preference for Political Change." *American Sociological Review* 37 (5): 560-74.

Pappas, Gregory, Susan Queen, Wilbur Hadden, and Gail Fisher. 1993. "The Increasing Disparity in Mortality Between Socioeconomic Groups in the United States, 1960 and 1986." *New England Journal of Medicine* 329: 103-109.

Plutzer, Eric. 2002. "Becoming a Habitual Voter: Inertia, Resources, and Growth in Young Adulthood." *American Political Science Review* 96 (1): 41-56.

Preston, Samuel H., Patrick Heuveline, and Michel Guillot. 2001. *Demography: Measuring and Modeling Population Processes.* Oxford, UK: Blackwell Publishers.

Robert, Stephanie A., and James S. House. 2000. "Social Inequalities in Health: An Enduring Sociological Problem." In *Handbook of Medical Sociology*, 5th ed, eds. Bird, Chloe E., Peter Conrad, and Allen M. Fremont. Upper Saddle River, New Jersey: Prentice Hall.

Rosenbaum, Paul R. 2010. *Design of Observational Studies*. New York, New York: Springer.

Rosenbaum, Paul R., and Donald B. Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika* 70 (1): 41-55.

Rosenston, Steven J. and John Mark Hansen. 1993. *Mobilization, Participation, and Democracy in America.* New York: Macmillan.

Ross, Catherine E., and Chia-Ling Wu. 1996. "Education, Age, and the Cumulative Advantage in Health." *Journal of Health and Social Behavior* 37: 104-20.

Satorra, Albert. 1992. "Asymptotic Robust Inferences in the Analysis of Mean and Covariance Structures." *Sociological Methodology* 22: 249-78.

Satorra, Albert, and Peter M. Bentler. 1994. "Corrections to test statistics and standard errors in covariance structure analysis." In *Latent Variables Analysis: Applications for Developmental Research,* eds. Alexander *v*on Eye and Clifford C. Clogg. Thousand Oaks, CA: Sage.

Seeman, Teresa, Burton Singer, Charles W. Wilkinson, and Bruce McEwen. 2001. "Gender Differences in Age-related Changes in HPA Axis Reactivity." *Psychoneuroendocrinology* 26 (3): 225-40.

Seeman, Teresa, Eileen Crimmins, Mei-Hua Huang, Burton Singer, Alexander Bucur, Tara Gruenewald, Lisa F. Berkman, and David B. Reuben. 2004. "Cumulative Biological Risk and Socio-economic Differences in Mortality: MacArthur Studies of Successful Aging." *Social Science & Medicine* 58: 1985-1997.

Seeman, Teresa, Elissa Epel, Tara Gruenawald, Arun Karlamangla, and Bruce McEwen. 2010. "Socio-economic Differentials in Peripheral Biology: Cumulative Allostatic Load." *Annals of the New York Academy of Sciences* 1186: 223-239.

Sekhon, Jasjeet S. 2007. "Multivariate and Propensity Score Matching Software with Automated Balance Optimization: The Matching Package for R." *Journal of Statistical Software* VV (II): 1-47.

Steiger, James H., Alexander Shapiro, and Michael W. Browne. 1985. "On the Multivariate Asymptotic Distribution of Sequential Chi-square Statistics." *Phsychometrika* 50 (3): 253-63.

Sydenstricker, Edgar. 1928. "The Prevalence of Ill Health." *Bulletin for the New York Academy of Medicine* 4 (2): 191-215.

Taylor, Shelley E., Rena L. Repetti, and Teresa Seeman. 1997. "Health Psychology: What is an Unhealthy Environment and How Does it Get Under the Skin." *Annual Review of Psychology* 48: 411-447.

Tingsten, Herbert. 1937. *Political Behavior: Studies in Election Statistics.* London: P.S. King & Son.

Treiblmaier, Horst, Peter M. Bentler, and Patrick Mair. 2011. "Formative Constructs Implemented via Common Factors." *Structural Equation Modeling*, Vol. 18: 1-17.

Verba, Sidney, and Norman H. Nie. 1972. *Participation in America: Political Democracy and Social Equality*. New York: Harper & Row.

Verba, Sidney, Norman H. Nie, and Jae-On Kim. 1978. *Participation and Political equality: A seven-Nation Comparison.* Cambridge: Cambridge University Press.

Welch, Susan. 1990. "The Impact of At-Large Elections on the Representation of Blacks and Hispanics." *Journal of Politics* 52: 1050-1976.

Wolfinger, Raymond E., and Steven J. Rosenstone. 1980. *Who Votes?* New Heaven, CT: Yale University Press.

Zhang, Junni L., and Donald B. Rubin. 2003. "Estimation of Causal Effects via Principal Stratification When Some Outcomes are Truncated by 'Death.'" *Journal of Educational and Behavioral Statistics* 28 (4): 353-68.

DIFFERENTIAL MORTALITY AND THE RACIAL COMPOSITION OF THE ELECTORATE

Javier M. Rodriguez
Department of Political Science
University of California, Los Angeles

Abstract

I set forth a conceptual model for the impact of mortality differentials between Blacks and Whites from 1970 to 2004 on the racial composition of the electorate in the presidential election of 2004. I use data from the Multiple Cause of Death files, the U.S. Census Bureau, the National Election Pool General Election Exit Polls, and the U.S. Elections Project in order to carry out four calculations: (1) The total number of excess deaths among Blacks due to mortality gaps between 1970 and 2004, (2) the probability that the excess deceased would have turned out to vote in the presidential election of 2004 had they survived, (3) the total number of hypothetical Black survivors should they had the mortality rates of Whites, and (4) the total number of Black votes lost to mortality gaps in the presidential election of 2004. Results show that there were 2.7 million excess deaths among Blacks due to mortality gaps between 1970 and 2004. Of those 2.7 million, 2.3 would have survived until 2004 should they had the mortality rates of Whites. These excess deaths resulted in 1.2 million Black votes lost in the presidential election of 2004. These findings suggest that mortality gaps between Blacks and Whites are affecting political determinants of election outcomes and the distribution of power that emanates from the political system.

Introduction

Humans need to be alive and in good health in order to participate in politics. Since the deceased cannot vote, and they are not represented in the electoral system, excess deaths exacerbate the imbalance of political power, especially between racial communities where health inequalities are most pronounced.

Much cumulative evidence in the health sciences speaks to the demographic contraction and health disadvantages of African Americans compared to Whites. Much less research has been developed in political science using this information, despite the fact that mortality gaps between African Americans and Whites determine the demographic structure of the United States and, accordingly, many basic features of our electoral democracy.

Contrary to the expectations of policy makers and political actors, there is no clear improvement of health disparities between Blacks and Whites in recent decades. This evidence has been found to be robust in spite of the enhancement in education and income attainment that resulted from the Civil Rights Movement. For instance, the standardized mortality rate (SMR) for African Americans was 1.47 in 1960 and 1.41 in 2000 (Satcher et al. 2005). This does not only mean that SMRs for African Americans have remained stable through the last decades, but also that African Americans still experience about 40% more deaths than expected if they had the mortality rate of Whites. On average, throughout the 20th century, the mortality rate of African Americans was about 60% greater than that of Whites (Kaufman et al. 1998).

The bulk of the evidence in scholarly research shows that Black-White gaps are substantial and persistent, and, in fact, that they have increased across numerous indicators of health (Dominguez et al. 2008). For example, the gap in standardized infant mortality rates between African Americans and Whites increased from 2.02 in 1960 to 2.52 in 2000 (Satcher et

al. 2005). Similarly, from 1979 to 1998, Black-White age- and sex-adjusted mortality ratios increased in eight of the nine leading causes of death responsible for 83% of all mortality in the United States (Levine et al. 2001). These types of disparities do not only refer to differences in the quality of life of individuals but also to the access that some communities have to public goods and services. For example, it was not until 1990 that the African American community reached the life expectancy that Whites had in 1950, and it was not until 1995 that African American males reached a life expectancy equal to the official age for retirement benefits determined since the 1930s.

Taken together, these statistics indicate that African Americans have historically died in greater proportions compared to Whites and that such excess deaths may have been translated into a demographic contraction that could constitute a non-trivial vanishing electorate. This possibility defines the main objective of this study: To estimate the impact that the excess number of deaths among African Americans, linked to historical mortality rate differences with Whites, has over the racial composition of the electorate in presidential elections—in specific, the presidential election of 2004.

Theoretical Framework

About 93% of all mortality that occurred in 2004 in the United States is attributable to health factors. The remaining 7% is attributable to other sources of death such as accidents (4.9%), suicide (1.4%), assault (0.7%), and all other external causes (0.2%). Health outcomes are, therefore, the major driving force behind mortality schedules in the United States. A great portion of the mortality gap between Blacks and Whites is due to the unequal distribution of morbidity between these races. The causes for said disparities are multiple and complex. Health science literature has identified pivotal factors that explain a good portion of the variation in

racial health disparities and has proposed some of the mechanisms through which social conditions "get under the skin".

There is a robust and powerful inverse association between socioeconomic status and health status that can be traced back to our earliest records (Williams and Collins 1995). Individuals with lower levels of education and income, those who work in poor occupational environments, and those who live in poverty and reside in highly segregated areas, among other indicators, have persistently shown higher levels of morbidity and consequently of mortality (Adler et al. 1994; Antonovsky 1967; Clougherty, Souza, and Cullen 2010; Grossack 1963; Kitagawa and Hauser 1973; Pappas et al. 1993; Seeman et al. 2010; Williams 1990; Winkleby et al. 1992). In the United States, for example, the difference in life expectancy at birth between the most advantaged and disadvantaged is about 20 years (Marmot 2005).

Residential segregation is related to the perception of safety, which in turn correlates with physical activity. Poor neighborhoods offer low quality health care services, higher exposure to pollution, scarce recreational facilities, a higher cost of healthy products and a prevalence of aggressive marketing investment by the tobacco and alcohol industries. These disadvantages of residential segregation are strongly correlated with race, especially among Blacks (Massey 2004; Williams and Jackson 2005) and have remained so for decades. For example, in their influential article, McCord and Freeman (1990) showed that in 1980 young black males in Harlem, New York City, were more likely to die before the age of 65 than young men in Bangladesh. More recent analyses of the 171 most important cities in the U.S. show that the most decayed urban environment in which Whites live is better than the average Black neighborhood (Sampson and Wilson 1995; Williams and Jackson 2005).

High disparity levels between African Americans and Whites are also present in the access to and quality of resources, including medical care. This is especially true across indicators such as health insurance coverage and preventive screening, diagnostic, treatment, and rehabilitation services (Williams and Mohammed, 2009). The low quality and lack of information, services, and attention provided to the Black community, confounded with socioeconomic and residential disadvantages, produce behaviors and practices that affect their health. Poor diet and malnutrition, the lack of physical activity, as well as the consumption of tobacco are among the most common (c.f. Taylor, Repetti, and Seeman 1997). Results from studies that examine behaviors that affect the health of the individual are not to be interpreted in isolation. To an important extent, they are the product of the confounded effects of differences in socioeconomic status and segregation.

Not surprisingly social factors explain a good portion of the variation manifested in many relevant biological indicators of health. Hierarchical social structures, as we know them, exist on the basis of a disproportionate distribution of resources and positive social values (e.g., wealth, power, prestige) across primarily differentiated social groups; for example, African Americans and Whites (Geronimus 2003; Pratto et al. 1994). Said distribution of social assets, the access individuals have to them, and the way these resources mold the daily experiences of individuals affect developmental factors across the lifespan, including the psychobiological and psychosocial pathways that interact with the environment and lead to disease (Adler et al. 1999). Constraints in the environment shape the way individuals adapt to challenges and consequently their endocrine responses to stress. The repetitive process through which an individual responds to a stressor and returns to equilibrium produces a cumulative physiological toll termed Allostatic Load (Crimmins et al. 2003; McEwen and Seeman 2006). Persistent challenges throughout the

life-course make the physiological systems of the body operate out of normal ranges, thus causing adverse health outcomes. Differences in Allostatic Load contribute to the explanation of how social conditions affect health and biological aging (Seeman et al. 2010) including differences on the basis of race (Geronimus et al. 2006).

Racial disparities in health may have a political origin, to the extent they concur more often than not with the processes that surround racial prejudice and discrimination. As social stratification processes—which we know are tempered by policy and political power—sort individuals into different socioeconomic strata on the basis of race, a disproportionate number of Blacks are exposed to the stress that emanates from social disadvantages, thus creating racial disparities in health.

Factors associated with differences in socioeconomic status, geographic or residential segregation, access to public goods and resources as well as high quality information, the composition of appropriate networks and work conditions, and decision-making processes and behaviors constitute the most relevant and persistent forces behind racial health disparities over and beyond biomedical predictors of health (Adler et al. 1993, 1994; Antonovsky 1967; Wight et al. 2010; Massey 2004). And, as it is well known, structural and individual discrimination affect most of these social determinants of health (Williams, Neighbors, and Jackson 2003). Historical health disparities between African Americans and Whites are a phenomenon mostly driven by social, not genetic, determinants of health (Laveist 2002; Sankar et al. 2004).

Disparities in health tend to be embedded in the three main demographic groups: age, sex/gender, and race/ethnicity. Health scientists have studied health differences across demographic groups with a focus on their biological origins. Health differences between the young and the elderly, linked to the physiological decline that increases with chronological

aging, are evinced by the incidence of frailty, comorbidity, and disability in the older years of the human lifespan (Fried et al. 2001). Disparities between women and men are often attributed to the physiological differences between their reproductive systems and to differences in cellular biochemistry related to XX/XY-chromosome imprinting (Wizemann and Pardue 2001). Much of recent research, however, shows that health differences between the races are only partially traced to genetic causes.

Recent research shows that environmental components of health disparities on the basis of race are stronger than genetic components. For example, Kittles and Weiss (2003) find that within-racial group genetic differences are higher than between-group differences. Samples including West Africans and African-origin Caribbeans, whose genetic origins resemble those of African Americans in the United States, show hypertension and diabetes prevalence rates two to five times lower than those of African Americans in the U.S.; that is, similar or better than those of U.S.-born Whites (Cruickshank et al. 2001). Dressler et al. (2005) find that immigrants of all racial/ethnic backgrounds in the U.S. have a better health standing than their same race/ethnic peers who were born in the U.S., and that health risks for these immigrants across many indicators of health increase with time of residence and generational status. These findings highlight the centrality of environmental factors affecting racial disparities in health and, therefore, open the possibility that part of the disparities observed between the races may be rooted in the political processes that surround racial prejudice and discrimination.

It is important to notice that most of the well-identified social determinants of health, found in health sciences literature, are deeply affected by politics. Political processes associated with racial prejudice are crucial drivers of inequalities in socioeconomic status, geographic or residential segregation, access to and quality of resources like medical care, income and

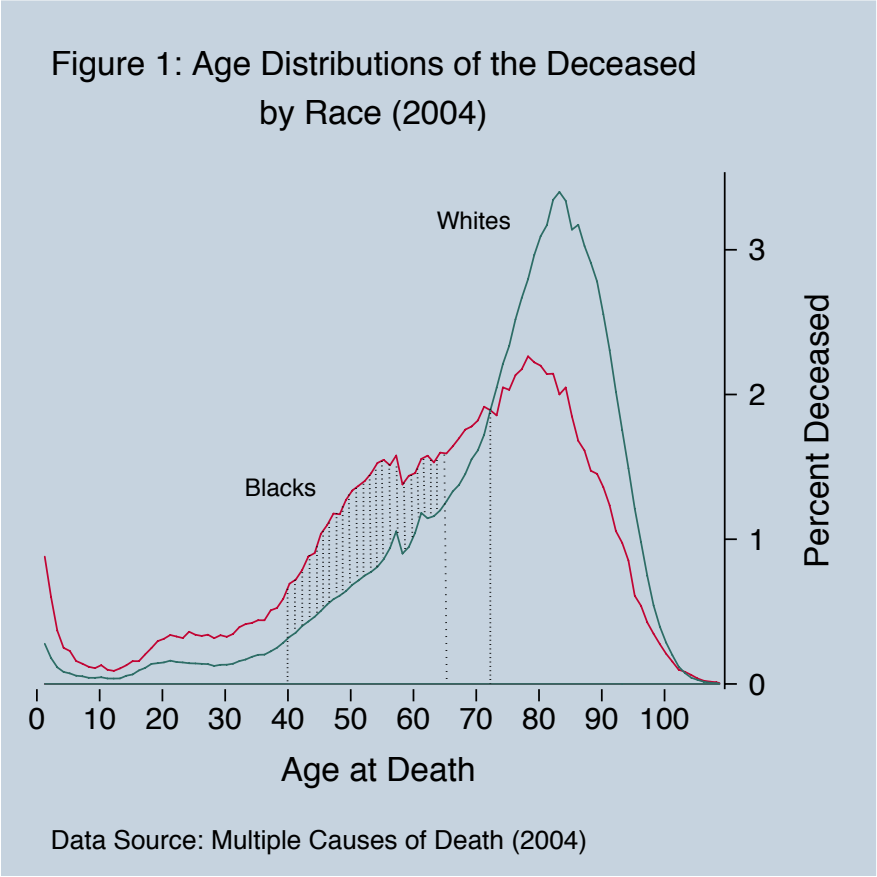
education. There is abundant evidence that racial prejudice is a political force that affects political attitudes (Henry and Reyna 2007; Sears and Kinder 1971); candidate preferences (Bobo and Dawson 2009; Valentino and Sears 2005); political ideology (Lane and Jost 2011; Pratto et al. 1994); racial and non-racial policy preferences (Rabinowitz et al. 2009; Tesler and Sears 2010); public opinion (Mendelberg 2008; Valentino et al. 2002); and political behavior (Enos 2011; Kinder and Sears 1981; Sidanius and Pratto 2001). Considering that racial disparities in health have traditionally been the subject matter of public health, epidemiology, and medical researchers, political processes surrounding racial prejudice and discrimination may constitute an important, yet overlooked, component of racial disparities in health.

Racial Mortality Gaps and Electoral Politics

A good number of forces that increase the probability of survival are the same ones that increase the probability of participating in the political system, including income and education. Considering that all of these variables are assessed on survivors alone, political science research has ignored the selective mortality process linked to health disparities that allows individuals—but not all individuals—to have an opportunity to cast their votes. Racial disparities in health may not only be caused by political forces but may also have explicit political consequences that are mediated by the process of electoral representation.

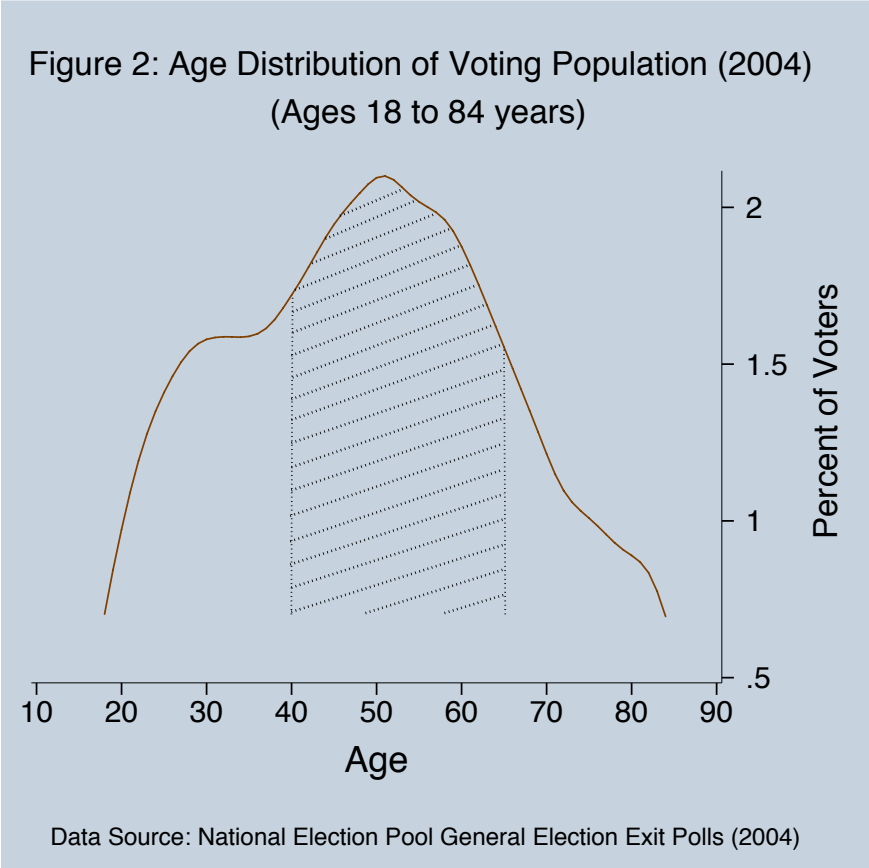
Figure 1 shows the age distributions of all individuals who died in 2004 by race. The area between the curves represents the mortality gap between Blacks and Whites. Notably, the distributions do not intersect until approximately the age of 73 years, evincing that mortality gaps between Blacks and Whites persist throughout most of the lifespan of Blacks.[33] Figure 1

[33] Life expectancy for Blacks in 2004 was, coincidentally, about 73 years, with an average of about 64 years for the period between 1970 and 2004.



also shows that the mortality gap between Blacks and Whites is greatest between the ages of 40 and 65, approximately. Research in social epidemiology has identified this period of the human lifespan as the one where mortality differentials due to inequalities in socioeconomic status are most strongly manifested (Seeman and Crimmins 2001).

This observation is important because, first, this is the age range in which the probability of turning out to vote is the highest, and, second, because an important portion of the voting population are individuals in this age range (Figure 2). This means that the mortality gap between Blacks and Whites creates a participatory disadvantage to the Black population, because they are dying off from the electorate at higher rates than Whites during the years in which they can make the most out of their voter turnout. In other words, Blacks are suffering a double negative effect from mortality gaps: They are not only losing members of their community in



excess compared to Whites, but they are also losing a good portion of their high-participation years of life.

The processes under observation are captured by the shape of the age distributions of both the deceased (i.e., those who disappeared from the electorate) and of voters (i.e., survivors). Figures 1 and 2 show that the age distribution of the deceased is skewed to the left, while the age distribution of the voting population is slightly skewed to the right, respectively. This difference in skewness exemplifies the dynamics between racial mortality gaps and voter turnout: Black individuals mostly die off from the voting-age distribution through the right tail of the distribution (most excessively between the ages of 40 and 65), while they can exclusively enter the voting-age distribution through the left tail at 18 years of age. Succinctly, in the Black community, the dynamics between mortality schedules and participation generate a rotation of

the population in which older high-level participants are being traded for young, low-level participants at a higher rate than Whites.

These observations bring to light the fact that the effects of mortality gaps between the races are not stationary but dynamic, and that they accumulate through time. This situation requires the contemplation of some empirical implications and theoretical considerations that pertain to possible disparities in the racial composition of the electorate that could have impacted the political atmosphere. The reason, as is evident, is that many individuals die before the age of 18 and consequently never joined the voting-age distribution of the population (i.e., never had the opportunity to vote). Following this line of reasoning, individuals who die after the age of 18, but before the age of life expectancy of their cohort, have a shorter "electoral life" than those who culminate their normal life cycle, thus diminishing the aggregate opportunity-years of their group to participate in politics. Higher premature death among the Black community represents a higher number of political-life years lost in comparison to Whites.

Premature death, therefore, does not only prevent individuals from voting in the presidential election following their date of death but also prevents them from voting in all subsequent presidential elections, making the effect cumulative. It is true that the group of Black individuals who, for example, died at age 45, in Michigan, in 1970, did not have the opportunity to vote in the presidential election of 1972. But it is also true that, should their death have been avoided in 1970, some of them would have survived to age 51 and have had the opportunity to participate in the presidential election of 1976. Similarly, some of them would have further survived to have the opportunity to vote in the presidential election of 1980, and some others to 1984, and so on. In other words, Blacks who die prematurely, and in excess compared to Whites, constitute a vanishing electorate. The political effect of premature death is cumulative, thus

diluting their political voice and putting them in both a representative and participatory disadvantage compared to Whites throughout time. Considering that representativeness, participation, and empowerment are some of the key ingredients of a healthy democracy, the demographic contraction of Blacks in comparison to Whites due to racial disparities in mortality schedules also carries democratic normative implications.

Research Design

This study proposes the counterfactual "what if" Blacks had survived at the same rates of Whites between the years 1970 and 2004. The basis of the counterfactual relies in the theoretical scenario in which Blacks are not exposed to their observed mortality rates between 1970 and 2004 but exposed to the mortality rates experienced by Whites in the same period. The construction of such counterfactual condition is done via stratification. The motivation behind the stratification of the data is that stratification helps with confounding, selection bias, and the validity of the evaluation and description of effect-measure modifications (Rothman and Greenland 1998). This hypothetical scenario inspires the research question of this study: How much did mortality gaps between Blacks and Whites from 1970 to 2004 affect the racial composition of the electorate in the presidential election of 2004?[34]

In order to answer this question it is needed estimate the following calculations: (1) The total number of Black excess deaths due to mortality gaps between 1970 and 2004; (2) The probability that the excess-deceased would have turned out to vote in the presidential election of 2004 had they survived; (3) From those who died in excess between 1970 and 2004, the total number of individuals who have hypothetically survived until 2004; and (4) The total number of Black votes lost to mortality gaps.

[34] The case study of the presidential election of 2004 is selected because it allows to capture the full cumulative effects of racial mortality gaps extant in the whole range of the data.

The observations come from four datasets. The mortality data are from the Multiple Cause of Death files (1970-2004), which records all deaths occurring within the United States. The data are based on information collected from death certificates. Approximately 73 million individuals died between 1970 and 2004, and this is the number of cases I analyze in this study. The population data are from the U.S. Census Bureau (various years from 1970-2004). The counted vote data are from the U.S. Elections Project, an institution specialized on reporting turnout statistics. And finally, the voter turnout data are from the National Election Pool General Election Exit Polls (2004). Their voter turnout data is comprised of representative samples in all 50 states and the District of Columbia. The interviews were conducted with voters as they left their polling places on elections day. The total sample size is 77,006 interviews.

I take advantage of the features of these data to develop four non-parametric models (one for each calculation enumerated above), which provide a high degree of validity especially as calculations using the Multiple Cause of Death, the U.S. Elections Project, and the U.S. Census data are representative of their respective universes, not samples. All calculations are carried out on the basis of sex, non-Hispanic race (i.e., non-Hispanic Black and non-Hispanic White), age (0 to 84 years), state of residence (35 states in total)[35], and year (1970 to 2004). These data are implemented according to the information available in death certificates.

Excess Black Deaths

"Excess deaths" makes reference to the number of deaths that results from differences in mortality rates between Blacks and Whites—this is to say, from mortality gaps. In order to create

[35] My calculations include 35 states, thus excluding the following 15 states: Alaska, Hawaii, Idaho, Maine, Montana, Nevada, New Hampshire, New Mexico, North Dakota, Oregon, Rhode Island, South Dakota, Utah, Vermont, and Wyoming. The exclusion of these states is on the basis of their low Black population, thus impeding valid estimations. Put together, the total Black population of these states constitutes about 1% of the national Black population; therefore, my calculations include about 99% of the national Black population in any given year between 1970 and 2004.

the counterfactual scenario, the mortality rate of Whites is directly applied to the Black population, which is then subtracted from the original Black mortality rate. The resulting number represents the number of Black individuals that should not have died if they had the mortality rates of Whites between 1970 and 2004. It is important to note that Black excess deaths constitute by itself an assessment of racial disparities in health. This *Total Black Excess Deaths* can be formally represented as follows:

$$TED^{B}_{(1970-2004)} = \sum_{Y=1970}^{2004} \sum_{Z=1}^{35} \sum_{S=1}^{2} \sum_{I=0}^{84} Pop^{B}_{i,s,z,y} \underbrace{\left(q^{B}_{i,s,z,y} - q^{W}_{i,s,z,y}\right)}_{Mortality-Gap} \qquad (1)$$

Where Y is the year of death (1970-2004); Z is the state of residence at death (1-35); S is the sex/gender group (1 = Male, 2 = Female); I is the individual's age at death (0 to 84 years); Pop^{B} is the *Black* population; q^{B} is the mortality rate of *Blacks*; and q^{W} is the mortality rate of *Whites*. The racial mortality rate is calculated by dividing the total number of deaths over the respective population for each and all levels of I, S, Z, and Y. The product between the Black population and the mortality gap equals the number of excess deaths. Thus, in essence, Equation 1 sums the excess number of Black deaths for each age, for both males and females, across all 35 states, from 1970 to 2004.

Probability of Turning Out to Vote in the Presidential Election of 2004 for Blacks

Taking into consideration well known inconveniences related with self-reported turnout (Burden 2000; Holbrook and Krosnick 2010; Vavreck 2007), I take advantage of the size of my sets of data to estimate a "physical" rate of turnout. The challenge, of course, is on the estimation of the total number of votes cast by a minority group like Blacks at all levels of I, S, and Z in the

presidential election of 2004. The probability of turning out to vote of those Black individuals who died in excess between 1970 and 2004, but who would have survived to the presidential election of 2004 should they have had the mortality rates of Whites, is calculated using the equation:

$$\Pr(V=1)^{B}_{z,s,g} = \left(\frac{V^{B}_{z,s,g}}{VAP^{B}_{z,s,g}}\right) = \left(\frac{E^{B}_{z,s,g} \bullet TCV_{z}}{VAP^{B}_{z,s,g}}\right) \qquad (2)$$

Where z is the state (1-35); s is sex/gender group (1 = Male, 2 = Female); g is the age group (1 = 18-29, 2 = 30-44, 3 = 45-64, and 4 = 65-84 years); V^B is the total number of votes cast by Blacks; E^B is percent of people exiting from the polls that were Black; TCV is the total counted vote; and VAP^B is the voting-age Black population.

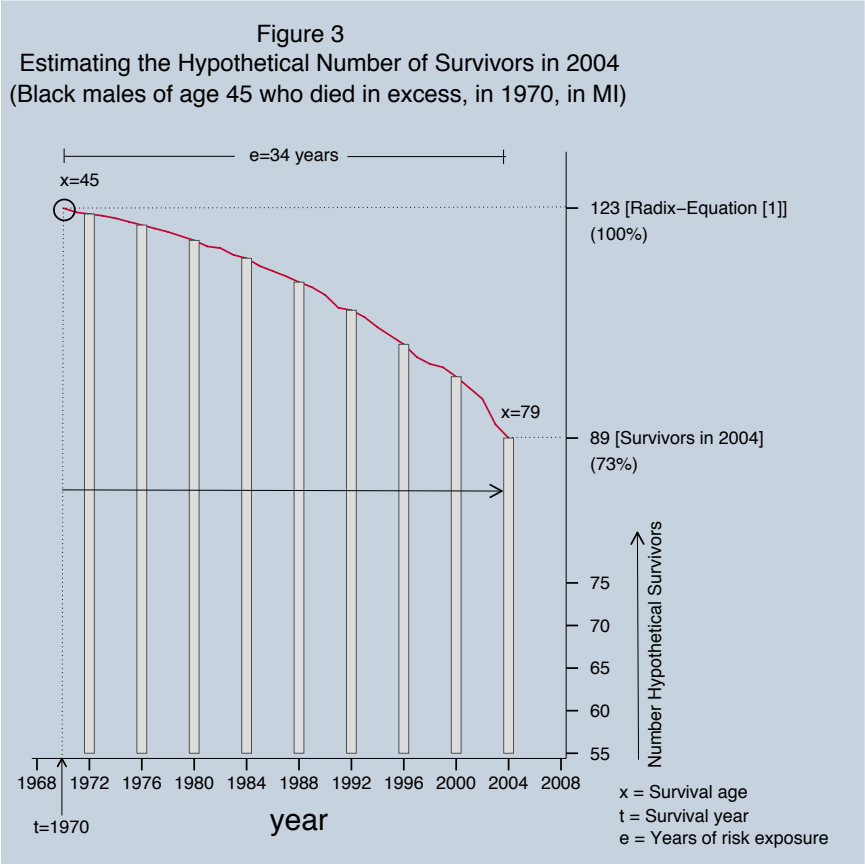
Note that equation 2 is only partially correcting for eligibility. Among those who are not eligible to vote are ex-felons (depending on state law), non-citizens, the mentally incompetent, and individuals who do not meet state residency requirements. With respect to the Black population, ineligibility is particularly important due to felony disfranchisement. In 2004, an estimated 9.3% of the voting-age population was ineligible to vote, from which 1.4 million were Black male ex-felons (McDonald 2004). Despite the fact that both terms E^B and TCV inherently take into consideration eligibility (for we know they voted, so they should have been eligible), the term in the denominator VAP^B does not correct for it. A recent estimation states that, by using the voting-age population instead of the voting eligible population, the national voter turnout rate in 2004 shows an underestimation of the turnout rate of 5.3% (McDonald 2004). Consequently, the denominator in Equation 2 is larger than it should be, and my calculation of the probability of turning out to vote for Blacks is, therefore, a conservative one.

Another consideration is that the probability of turning out to vote varies by cohort. Let us imagine Black males of age 45 who died in excess, in 1970, in Michigan. Since the population is truncated by death, the only (and most realistically close) data available to estimate their turnout rate are that of their very same age, state, sex, and race. Accordingly, I assume that their probability of turning out to vote in the 2004 presidential election would be that of their surviving cohort; that is, in the example above, those who by 2004 were 79-year-old Black males in Michigan. Succinctly, under the counterfactual condition in which they would have survived at the mortality rates of Whites, the model assumes that hypothetical Black survivors would have been similar to those Blacks who did not die. Please note, however, that the probability of turning out to vote is only applied to those Black individuals who died *in excess*. This group of Blacks—who should not have died in first instance should they had the mortality rates of Whites—should show, in principle, very similar turnout rates to those of Blacks who did not die. Accordingly, bias due to non-comparability effects should be trivial.

Number of Hypothetical Survivors in 2004

Thus far I have not offered a model to estimate how Black excess deaths may continue to affect the composition of the electorate through time. Two points are worth repeating (a) Individuals who die under the age of 18 never had the opportunity to vote, and (b) Individuals who die after the age of 18, but before the age of life expectancy of their cohort, have a shorter electoral life than those who culminate their normal life cycle. As stated above, this suggests that mortality differences have cumulative consequences through time, and thus far I have ignored that many of the dying population could have survived until the presidential election of 2004.

This calculation is based on well established methods developed in demography, usually for survival projections based on manipulation of life tables' formal calculations. The dynamics



of interest pertaining to the present calculation rest on projecting the dying cohorts onto hypothetical surviving populations in the future. Population projections are particularly useful for the assessment of "what if?" type of counterfactuals, and they have been extensively implemented to calculate hypothetical scenarios in which mortality rates constitute the very basis of the counterfactual condition (Preston, Heuveline, and Guillot 2001; White and Preston 1996). Implicit in this exercise is the notion that, even if Blacks have experienced the mortality rates of Whites, there was always the possibility that some of them would have died from one year to the next due to the ordinary dynamics of death in the human lifespan. If a Black death is "avoided," then this surviving individual should equally be exposed to the mortality risks to which Whites were exposed to, as predetermined by the counterfactual condition. Consequently, there is a peripheral dying population that should be discounted.

Table 1: Survival Projection Table
(Black males survivors in 2004 who were 45 years old and died in excess, in 1970, in Michigan)

Survival Year	Survival Age	Mort. Rate (Whites)	Percent Surviving	Number Surviving
t	***x***	$q^{w}_{x,t}$	***% $l_{x,t}$***	***$l_{x,t}$***
1970	45	0.00290	100.0	**123**
1971	46	0.00310	99.7	123
1972	47	0.00330	99.4	122
1973	48	0.00352	99.1	122
1974	49	0.00375	98.4	121
1975	50	0.00401	98.0	121
1976	51	0.00426	97.6	120
⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮
1998	73	0.01743	82.6	102
1999	74	0.01828	80.2	99
2000	75	0.01931	78.8	97
2001	76	0.02002	77.3	95
2002	77	0.02201	75.8	93
2003	78	0.02306	74.2	91
2004	79	0.02508	72.7	***89***

For illustrative purposes, Figure 3 portrays the survival curve of 123 Black males of age 45 who died in excess, in 1970, in the state of Michigan. As time progresses, this hypothetical group of survivors are exposed to the mortality risks of Whites with their same characteristics. Accordingly, 89 of them (73% of the initial 123) would have survived until 2004, when they would have been 79 years old. Figure 3 also depicts how, as time progresses, the survival curve meets the bars (which represent the presidential elections between 1970 and 2004), signaling the percentage of Black individuals that would have survived at each instance and would have had the opportunity to participate in each of the presidential elections.

Table 1 shows the mechanism behind the calculation of the total number of hypothetical survivors. Note that the "radix" (i.e., the starting number of individuals in the table) equals the

total excess deaths of Black males of age 45, in 1970, in Michigan, estimated using Equation 1. In mathematical terms, for this specific example:

$$Radix = l_{x,t,(e=0)} = TED^{B}_{i=45,s=1,z=18,y=1970}$$

Where x is the survival age, t is the survival year, and e is the number of years of risk for death exposure.[36] Table 1 also illustrates how, as a fraction of the initial 123 Black individuals survive into a future year t, their age x advances accordingly. For example, 121 individuals (98.4% out of the initial 123) would have survived until 1974, when they would have been 49 years old.

The estimation of the projected surviving population 1 year into the future can be expressed in this form:

$$l_{(x+1)(t+1)} = l_{x,t} - \left(l_{x,t} \bullet q^{W}_{x,t}\right)$$

Where the term $q^{w}_{x,t}$ denotes the respective mortality rate of Whites at age x in year t. Accordingly, the term $(l_{x,t} \bullet q^{w}_{x,t})$ represents the dying population at age x in year t, which is discounted from the preceding population $l_{x,t}$.

Taking these into account, the total number (out of the initial 123) of Black hypothetical survivors that would have reached the year 2004 can be expressed in the following terms:

$$l_{(x+e=79)(t+e=2004)} = r_{i=45,s=1,z=18,t=1970}$$

Where r denotes the total number of Black male survivors in 2004 from those who were 45 years old and died in excess, in 1970, in Michigan. Accordingly, the total number of hypothetical survivors in 2004 can be formalized as follows:

[36] Note that, in this particular instance, x is equal to the age at death (i), and t is equal to the year of death (y). The years of risk for death exposure can be represented by this equation: $e = (2004 - t)$.

$$THS^{B}_{2004} = \sum_{Y=1970}^{2004} \sum_{Z=1}^{35} \sum_{S=1}^{2} \sum_{I=0}^{84} \left(r_{i,s,z,y} \right) \qquad (3)$$

and,

$$r_{i,s,z,y} = \sum_{x} \sum_{t} \sum_{e} l_{(i,s,z,y),e,t,x} \qquad (3A)$$

where *i*, *s*, *z*, and *y* in Equation 3A are fixed. In essence, what Equation 3 represents is the summation of all the *yellow* cells in Table 1 throughout each and all levels of *i*, *s*, *z*, and *y* (i.e., the sum of all hypothetical survivors in 2004 from those Blacks who died in excess between 1970 and 2004, in all 35 states, males and females, from age 0 (under 1) to age 84).

It is worth noting that Equation 3 does not account for internal migration. The model assumes that the net effects of internal migration are close to 0. It assumes in this regard that, for example, as individuals migrate from Ohio to Illinois a similar number would migrate from Illinois to Ohio. It is also important to observe that the most mobile population during the period covered in this study are immigrants, especially from Latin America, who are discounted from the Black and White populations used in the calculations. Finally, it is expected that this assumption holds because the period covered in this study begins when the Second Great Migration of Blacks already ended by 1970.

Another particularity of the model is that it does not account for fertility. An important portion of Blacks who die in excess are their reproductive years, and, therefore, it would have been ideal to control for their population growth. It is expected that in a wide range of time, Black population growth due to reproduction among survivors would slightly surpass the

population discounted due to risks for death, in accordance with standard population dynamics. Thus, by not accounting for fertility among the female population who died in excess, Equation 3 is underestimating the total number of survivors in 2004, which, again, makes the estimation a conservative one.[37] In essence, a model that accounts for fertility should show that the total number of hypothetical survivors until the year 2004 would have been higher to the total number of excess deaths.

Equation 3 has another underlying assumption worth mentioning. There may be an overestimation of the excess deceased Blacks' probability of survival by using the average mortality rate of Whites as a discount factor (third column, Table 1). The model assumes that, should excess deceased Blacks have not died, their latent (unmeasured) health statuses correspond to latent health statuses of average White survivors of similar characteristics. This is a misspecification of the model. After matching on age, sex, state, and year, the average White survivor may still have a better health status than a similar Black that should not have died. In an ideal world, the mortality rate that would have been used as a discount factor is that of Whites with the same health statuses of Blacks who should not have died.

Total Black Votes Lost in the Presidential Election of 2004 to Mortality Gaps

The multiplication of Equation 2 times Equation 3 (which was derived from calculations obtained from Equation 1) produces the number of Black votes lost in the presidential election of 2004 due to mortality gaps between the races for a specific configuration of *I*, *Z*, *S*, and *Y*. Therefore, by summing this product across all levels of *I*, *Z*, *S*, and *Y*, we obtain the total number of Black votes lost in 2004:

[37] The size of the impact resulting from population growth due to fertility among survivors on the presidential election of 2004 might be intermediate to modest. In the 35-year period covered in this study most of the children of survivors would have been young by 2004 (i.e., they would have been low to mid-level participants). Further, a portion of them would not have yet reached voting age by 2004.

$$TVL^{B}_{2004} = \sum_{Y=1970}^{2004} \sum_{Z=1}^{35} \sum_{S=1}^{2} \sum_{I=0}^{84} \left(r_{y,z,s,i} \bullet \Pr(V=1)^{B}_{z,s,g(a)} \right) \quad (4)$$

Where *g* represents the age group (1 = 18-29, 2 = 30-44, 3 = 45-64, and 4 = 65-84 years). Note that in Equation 4 $g(a) = f(i,y),$ and that $g(a)$ is the respective age group that matches age *a*. Accordingly, *a* can be estimated as $a = (2004 - y) + i, \ \forall a \geq 18$. However, this is a mere notation artifact. By implementing the probability of turning out to vote among Blacks $\Pr(V=1)^{B}$ we are inherently excluding hypothetical survivors who did not reach the voting age of 18 by year 2004. Concisely, Equation 4 assigns to each and all hypothetical survivors their specific probability of turning out to vote in 2004 and sums across all levels of *I*, *Z*, *S*, and *Y*.

Results and Discussion

Implementing Equation 1, the total number of excess deaths among Blacks (both males and females) between the ages 0 and 84 years, in the 35 states included in the study, between 1970 and 2004, is 2,701,377. Considering that the total Black population in 1970 was 22.6 million and 36.1 million in 2004, this number represents 20% of the total national Black population growth in this period. Should Blacks have had the mortality rates of Whites between 1970 and 2004, the total number of Black deaths would have been reduced from 8.5 million to 5.8 million. In plain words, Blacks would have suffered about 32% fewer deaths if their mortality rates had been the same as Whites; 1 of every 3 Black deaths would have not happened.

Results from Equation 3 show that, out of the 2.7 million Black excess deaths between 1970 and 2004, the total number of hypothetical survivors in 2004 would have been 2,342,478. This means that, should Blacks have had the mortality rates of Whites between 1970 and 2004,

its population in 2004 would have not been 36.1 million but 38.5 million. Over 6% of the total Black population that would have been expected in 2004 at Whites' rates of survival did not survive to 2004. To exemplify the substantive significance of these numbers, consider that, should the American population have had the mortality rates of Blacks between 1970 and 2004, the population of the United States in 2004 would have been about 16 million fewer. Much attention has been given in the social sciences to the possible high number of Black (male) votes that are loss due to racial disparities in incarceration rates. According to the Sentencing Report, 1.4 million Black males were ineligible to vote in the presidential election of 2004 due to felony disfranchisement (McDonald 2004). The total number of hypothetical survivors in 2004 is 67% higher than the number of disfranchised Black male felons.

According to Equation 4, the total number of votes lost by the Black community in the presidential election of 2004 is 1,196,365. This number is about 1% of the national turnout in the 2004 presidential election. For the Black community, specifically, these 1.2 million votes represent about 8.5% of the national Black vote in the 2004 presidential election. This number exceeds the total increase in Black votes between the four presidential elections from 1984 through 1996 (from 10.3 million to 11.3 million, respectively) or the total increase in Black male votes between the five presidential elections from 1984 through 2000 (from 4.2 million to 5.3 million, respectively).

Apart from the substantive significance that these numbers represent for the Black community and for American democracy in general, the excessive demographic contraction of Blacks spills over into other areas of the political arena. Another set of political implications that derive from these numbers is linked to the fact that Blacks are overwhelmingly Democratic Party voters (at about 90%). Blacks practically do not split their tickets. Mortality gaps between the

races, therefore, may disproportionally attrit the Democratic voting base. These effects may also be notable in Democratic primaries, especially as Blacks manifest a higher vote share in key Southern states (some over 40%). If this is so, health disparities between Blacks and Whites may have had meaningful political consequences in local and possibly general elections. Lost Black votes may have been pivotal in local elections including political control for cities, counties, and congressional districts.

It is important to highlight that results are based on only 35 years of mortality exposure. All cumulative effects of excess deaths that occurred before 1970 are excluded from the analysis. Thirty-five years represent just about half of the life cycle of Blacks during the period between 1970 and 2004. Ideally, an analysis of this nature would include enough years to exhaust a complete cohort of the population under observation. The effects of mortality disparities between Blacks and Whites presented herein are therefore conservative. Apart from including a short period compared to a complete, average life cycle of Blacks, results do not account for fertility; for example, females who died prematurely cannot have children thereafter, and the calculations did not account for this fact. The estimation of the probability of turning out to vote only partially accounts for felons, thus, as noted, my denominator (the voting-age population) is significantly larger than it should be.

Conclusions

There are no current estimations of the impact that excessive mortality of Blacks in comparison to Whites exerts over political participation in the United States. This study attempts to fill that void by setting forth an assessment of said impact over the racial composition of the electorate in the presidential election of 2004. The quality and the size of the sets of data implemented in this study permitted me to physically *count* the total number of excess deaths

while accounting for various yet significant demographic aspects of racial mortality schedules that affect electoral politics.

Results show that racial mortality gaps are responsible for about 2.7 million Black excess deaths between 1970 and 2004. Out of these 2.7 million Black individuals, about 2.3 would have survived to 2004, should Blacks have had the same mortality rates of Whites during the period under observation. These 2.3 million hypothetical survivors represent a substantive vanishing electorate among the Black community, calculated at about 1.2 million Black votes that were lost in the presidential election of 2004. The premature mortality of Blacks throughout time generates an important number of years of political life lost among Blacks that translate into a racial differential in the composition of the electorate, at least in presidential elections.

Because the dead cannot vote or voice opinions, excess deaths due to disparities in health may have contributed to the imbalance of political power and representation especially between Blacks and Whites, where these disparities are most pronounced. It is possible that these effects may be related to the political processes surrounding racial prejudice and discrimination, especially as the bulk of the evidence shows that disparities in health between Blacks and Whites are persistent at all levels of the socioeconomic spectrum (Braveman et al. 2006; Geronimus et al. 1999; Geronimus, Bound, and Waidmann 2001, Geronimus et al. 2006; Schoendorf et al. 1992; Williams 1996). Considering that racial prejudice and discrimination are powerful political forces that impact most social determinants of health, it is possible to hypothesize that health disparities between Blacks and Whites may have explicit politically related causes and consequences.

References

Adler, Nancy E., Thomas Boyce, Margaret A. Cheney, Susan Folkman, and Leonard Syme. 1993. "Socioeconomic Inequalities in health: No easy solution." *The Journal of the American Medical Association*, 269(24): 3140-3145

Adler, Nancy E., Thomas Boyce, Margaret A. Cheney, Sheldon Cohen, Susan Folkman, Robert L. Kahn, and S. Leonard Syme. 1994. "Socioeconomic Status and Health: The Challenge of the Gradient." *American Psychologist* 49(1): 15-24.

Adler, Nancy E., and Joan M. Ostrove. 1999. "Socioeconomic Status and Health: What We Know and What We Don't." *Annals of the New York Academy of Sciences*, 869:3-15

Antonovsky, Aaron. 1967. "Social Class, Life Expectancy and Overall Mortality." *The Milbank Memorial Fund Quarterly* 45(2): 31-73.

Bobo, Lawrence D., and Michael C. Dawson. 2009. "A Change has Come: Race, Politics, and the Path to the Obama Presidency. *Du Bois Review*, 6(1): 1-14.

Braveman, Paula A., Catherine Cubbin, Susan Egerter, David R. Williams, and Elsie Pamuk. 2006. "Socioeconomic Disparities in Health in the United States: What the Patterns Tell Us." *American Journal of Public Health* 100: 186-96.

Burden, Barry C. 2000. "Voter turnout and the National Election Studies." *Political Analysis*, Vol. 8, No. 4: 389-398.

Clougherty, Jane E., Kerry Souza, and Mark Cullen. 2010. "Work and Its Role in Shaping the Social Gradient in Health." *Annals of the New York Academy of Sciences* 1186(1): 102-24.

Crimmins, Eileen M., Melanie Johnston, Mark Hayward, and Teresa Seeman. 2003. "Age Differences in Allostatic Load: An Index of Physiological Dysregulation." *Experimental Gerontology*, Vol. 38, No. 7: 731-734.

Cruickshank, Kennedy Jean Claude Mbanya, Rainford Wilks, Beverly Balkau, N. McFarlane-Anderson, and Terrence Forrester. 2001. "Sick Genes, Sick Individuals or Sick Populations with Chronic Disease? The Emergence of Diabetes and High Blood Pressure in African-origin Populations. *International Journal of Epidemiology*, Vol. 30, No. 1: 111-117.

Dominguez, Tyan Parker, Christine Dunkel-Schetter, Laura M. Glynn, Calvin Hobel, and Curt A. Sandman. 2008. "Racial Differences in Birth Outcomes: The Role of General, Pregnancy, and Racism Stress." *Health Psychology* 27(2): 194-203.

Dressler, William W., Kathryn S. Oths, and Clarence C. Gravlee. 2005. "Race and Ethnicity in Public Health Research: Models to Explain Health Disparities." *Annual Review of Anthropology*, 34: 231-252.

Enos, Ryan D. 2011. "Racial Threat: Field-experimental Evidence That a Proximate Racial Outgroup Activates Political Participation." Working paper. Harvard University.

Linda P. Fried, Catherine M. Tangen, Jeremy Walston, Anne B. Newman, Calvin Hirsch, John Gottdiener, Teresa Seeman, Russell Tracy, Willem J. Kop, Gregory Burke, and Mary Ann McBurnie. 2001. "Frailty in Older Adults: Evidence for a Phenotype." *Journal of Gerontology*, Vol. 56A, No.3: 146-156.

Geronimus, Arline T. 2003. "Damned if you do: Culture, identity, privilege, and teenage childbearing in the United States." *Social Science & Medicine* 57, 881-893.

Geronimus Arline T., John Bound, Timothy A. Waidmann, Marianne M. Hillemeier, and Patricia B. Burns. 1999. "Excess Mortality among Blacks and Whites in the United States." *New England Journal of Medicine* 335: 1552-58.

Geronimus, Arline T., John Bound, and Timothy A. Waidmann. 2001. "Inequality in Life Expectancy, Functional Status, and Active Life Expectancy across Selected Black and White Populations in the United States." *Demography* 38, 227-51.

Geronimus, Arline T., Margaret Hicken, Danya Keene, and John Bound. 2006. "Weathering" and Age Patterns of Allostatic Load Scores among Blacks and Whites in the United States." *American Journal of Public Health* 96(5): 826-33.

Grossack, Martin M. 1963. *Mental Health and Segregation*. New York, New York: John Wiley.

Henry, P. J. and Christine Reyna. 2007. "Value Judgments: The Impact of Perceived Value Violations on American Political Attitudes." *Political Psychology*, Vol. 28, No. 3: 273-298.

Holbrook, Allyson L. and Jon A. Krosnick. 2010. "Social desirability bias in voter turnout reports: Tests using the item count technique." *Public Opinion Quarterly*, Vol. 74, No. 1: 37-67.

Kristin A. Lane, and John T. Jost. 2011. "Black Man in the White House: Ideology and Implicit Racial Bias in the Age of Obama." In Gregory Parks, Matthew Windust Hughey, Eds. *The Obamas and a (post) racial America?* New York, New York: Oxford University Press

Kaufman, Jay S., Andrew E. Long, Youlian Liao, Richard S. Cooper, and Daniel L. McGee. 1998. "The Relation between Income and Mortality in U.S. African Americans and Whites." *Epidemiology* 9(2): 147-55.

Kinder, Donald R., and David O. Sears. 1981. "Prejudice and Politics: Symbolic Racism Versus Racial Threats to the Good Life." *Journal of Personality and Social Psychology*, Vol. 40, No. 3: 414-431.

Kitagawa, E. M., and Philip Hauser M. 1973. *Differential Mortality in the United States: A Study of Socioeconomic Epidemiology.* Cambridge, MA: Harvard University Press.

Kittles, Rick A., and Kenneth M. Weiss. 2003. "Race, Ancestry, and Genes: Implications for Defining Disease Risk." *Annual Review of Genomics and Human Genetics*, Vol. 4: 33-67.

Laveist, Thomas Alexis. 2002. *Race, ethnicity, and health: A public health reader.* San Francisco: Jossey-Bass.

Levine, Robert S., James E. Foster, Robert E. Fullilove, Mindy T. Fullilove, Nathaniel C. Briggs, Pamela C. Hull, Baqar A. Husaini, and Charles H. Hennekens. 2001. "Black-White Inequalities in Mortality and Life Expectancy, 1933-1999: Implications for Healthy People 2010." *Public Health Reports* 116: 474-83.

McCord, Colin and Harold P. Freeman. 1990. "Excess Mortality in Harlem." *The New England Journal of Medicine* 322: 173-177.

McDonald, Michael P. 2004. "Every Eligible Voter Counts: Correctly Measuring American Turnout Rates." *Brookings Institution Report*. Available online at http://www.brookings.edu/views/papers/20040909mcdonald.pdf

McEwen, Bruce S. and Teresa Seeman. 2006. "Protective and damaging effects of mediators of stress: Elaborating and testing the concepts of allostasis and Allostatic load." *Annals of the New York Academy of Sciences* 896: 30-47.

Marmot, M. 2005. "Social Determinants of Health Inequalities." *Lancet* 365(9464): 1099-1104.

Massey, Douglas S.. 2004. "Segregation and Stratification: A Biosocial Perspective." *Du Bois Review*, Vol. 1: 7-25.

Mendelberg, Tali. 2008. "Racial Priming Revived." *Perspectives on Politics*, Vol. 6, No. 1: 109-123.

Pappas, Gregory, Susan Queen, Wilbur Hadden, and Gail Fisher. 1993. "The Increasing Disparity in Mortality between Socioeconomic Groups in the United States, 1960 and 1986." *New England Journal of Medicine* 329: 103-09.

Pratto, Felicia, James Sidanius, Lisa M. Stallworth, and Bertram F. Malle. 1994. "Social dominance orientation: A personality variable predicting social and political attitudes." *Journal of Personality and Social Psychology*, Vol. 67, No. 4: 741-763.

Preston, Samuel H., Patrick Heuveline, and Michael Guillot. 2001. *Demography: Measuring and Modeling Population Processes.* Malden, MA: Blackwell Publishers.

Rabinowitz, Joshua L., David O. Sears, James Sidanius, and Jon A. Krosnick. 2009. "Why do White Americans Oppose Race-Targeted Policies? Clarifying the Impact of Symbolic Racism." *Political Psychology*, Vol. 30, No. 5: 805-828.

Rothman, Kenneth J., and Sandre Greenland. 1998. *Modern Epidemiology*. 2nd ed. Philadelphia: Lippincott Williams & Wilikins.

Sampson, Robert J., and William Julius Wilson. 1995. "Toward a Theory of Race, Crime, and Urban Inequality." In John Hagan and Ruth D. Peterson, Eds. *Crime and Inequality*. Stanford, CA: Stanford University Press.

Sankar, Pamela, Mildred K. Cho, Celeste M. Condit, Linda M. Hunt, Barbara Koenig, Patricia Marshall, Sandra Soo-Jin Lee, and Paul Spicer. 2004. "Genetic Research and Health Disparities." *The Journal of the American Medical Association* 291 (24): 2985-2989.

Satcher, David, George E. Fryer Jr., Jessica McCann, Adewale Troutman, Steven H. Woolf, and George Rust. 2005. "What If We Were Equal? A Comparison of the Black-White Mortality Gap in 1960 and 2000." *Health Affairs* 24(2): 459-64.

Schoendorf, Kenneth C., Carol J. R. Hogue, Joel C. Kleinman, and Diane Rowley. 1992. "Mortality among Infants of Black as Compared with White College-educated Parents." *New England Journal of Medicine* 326: 1522-26.

Sears, David O. and Donald R. Kinder. 1971. "Racial Tensions and Voting in Los Angeles." In Werner Z. Hirsch, Ed. *Los Angeles: Viability and Prospects for Metropolitan Leadership*. New York, New York: Praeger.

Seeman, Teresa E., Bruce S. McEwen, John W. Rowe, and Burton H. Singer. 2001. "Allostatic load as a marker of cumulative biological risk: MacArthur studies of successful aging." *Proceedings of the National Academy of Sciences*, Vol. 98, No.8: 4770-4775.

Seeman, Teresa, and Eileen Crimmins. 2001. "Social Environment Effects on Health and Aging: Integrating Epidemiologic and Demographic Approaches and Perspectives." *Annals of the New York Academy of Sciences* 954: 88-117.

Seeman, Teresa, Elissa Epel, Tara Gruenewald, Arun Karlamangla, and Bruce S. McEwen. 2010. "Socio-economic Differentials in Peripheral Biology: Cumulative Allostatic Load." *Annuals of the New York Academy of Sciences* 1186, 223-39.

Sidanius, James, and Felicia Pratto. 2001. *Social Dominance: An Intergroup Theory of Social Hierarchy and Oppression.* New York, New York: Cambridge University Press.

Taylor, Shelley E., Rena L. Repetti, & Teresa Seeman. 1997. "Health Psychology: What Is an Unhealthy Environment and How Does It Get Under the Skin?" *Annual Review of Psychology* 48, 411-47.

Tesler, Michael, and David O. Sears. 2010. *Obama's Race: The 2008 Election and the Dream of a Post-Racial America.* Chicago, IL: University of Chicago Press.

U.S. Dept. of Health and Human Services, National Center for Health Statistics. MULTIPLE CAUSE OF DEATH PUBLIC USE FILES, [Computer file]. Hyattsville, MD: U.S. Dept. of Health and Human Services, National Center for Health Statistics [producer], 2004. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2007-12-20. doi:10.3886/ICPSR20623.v1

Valentino, Nicolas A., and David O. Sears. 2005. "Old Times There Are Not Forgotten: Race and Partisan Realignment in the Contemporary South." *American Journal of Political Science*, Vol. 49, No. 3: 672-688.

Valentino, Nicolas A., Vincent L. Hutchings, and Ismail K. White. 2002. "Cues that Matter: How Political Ads Prime Racial Attitudes During Campaigns." *American Political Science Review*, Vol. 96, No. 1: 75-90.

Vavreck, Lynn. 2007. "The exaggerated effects of advertising on turnout: The dangers of self-reports." *Quarterly Journal of Political Science* 2: 325-343.

Wight, Richard G., Janet R. Cummings, Dana Miller-Martinez, Arun S. Karlamangla, Teresa E. Seeman, and Carol S. Aneshensel. 2008. "A Multilevel Analysis of Urban Neighborhood Socioeconomic Disadvantage and Health in Late Life." *Social Science and Medicine*, Vol. 66, No. 4: 862-872.

White Kevin M., and Samuel H. Preston. 1996. "How Many Americans are Alive because of Twentieth-Century Improvements in Mortality?" *Population and Development Review* 22(3): 415-29.

Williams, David R. 1990. "Socioeconomic Differences in Health: A Review and Redirection." *Social Psychology Quarterly*, Vol. 53, No. 2: 81-99.

———. 1996. "Race/Ethnicity and Socioeconomic Status: Measurement and Methodological Issues." *International Journal of Health Services*, Vol. 26, No. 3: 483-505.

Williams, David R., and Chiquita Collins. 1995. "US Socioeconomic and Racial Differences in Health: Patterns and Explanations." *Annual Review of Sociology*, Vol. 21: 349-86.

Williams, David R., Harold W. Neighbors, and James S. Jackson. 2003. "Racial/ethnic discrimination and health: Findings from community studies." *American Journal of Public Health*, Vol. 93, No. 2: 200-208.

Williams, David R., and Pamela Braboy Jackson. 2005. "Social Sources of Racial Disparities in Health." *Health Affairs*, Vol. 24, No. 2: 325-334.

Williams, David R., and Selina A. Mohammed. 2009. "Discrimination and Racial Disparities in Health: Evidence and Needed Research." *Journal of Behavioral Medicine*, Vol. 32, No. 1: 20-47.

Winkleby, Mailyn A., Darius E. Jatulis, Erica Frank, and Stephen P. Fortmann. 1992. "Socioeconomic Status and Health: How Education, Income, and Occupation Contribute to Risk Factors for Cardiovascular Disease." *American Journal of Public Health* 82(6): 816-20.

Wizemann, Theresa M., and Mary-Lou Pardue. Eds. 2001. *Exploring the Biological Contributions to Human Health: Does Sex Matter?* Washington DC: National Academy Press.

INFANT MORTALITY AND THE PRESIDENT'S PARTY

Javier M. Rodriguez
Department of Political Science
University of California, Los Angeles

Abstract

It is hypothesized that variation in infant mortality rates in the United States is linked to the party that controls the presidency. Data to test this hypothesis are from various sources including the Department of Health and Human Services and the Census Bureau, and extend from 1965 to 2010. Statistical analyses show that national infant mortality rates for whites and blacks decrease under Democratic administrations and increase under Republican administrations. Variations in low birthweight and preterm birth, which are known to precede infant mortality, are also related to the president's party. The different policy preferences of Democratic and Republican presidents show to be one of the most influential factors impacting persistent infant mortality rate disparities between blacks and whites and the underperformance of the American health system at reducing infant mortality compared to other world nations and territories.

Introduction

The study of infant mortality rates (defined as deaths of infants under 1 year of age per 1,000 live births), racial disparities in these rates, and policy interventions to address them, has traditionally been the effort of epidemiology, public health, and medical researchers. Unfortunately, political science researchers have played a peripheral role in this endeavor despite the fact that political actors, bureaucracies, and institutions ultimately set the policy agenda and its procedures. Multidisciplinary efforts to explain infant mortality, which bridge biological and social (but not political) phenomena, have effectively shed light on how environment-physiology interactions affect infant mortality, sometimes more disproportionally affecting socioeconomically disadvantaged communities. This article attempts to connect these two historically disconnected domains of research. It adds to the extant record on infant mortality and racial disparities in infant mortality rates by indicating that the omission of political variables from this research agenda made researchers fail to notice one of the most influential factors impacting infant mortality and racial disparities in the United States over the last half-century: The divergent policy formulations of Democratic and Republican presidents.

In spite of some inherent limitations, infant mortality rate has traditionally been the focus of health policy as it is regarded a sensitive measure of population health (*1*). This is because vulnerable communities, such as pregnant women and the newborn, are more likely than others to be affected by variations in the environment. Much research demonstrates that biological indicators of health are particularly sensitive to contextual factors. Acknowledging that research in genetics is crucial for the improvement of overall population health and disparities in health status, social and environmental determinants of health are nevertheless broadly signaled as the major contributors to health status and racial disparities in health (*2-4*). Socioeconomic adversity

and inequality, racial discrimination, lack of access to high-quality health services and resources, and a wide range of correlates (e.g., residential segregation, unhealthy air quality, malnutrition) interact with the biological endowment of the individual, thus affecting health. The continual process of adaptation the human body undergoes in response to daily life challenges has considerable effects on physiological indicators of health risks (*5*) including reproductive outcomes, especially among minority communities (*6*). Research in political science has demonstrated that American political elites have become increasingly polarized along partisan lines (*7*). The frequency of cross-party coalitions in most key policy issues has diminished since the 1970s, and more radically during the last decade. In recent times, health policy has arguably been the most polarizing issue between the parties. Given that policy generally affects the distribution of public goods and services across the entire population, it is a major mechanism impacting the conditions to which people must respond. Divergent policy prescriptions by the party in power thus mold the ebbs and flows of health indicators like infant mortality rates.

By executing his many legislative, regulatory, and budgetary powers, the President of the United States is one of the main actors at decreeing the platform of his party while in office. Research in political science has shown that postwar macroeconomic policy in the United States reflects the economic interests of the party that controls the presidency (*8*). Building upon this literature, new studies of income inequality in the American context lean toward political explanations as opposed to technological and globalization paradigms (*9*). Indeed, early work showed concern in that partisan, short-term manipulations of economic policy, especially during electoral periods, could become the source of long-term social costs (*10*). Recent work developed by Larry Bartels shows that contrasting patterns in economic growth linked to presidential partisan regimes have played a central role in the rise of income inequality in the

United States after World War II (*11*). Political ideology informs public policy that affects health directly or indirectly (*12*), and infant mortality is no exception (*13*).

The history of infant mortality in the United States, at least in the last half-century, exemplifies some of the contrasts that exist between policy goals and health outcomes. For example, the United Nations Population Division recently ranked the United States 43rd among world nations and territories for infant mortality between the years 2005 and 2010; for infant mortality between 1960 and 1965 the United States ranked 13th (*14*). In the United States, in 2010, the infant mortality rate for blacks (BIMR) was similar to the infant mortality rate that whites (WIMR) had over 30 years before in the late 1970s. Indexes of racial disparities in health, like the infant mortality race ratio, increased from 1.9 in 1965 to 2.3 in 2010. Key predictors of infant mortality rate (IMR) like low birthweight rates (LBW, defined as percent of infants weighting less than 2,500 grams at birth) and preterm birth rates (PB, defined as percent of infants born at less than 37 weeks of gestation) increased from 7.9 and 9.3 in 1970 to 8.2 and 12.0 in 2010, respectively. In spite of overall improvements in public health in the United States (Fig. S1), advances in IMR in other developed (or even non-developed) nations during the last half-century signals a comparative underperformance of the American health system (*15*). Within the United States, the health status of whites suggests that there are successful aspects of the American health system that are still to be effectively delivered to other sectors of the population. This article argues that said underperformance in, and misdistribution of, health outcomes are related to the distribution of resources that results from the implementation of policies endorsed by the party in power.

Much evidence shows that the political realignment of the 1960s marked a critical, issue-based departing point between the Democratic and Republican parties (*16*). The way the parties

recast their political platforms according to the advancements of the Civil Rights Movement, the 1964 Civil Rights Act, the Voting Rights Act of 1965, and the Social Security Act of 1965 changed the course of American politics. Liberal civil rights initiatives during the late 1960s and early 1970s, which paired socially-oriented issues with racially-based ones, paved the way for new elite views and new constituencies of supporters and opponents that rallied along party lines (*17*). The Democratic Party stood apart from its long-protected Jim Crow system in the South and reframed its policy platform around civil rights while the Republican Party focused its platform on economic conservatism (*18*). These divergent positions, promoted by political actors, still determine the fate of health policy and outcomes today (*19, 20*). For example, following the Supreme Court decision to uphold most of the Affordable Care Act on June 28, 2012, Republican candidate Mitt Romney publicly declared that, if elected, he would repeal "Obamacare" on the first day of his presidency (*21*).

Data and methodological approach

Data from various sources, including the Department of Health and Human Services, and the U.S. Census Bureau (for detailed data sources see SOM), are used to test the hypothesis that variations in infant mortality and racial disparities in IMR are linked to the president's party. The variables included in the statistical analyses are deemed to be among the most relevant to infant mortality and racial disparities in IMR according to current literature (Table 1; also see SOM). The data mostly extend from 1965 to 2010, depending on data availability. Partisan control of the presidency is lagged by one year, since it is not expected that policy enacted by a president would affect infant mortality in his first year. One-year lags of partisan control are based on research showing that effects of health-related policies manifest about one year after policies are introduced (*22*). This includes the effects of programs targeting the reduction of infant mortality

(*23*). This assumption has also been successfully applied in other research studying the effects of presidential fiscal policy implementation (*11, 24*). Since infant mortality is strongly related to long-term factors such as the advancement of medical technology, sanitation and public health supervision, improvements in income and education, and programs like Medicaid among many other factors, long- and short-term effects were disaggregated. All variables, therefore, were de-trended to clean the data from structural inertia (except "president's party" and "recessions"). Residuals were recovered after fitting a median cubic spline (Fig S1; SOM). Variations in IMR, WIMR, and BIMR as well as regression coefficients are therefore net of variations attributable to historical trends.

To avoid the ambiguity and multi-dimensionality of measuring health disparities (*25*) through absolute and/or relative indexes of inequality (*26*), IMR, WIMR, and BIMR were analyzed separately. Natural scales were thus preserved. Data analyses such as comparison of medians and means of de-trended data (Table 1), quantile regression on the median and ordinary linear regression, and influence Delta-beta analysis were used to evaluate the relationship between infant mortality and the president's party. Least Angle Regression (LAR) (*27*) was implemented to explore the importance of the president's party at explaining infant mortality rates when subjugated to a model selection algorithm. Sensitivity analysis on linear Seemingly Unrelated Regressions (SUR) (*28*) was used to analyze the robustness of the collective association of the president's party with IMR, WIMR, and BIMR. Sensitivity analyses on structural equation models were used to explore the mediation of relevant covariates between the president's party and infant mortality.

The relationship between infant mortality rates and the president's party

Table 1 evinces a drastic difference between the political parties across de-trended IMR-

Table 1. Medians, means, and test statistics of de-trended variables by president's party

Variable *De-trended*	Median Democratic	Median Republican	Wilcoxon Rank-sum Test (*p*-value)	Probability Random draw Dem > Rep*	Mean Democratic	Mean Republican	T-test (*p*-value)	N
IMR (all races)	-.191	.131	.00	.12	-.174	.112	.00	46
Black IMR	-.461	.208	.00	.21	-.305	.196	.00	46
White IMR	-.188	.112	.00	.10	-.150	.096	.00	46
Low Birthweight Rate (all races)	-.022	.005	.12	.36	-.024	.015	.05	46
Preterm Birth Rate (all races)	-.140	.061	.00	.20	-.131	.066	.00	42
DW-Nominate Score	-.52	.73	.00	.00	-.56	.70	.00	46
Recessions**	.00	.00	.36	.44	.17	.29	.37	46
Family Income Gini Index	-.0002	.0002	.51	.44	-.0002	.0001	.85	46
Black Family Income Gini Index	-.0047	.0014	.00	.23	-.0045	.0027	.00	45
White Family Income Gini Index	.0001	.0005	.82	.48	.0003	-.0002	.74	46
Income Share Ratio Top 5% / Bottom 20%	-.08	.09	.10	.36	-.08	.05	.09	46
Percent Poverty	-.05	.06	.25	.40	-.16	.11	.20	46
Percent Under 1.25 Poverty Threshold	-.26	.04	.29	.41	-.17	.10	.29	45
Percent Female Poverty	-.16	.08	.28	.40	-.15	.09	.27	45
Percent Female Without Husband Poverty	-.79	.50	.03	.31	-.53	.34	.04	46
Percent Black Poverty	-.82	.27	.02	.30	-.69	.42	.01	45
Percent White Poverty	-.11	.06	.40	.43	-.11	.07	.35	46
Percent Black Females Without Husband Poverty	-1.01	.62	.01	.25	-.87	.53	.01	45
Percent White Females Without Husband Poverty	-.41	.25	.13	.37	-.42	.27	.11	46
Income Share of Top 5%	.09	.00	.59	.55	.06	-.04	.45	46
Black Income Share of Top 5%	-.40	.17	.03	.31	-.33	.20	.03	46
White Income Share of Top 5%	.07	-.12	.29	.59	.06	-.04	.61	46
Income Share of Bottom 20%	.03	-.06	.01	.72	.05	-.03	.01	46
Black Income Share of Bottom 20%	.09	-.06	.00	.81	.10	-.06	.00	45
White Income Share of Bottom 20%	.04	-.03	.01	.72	.04	-.03	.02	46
Mean Income of Bottom 20%	229	-113	.02	.71	247	-150	.01	45
Black Mean Income of Bottom 20%	203	-94	.00	.79	282	-171	.00	45
White Mean Income of Bottom 20%	246	-99	.03	.69	210	-128	.03	45
Women's % of Men's Earnings	-.29	.31	.02	.30	-.43	.28	.01	46
Consumer Price Index-URS	-.50	.36	.24	.40	-.98	.63	.10	46
Change in Consumer Price Index-U Medical Care	-.04	-.14	.82	.48	-.06	.05	.80	39
Unemployment	-.15	.12	.31	.41	-.22	.14	.23	46
Black Unemployment	-.03	.31	.27	.40	-.32	.21	.25	46
White Unemployment	-.20	.08	.36	.42	-.20	.13	.24	46
Percent with High School Degree	.03	-.03	.14	.63	.08	-.05	.10	46
Percent Black with High School Degree	.21	-.10	.07	.66	.22	-.14	.05	46
Percent White with High School Degree	.02	-.05	.15	.63	.08	-.05	.08	46
Percent Black Female with High School Degree	-.04	.02	.62	.54	.09	-.06	.54	46
Percent White Female with High School Degree	.04	-.05	.30	.59	.03	-.02	.41	46
Total Maternal/Child Expenditure per capita	.02	-.02	.19	.62	.06	-.04	.11	46
Federal Maternal/Child Expenditure per capita	.02	-.01	.32	.59	.02	-.02	.21	46
Total Maternal/Child Health Expenditure as %GDP	.00	.01	.77	.53	.03	-.17	.27	46
Federal Maternal/Child Health Expenditure as %GDP	-.03	-.01	.49	.44	.00	.00	.99	46
Percent Uninsured	-.25	.04	.25	.37	-.14	.08	.26	31
Alcohol Consumption	.00	.01	.47	.43	.00	.00	.53	43
Tobacco Consumption	27.7	-9.6	.38	.58	7.98	-5.42	.48	42
Median Age of Mother at 1st birth	.025	-.020	.34	.59	.012	-.009	.38	39
Black Median Age of Mother at 1st birth	.041	-0.022	.13	.64	.032	-.025	.09	39
White Median Age of Mother at 1st birth	.017	.004	.80	.52	.002	-.002	.84	39
Percent Rural Population	.04	.03	.96	.50	-.02	.01	.19	46
Abortion Rate	.39	.24	.78	.48	-.05	.03	.86	44
Abortion Ratio	4.04	-4.59	.39	.58	1.57	-.99	.62	44
Abortion Percent	.27	-.18	.74	.53	-.06	.04	.86	44

* Probability that a random draw from Democratic president values would be larger than a random draw from Republican president values. ** Not de-trended.

relevant variables. The data show a general underperformance of Republican presidents at improving the social determinants of health compared to Democratic presidents. All medians and means of de-trended IMR, WIMR, BIMR, LBW and PB are negative (i.e., decrease) under Democratic presidents and positive under Republican presidents. These differences are statistically significant. Figure 1 confirms this pattern. De-trended IMR, WIMR, and BIMR decrease and increase episodically depending on the party that controls the presidency. IMR,

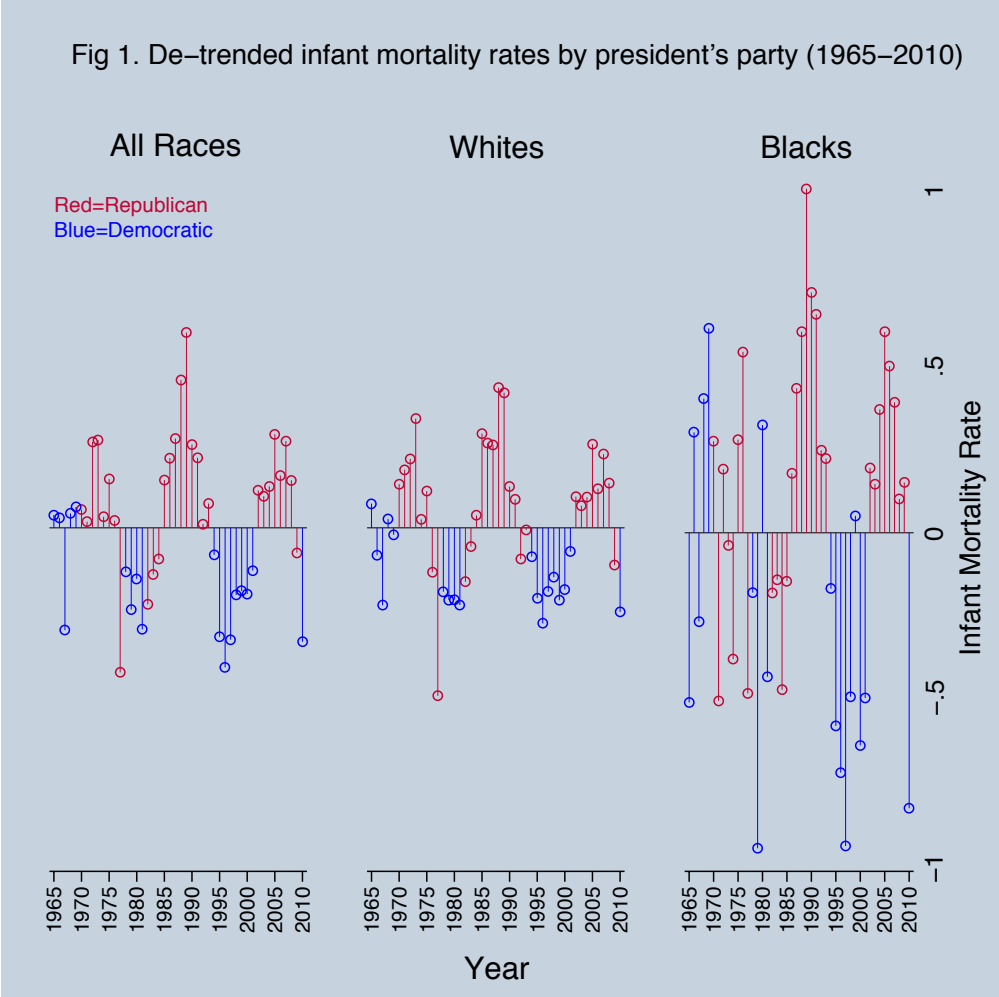


Table 2. Sum of IMR residuals by partisan presidential period (1965-2010)

Presidential Partisan Period	Sum IMR Residuals	Sum WIMR Residuals	Sum BIMR Residuals
Obama*	-.34	-.25	-.82
W. Bush	1.08	.85	2.45
Clinton	-1.85	-1.37	-3.97
Reagan/H. W. Bush	1.70	1.51	3.15
Carter	-.82	-.84	-1.23
Nixon/Ford	.35	.33	-.11
L.B. Johnson	-.13	-.23	.54
Annual Avg. Dems	*-.17*	*-.15*	*-.30*
Annual Avg. Reps	*.11*	*.10*	*.20*

* Calculations account for data between 2009 and 2010

WIMR, and BIMR decrease under Democratic presidents and increase under Republican presidents. The pattern for BIMR (but not for WIMR) shows instability between 1965 and the mid 1970s evincing the period of racially-based issue adaptation of the parties linked to the political realignment. After the mid 1970s, variation in BIMR replicates the cyclical regularity of WIMR and of overall IMR.

Table 2 quantifies these differences. The data show that infant mortality partisan cycles hold for overall IMR, WIMR, and BIMR. There is an exception for BIMR during the period of party adaptation, when the Democratic Party was just beginning to incorporate civil rights into its platform under L.B.J., and during the Nixon/Ford period, who are known to be the most liberal, if not "Democratic-like" Republican presidents (*29*). The data also show that, on average, BIMR declined twice as fast as WIMR under Democratic administrations (-.30 *vs.* -.15), and that BIMR increased twice as fast as WIMR under Republican administrations (.20 *vs.* .10) net of variations attributable to historical trends. If this relationship is real, then absolute inequality indicators like the *gap* in IMR between blacks and whites decreased during Democratic administrations and increased during Republican administrations, on average. Relative inequality indicators like the black-white IMR *ratio* increased under both Democratic and Republican administrations *iff* IMR race ratios were over 2.0, on average. Should the IMR race ratio increased under Democratic presidents, it happened while *diminishing* racial IMRs (i.e., while improving overall health), whereas if the IMR race ratio increased under Republican presidents it happened while *increasing* racial IMRs (i.e., while deteriorating overall health), on average (*30*). Table 2 evinces that, on average, during the last half-century, racial disparities in IMR were not induced by favoring whites while not favoring blacks; rather, racial disparities in IMR were induced during Republican administrations by favoring neither whites nor blacks, with blacks

bearing the worst part. Racial disparities in IMR are not the product of the unequal distribution of privileges but of the unequal distribution of adversity.

Table 2 also shows that average overall IMR partisan differences (-.17 for Democratic presidents *vs.* .11 for Republican presidents) are at the core of the underperformance of the American health system at reducing IMR compared to other world nations and territories. The data suggest that this underperformance is due in great part to 28 years of Republican administration. Parametric estimations of these cross-party differences via quantile regressions at the median and ordinary linear regressions corroborated these differences (Table S1). Delta-beta analysis ensured that these findings were not driven by any specific president-year (Table S2). A cross-validation using president DW-Nominate scores instead of "president's party" delivered practically identical results (Table S1).

Least Angle Regression (LAR) results depicted in Figure 2 show that the president's party is among the most important variables predicting IMR, WIMR, and BIMR. It was not only among the earliest variables chosen by the LAR algorithm to enter the model but also the size of its coefficients is substantial. The sign of these coefficients is always positive and fairly stable in spite of small sample size (*31*) and collinearity between some of the covariates (Tables S3a-S3c). Final models fit the data very well (Tables S4a-S4c; Table S5a). The importance of the president's party as a predictor of IMRs accentuates once it is considered that significant changes in the social determinants of health (e.g., a change equivalent to one standard deviation in education) happen at a slow pace while presidential partisan cycles happen every 4 or 8 years.

A sensitivity analysis was run taking advantage of the nice small-sample properties of Seemingly Unrelated Regressions (*32*). Results are depicted in Table 3. Variables to be included in the models were selected using LAR's output. For the sake of simplicity, Model 6 is

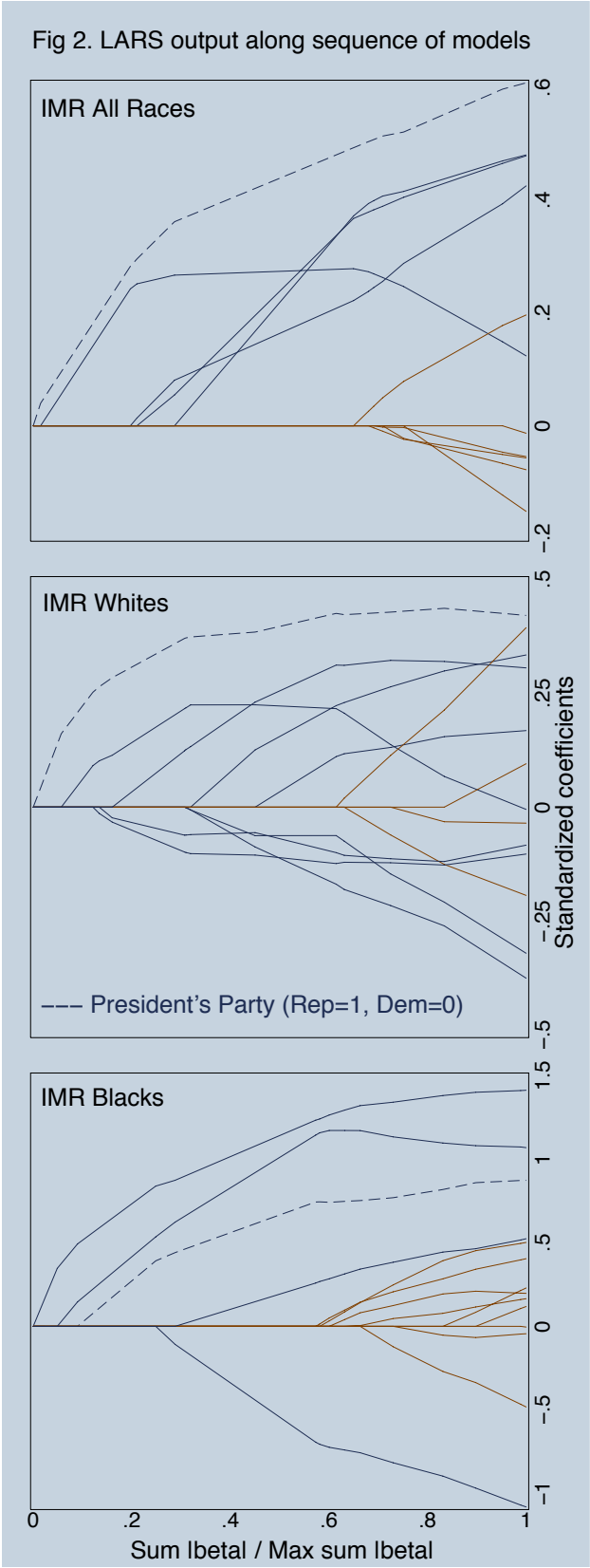


Fig 2. Least Angle Regression (LAR) is a model selection algorithm that provides all possible LASSO (l_1 penalized regression) estimates while efficiently implementing a Forward Stagewise mechanism. It is mainly focused on parsimony and prediction accuracy. In the graphs, the dark lines refer to variables included in the [final] model that minimizes Mallows' Cp statistic, an estimate of prediction error (Fig. S2; Tables S4a-S4c). From left to right of X-axis variables are included in order of importance (Tables S4a-S4c); their effects on the dependent variable are gradually manipulated in the direction of prediction accuracy. Y-axis refers to standardized size of coefficients. Different sets of variables were used in models, and were chosen to diminish multicollinearity (see Tables S3a-S3c). Variables included to predict IMR for all races were (in order of importance): (1) President's party, (2) PB, (3) Percent rural population, (4) LBW, (5) Alcohol consumption, (6) Income share of top 5%, (7) Abortion ratio, (8) Federal per capita expenditure in maternal/child care (selected items, see SOM), (9) Percent with high school, (10) Income ratio women/men, and (11) Income share of bottom 20%. Variables included to predict IMR for whites were: (1) President's party, (2) PB, (3) Percent whites with high school, (4) Abortion ratio, (5) Alcohol consumption, (6) Federal per capita expenditure in maternal/child care (selected items), (7) White's income share of bottom 20%, (8) LBW, (9) White's income share of top 5%, (10) Percent rural population, (11) Income ratio women/men, (12) Cigarette consumption, and (13) Total expenditure maternal/child care as percent of GDP. Variables included to predict IMR for blacks were: (1) LBW, (2) Percent rural population, (3) President's party, (4) Abortion percent, (5) Alcohol consumption, (6) Black's income share of top 5%, (7) Recession years, (8) Total expenditure maternal/child care as percent of GDP, (9) CPI, (10) PB, (11) Income ratio women/men, (12) Black's income share of bottom 20%, (13) Percent black households with females without husband, (14) Percent blacks with high school. Analyses run in STATA 11.

Table 3. Sensitivity Analysis: Seemingly Unrelated Regressions parameter estimates for president's party

MODELS	IMR All Races	IMR Whites	IMR Blacks
Model 1: President's Party (Rep=1, Dem=0)	.29**	.25**	.50**
Model 2: adds Low Birthweight Rate to model 1	.23**	.22**	.34**
Model 3: adds Preterm Births Rate to model 2	.22**	.20**	.38**
Model 4: adds Percent Rural Population to model 3	.21**	.20**	.34**
Model 5: adds Alcohol Consumption per capita to model 4	.18**	.17**	.26*
Model 6: adds Abortions Percent to model 5	.18**	.17**	.29**
Model 7: adds Education to model 6	.18**	.17**	.29**
Model 8: adds Income to model 7	.19**	.17**	.30*

Table 3. Seemingly Unrelated Regressions (SUR) refer to systems of equations linked by their correlated disturbances. Because IMR, WIMR, and BIMR are indicators of the same inherent phenomenon (i.e., infant mortality), and are observed at the same time points (i.e., they are exposed to the same exogenous conditions), then their descriptive regression equations are likely to have correlated disturbances (Tables S5b and S5c). SUR allow to analyze these three individual regression equations simultaneously while integrating the cross-equation correlations into the estimation procedure. SUR also increase the efficiency of parameter estimates compared to OLS and maximum likelihood even in small sample situations. In the table, each row represents a model. Cells are the effect of the president's party on IMR, WIMR, and BIMR, respectively. Number of data points varies across models from 46 to 39 depending on data availability. Guided by LAR's results Model 6 corresponds to the most parsimonious model. In models 7 and 8, adding education and income covariates did not add explanatory power to the model and results were practically identical. Excluding one covariate at a time from Model 6 showed that the president's party contributes the most to the fitting of the models (Table S6). R-squared for models range from .39, .37, and .26 (for IMR, WIMR, and BIMR, respectively) in Model 1 to .73, .65, and .77 in Model 6, respectively. "Abortions percent" in Model 6 is defined as abortions as percentage of pregnancies (excluding fetal deaths and miscarriages). "Education" in Model 7 refers to percent with high school (HS) education for IMR, percent of whites with HS for WIMR, and percent of blacks with HS for BIMR. "Income" in Model 8 refers to income share of bottom 20% for IMR, white's income share of bottom 20% for WIMR, and black's income share of bottom 20% for BIMR. Standard errors are adjusted for small sample size. All variables are de-trended (except "President's Party"). Statistical significance code: ** $p<0.01$, * $p<0.05$. Analyses run in STATA 11.

composed of those variables that showed to be the most significant for all IMR, WIMR, and BIMR (for a SUR analysis of the final models selected through LAR see Table S5a). The data for Model 6 expand from 1969 to 2007 due to data availability. Models 7 and 8 are included in the table because of the centrality of education and income variables in infant mortality literature. Estimates are unstandardized and corroborate LAR's output. Results show that estimates of the differences in IMR, WIMR, and BIMR between the parties are robust, statistically significant, and substantial. Model 6 coefficients state that net of variations attributable to historical trends, covariates held constant, and controlling for cross-equation disturbance collinearity, Democratic presidents would have saved 1 extra infant life (under the

age of 1 per 1,000 live births) for every 5-6 years of Republican administration, on average. Had Republican presidents had the record of Democratic presidents for IMR, most (if not all) IMR differences between the U.S. and the rest of the world would have never existed, all else equal (*33*). A similar scenario arises for racial IMRs. WIMR and BIMR were .17 and .29 units higher, respectively, during a typical Republican-president-year compared to a typical Democratic-president-year. Indicators of health disparities like the racial gap in IMR increased by about 1 infant death (under the age of 1 per 1,000 live births) for every 8 years of Republican administration, on average. These figures have powerful implications for racial disparities in IMR. For example, observed total increase in the black-white IMR gap between 1965 and 2010 (net of factors affecting historical trends) was 7.6 units, of which 3.4 (about 45%) could be attributed to 28 typical years of Republican administration alone.

Of course these estimates are for independent, direct effects of the president's party on IMRs where no mediation dynamics were considered. And mediation analysis is a necessary step presumably because policy implementation by the president's party affects variables that in turn affect IMR; effects of the president's party on IMR are inherently indirect. A baseline structural equation model (Fig. 3a) helped to explore which variables mediate said effects. The model is saturated, yet its simplicity allows us to explore mediation dynamics before small sample constraints. Because the ratio of data points to possible mediators is exiguous, thus limiting the number of variables that could be included in the model and still get reliable estimates, sensitivity analyses allowed mediators to "rotate" alternatively, with each mediator sharing the model with all other possible mediators one time.

Results from the sensitivity analyses (Fig. 3b) show that LBW and PB, two key biological outcomes known to precede infant mortality, mediate to some extent the effect of the

Fig 3a. Baseline structural equation model

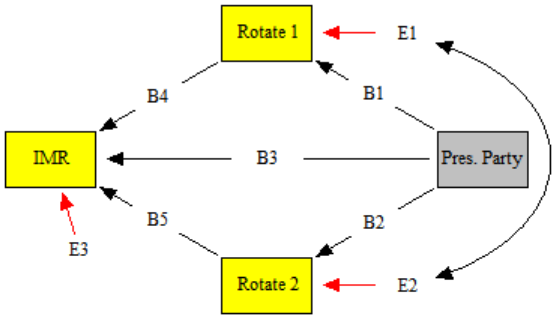


Fig 3a. Sensitivity analyses were run on this structural equation model (SOM). The models were run using maximum likelihood estimation. The dependent variables were IMR (all races), WIMR, and BIMR, one at a time. Rotating variables (29 in total; Tables S3a-S3c) refer to variables that changed from one model to the next until exhausting all possible combinations (406 models for IMR, WIMR, and BIMR, respectively). Each variable, therefore, appears in 28 different models, thus being exposed to all other covariates one time. Coefficients to calculate indirect and direct effects were estimated each run. Indirect effect via Rotate 1 = B1 * B4; indirect effect via Rotate 2 = B2 * B5; direct effect = B3. The model accounts for the correlated error terms of the rotating variables (error parameter not shown). Analyses were run using the "R interface to EQS" on EQS6.2.

Fig 3b. Sensitivity analyses: SEM estimates of president's party effects on IMRs

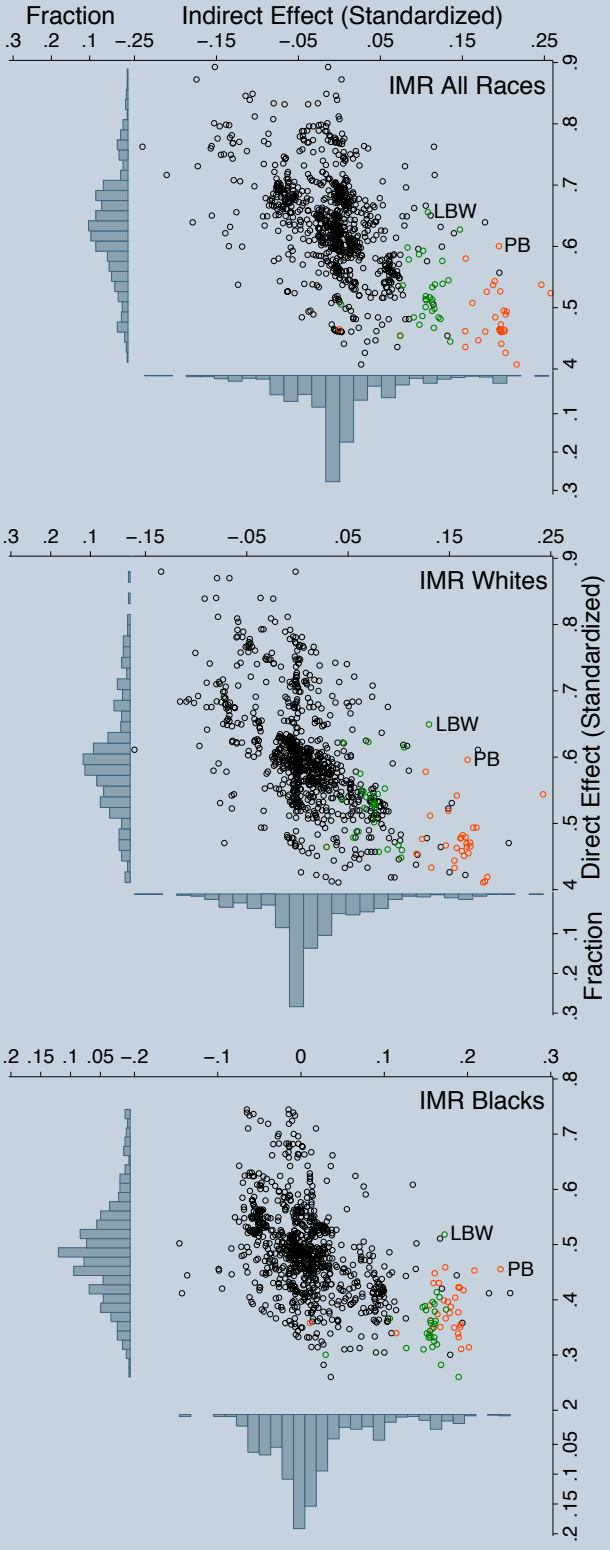


Fig 3b. Results for sensitivity analyses. In the graphs, each data point represents the standardized indirect (X-axis) and standardized direct (Y-axis) effects of the president's party on the dependent variable (IMR, WIMR, or BIMR) when a given [rotating] variable was selected into the model. As indirect effects increase and direct effects decrease, it is possible to identify the variables that are mediating the relationship between the president's party and IMRs. These variables are readily identifiable as they form clusters located at the lower-right quadrant of the scatterplot. Standardized coefficients are comparable within panels but not between panels, as coefficients are standardized using the standard deviation (thus the units of measurement) of the respective dependent variable (and $SD_{IMR} \neq SD_{WIMR} \neq SD_{BIMR}$).

president's party on infant mortality. This finding adds to the literature that shows that contextual, socially related factors affect biological outcomes. The direct effect of the president's party repeatedly diminished—while its indirect effect repeatedly increased—mostly under the presence of LBW and PB, thus producing the clusters localized in the lower-right quadrants of the scatterplots. Histograms evince that standardized indirect effects follow efficient distributions (SD_{IMR}=.07, SD_{WIMR}=.05, and SD_{BIMR}=.06) where most of the effects are around zero. Distributions of direct effects are less efficient, however (SD_{IMR}=.1, SD_{WIMR}=.08, and SD_{BIMR}=.08), with average standardized direct effects on IMR, WIMR, and BIMR of .62, .59, and .48 respectively. Findings evince that, [whichever] the social factors altered by presidential policy, persistently affect IMR, WIMR, and BIMR more so through PB than through LBW (average indirect effects were .16 *vs*. .08, .18 *vs.* .11, and .17 *vs.* .15 respectively). Using average direct effects of the president's party on IMR, WIMR, and BIMR as a baseline, these effects diminished by 15, 12, and 26% respectively when LBW was selected into the model, and by 22, 19, and 20% respectively when PB was part of the model. The direct effects of the president's party on IMRs were not completely attenuated by mediation dynamics. These partial mediations were probably due to a model misspecification where mediators of, and paths between, the president's party and LBW/PB were omitted due to small sample constraints. Estimation results, nevertheless, are consistent with current research on infant mortality and racial disparities in IMR.

These findings also suggest that much of the relationship between the president's party and IMRs remains unexplained. Cross-party variations in IMRs are mediated through more complex, idiosyncratic mechanisms not explored herein. This article, however, sheds light on the existence of a pivotal yet previously unexplored relationship: Infant mortality and the president's

party. Future research should determine further structural mechanisms through which the party that controls the presidency affects infant mortality.

Conclusion

The prevention of infant mortality is of the essence for the advancement of public health; the death of an infant delivers the maximum possible number of human years of life lost. Political authorities do not control health per se but they do shape the prescription of health policy, including interventions that affect infant mortality. This article shows that, net of infant mortality variations attributable to historical trends, infant mortality and racial disparities in IMR are to a considerable degree influenced by *political* phenomena. A good portion of the comparative international underperformance of the U.S. for IMR as well as of persistent black-white disparities in IMR is related to variations imparted under Republican presidents. This relationship between infant mortality and the president's party was tested against the most relevant variables in current literature, using a variety of functional forms, model specifications, and methodological techniques, and it showed to be robust, substantial, and very unlikely to be the product of chance.

References and Notes

1. D. Reidpath, P. Allotey, Infant mortality rate as an indicator of population health. *Journal of Epidemiology and Community Health* **57**, 344 (2003).

2. P. Sankar *et al.*, Genetic research and health disparities. *JAMA: the journal of the American Medical Association* **291**, 2985 (2004).

3. R. Cooper, A note on the biologic concept of race and its application in epidemiologic research. *American Heart Journal* **108**, 715 (1984).

4. T. A. LaVeist, Race, ethnicity, and health: A public health reader. *Recherche* **67**, 02 (2002).

5. B. S. McEwen, T. Seeman, Protective and damaging effects of mediators of stress: elaborating and testing the concepts of allostasis and allostatic load. *Annals of the New York Academy of Sciences* **896**, 30 (1999).

6. A. T. Geronimus, Black/white differences in the relationship of maternal age to birthweight: a population-based test of the weathering hypothesis. *Social science & medicine* **42**, 589 (1996).

7. M. J. Hetherington, A. M. D. Owens, Listen: How voters view authority speaks volumes about political views-Marc Hetherington co-authors book on increasing polarization between GOP and Dems. (2009).

8. D. A. Hibbs, Political parties and macroeconomic policies and outcomes in the United States. *The American Economic Review* **76**, 66 (1986).

9. J. S. Hacker, P. Pierson, *Winner-take-all politics: how Washington made the rich richer, and turned its back on the middle class*. (Simon & Schuster, 2011).

10. E. R. Tufte, *Political control of the economy*. (Princeton Univ Pr, 1980).

11. L. M. Bartels, *Unequal democracy: The political economy of the new gilded age*. (Princeton Univ Pr, 2010).

12. V. Navarro, L. Y. Shi, The political context of social inequalities and health. *Social science & medicine* **52**, 481 (Feb, 2001).

13. V. Navarro *et al.*, Politics and health outcomes. *The Lancet* **368**, 1033 (2006).

14. U. Nations, Department of Economic and Social Affairs, Population Division. World Population Prospects: The 2010 Revision, CD-ROM Edition., (2011).

15. G. F. Anderson, J. P. Poullier, Health spending, access, and outcomes: Trends in industrialized countries. *Health Affair* **18**, 178 (May-Jun, 1999).

16. J. H. Aldrich, Southern Parties in State and Nation*. *Journal of Politics* **62**, 643 (2000).

17. D. O. Sears, Symbolic racism. *Eliminating racism: Profiles in controversy*, 53 (1988).

18. N. A. Valentino, D. O. Sears, Old times there are not forgotten: Race and partisan realignment in the contemporary South. *American Journal of Political Science* **49**, 672 (2005).

19. L. The, The pre-election health environment in the USA. *The Lancet* **380**, 187 (2012).

20. J. E. McDonough, The Road Ahead for the Affordable Care Act. *New England journal of medicine* **367**, 199 (2012/07/19, 2012).

21. J. Calmes, J. Steinhauer, and J. Rutenberg, G.O.P. vowing to take battle into November. New York Times, pp. A1 (2012, June 29).

22. A. Finkelstein, Static and dynamic effects of health policy: Evidence from the vaccine industry. *the Quarterly Journal of economics* **119**, 527 (2004).

23. J. Macinko, F. C. Guanais, M. F. M. de Souza, Evaluation of the impact of the Family Health Program on infant mortality in Brazil, 1990–2002. *Journal of Epidemiology and Community Health* **60**, 13 (2006).

24. O. Blanchard, R. Perotti, An empirical characterization of the dynamic effects of changes in government spending and taxes on output. *the Quarterly Journal of economics* **117**, 1329 (2002).

25. S. Harper, J. Lynch, Methods for measuring cancer disparities: using data relevant to Healthy People 2010 cancer-related objectives. (2005).

26. M. Ravallion, World Bank., in *Policy research working paper 3243*. (World Bank,, Washington, D.C., 2004).

27. B. Efron, T. Hastie, I. Johnstone, R. Tibshirani, Least angle regression. *The Annals of statistics* **32**, 407 (2004).

28. A. Zellner, An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. *Journal of the American statistical Association*, 348 (1962).

29. For example, the DW-Nominate score—a measure of the president's ideological position where -1 is extremely liberal and 1 extremely conservative—for Nixon is .49 and for Ford is .50 while for Reagan, Bush Sr., and Bush Jr. are .73, .63, and .90, respectively.

30. This is, of course, due to differences in racial IMR growth rates. During Democratic presidents, despite that the absolute decrease of BIMR was higher compared to the absolute decrease of WIMR, BIMR relative decrease was lower than that of WIMR. Accordingly, the rate denominator contracts faster than the nominator, and the IMR race ratio increases in spite of diminishing overall IMRs. Since rate changes between BIMR and WIMR are at about 2.0 (i.e., double increase under Republican presidents and double decrease under Democratic presidents), IMR race ratios continue to increase even under contraction periods iff IMR race ratios were over 2.0.

31. Thus small changes in the data could impact estimates.

32. J. Kmenta, R. F. Gilbert, Small sample properties of alternative estimators of seemingly unrelated regressions. *Journal of the American statistical Association*, 1180 (1968).

33. For example, using data from the United Nations Population Division, IMRs between 2005 and 2010 for Singapoure, China (Hong Kong, SAR), Iceland, Luxemburg, and Sweden (the 5 world nations/territories with the lowest IMR) were 1.9, 2.0, 2.1, 2.3, and 2.6 respectively, while the IMR for the United States was 6.8. The difference between the IMR of the U.S. and the average IMR of these top-5 nations/territories is 4.6 units. The total IMR increase in 28 typical years of Republican administration was 5.04 units (.18 * 28 = 5.04).

APPENDICES

Propensity Scores Estimation for

Age, SES, Mortality, and Socio-political Participation

Variables Included in the Estimation of the Propensity Scores

Survival status. This is the dependent variable, taken from the 10-year epidemiological/mortality follow up. This variable is dichotomous (0 = person died between 1995/1996 and 2005, 1 = person survived).

Race/Ethnicity. There is a bulk of evidence addressing health disparities and mortality gaps between Blacks and other ethnicities, especially with Whites (Hayward et al. 2000; Xu, Murphy and Tejada-Vera 2010). This variable is dichotomous (2 = Black, 1 = Other races/ethnicities).

Self-rated physical health. Current research finds that self-rated physical health is associated with subsequent mortality, showing an independent effect from biomedical variables (Finch et al., 2001; Idler and Benyamini 1997). This is a 5-category variable (1 = Poor self-rated physical health, 2 = Fair, 3 = Good, 4 = Very good, and 5 = Excellent). Separate dummies for each category are used in the model.

Mental or emotional health. Self-reported mental health is commonly associated with other health conditions mostly linked to distress, behaviors, and social interactions that increase the susceptibility to physical illnesses related to mortality (Berkman et al. 2000; Prince et al. 2007). This is a 5-category variable (1 = Poor self-rated mental or emotional health, 2 = Fair, 3 = Good, 4 = Very good, and 5 = Excellent).

Self-rated present general health. Self-reported general health is an independent, substantive predictor of mortality and health status (Phillips, Der and Carroll 2010; Schoenfeld et al. 1993). This variable ranges from 0 to 10, with 0 being worst possible general health and 10 being best possible.

Sum of chronic conditions. This variable is the sum of 29 chronic conditions the respondent experienced or was treated for, including cardiovascular, autoimmune, and neurological disorders among others. The resulting index range from 0 to 27, and it was subsequently censored into an 11-category variable with values from 0 (no chronic conditions) to 10 (10 or more chronic conditions). Separate dummies for each category are used in the model.

Alcohol consumption. Research has found both detrimental and beneficial effects of alcohol consumption over health and mortality; a clearer adverse impact on health outcomes has been related to alcoholism or high alcohol consumption (Emberson et al. 2005; Liao et al. 2000). This variable is the sum of 5 questions that tap on alcohol hazards and consumption as related to chances of getting hurt, emotional problems, and alcohol addiction. The result is a 6-category variable ranging from 0 to 5. This variable is dichotomized across its values in the model.

Waist-to-hip ratio (WHR) and body mass index (BMI). These variables are measures of under/over weight and obesity, both associated with higher mortality (Flegal et al. 2005; Visscher et al. 2000). BMI is a measure of overall obesity while WHR assesses the distribution of abdominal fat, which is related to certain diseases independent from overall obesity (Epel et al. 2000; Willett 1998). The WHR was measured as the ratio of waist circumference (at the umbilicus) to hip circumference (at the largest point). The BMI is the weight (in kilograms)

divided by the square of the height (in meters). Both are continuous variables and range from 14.4 to 64.0 and 0.46 to 1.39, respectively.[38]

Hospitalization and physician visits index. A coefficient of reliability (Cronbach's alpha) computed on 3 indicators of hospitalization and visits to physicians related to illness ($\alpha = 0.64$). The 3 items assessed the number of times the respondent had been hospitalized, total number of days spent hospitalized, and total number of visits to physicians. This is a continuous variable ranging from 0 (no hospitalizations and no visits to physicians) to 4 (maximum level of hospitalizations and visits to the physician).

Neighborhood quality. This is a coefficient of reliability (Cronbach's alpha) computed on 4 items that tap on the respondent's perception of the quality of his/her neighborhood ($\alpha = 0.68$). The items prompt participants on feelings of safety, social support, and interpersonal trust in the neighborhood where they live. This latent variable represents an overall assessment of the neighborhood's social environment, which has been found to correlate with mortality (Kawachi and Berkman 2003; Wen, Hawkley and Cacioppo 2006). This is a continuous variable that ranges from 1 (lowest perceived neighborhood quality) to 4 (highest perceived neighborhood quality).

Smoking. There is a bulk of evidence that shows that smoking is related to morbidity and higher mortality risk (Doll et al. 2004; Jacobs et al. 1999). Research has also shown that mortality linked to smoking differs between males and females (Hunt et al. 2011). For this reason, this variable is interacted by gender in the model. This variable assesses the average daily number of cigarettes the respondent smoked during the year in which he/she smoked most

[38] Underweight is associated with a body mass index <18.5, and a high level of obesity with a body mass index ≥30 (c.f. Flegal et al., 2005; Katzmarzyk et al., 2001). A Waist-to-hip ratio over 0.85 for women and over 0.9 for men is usually considered a high-risk cut-point above which obesity-related health problems begin to manifest (Alberti and Zimmet, 1998; Seeman et al., 2008).

heavily. It was censored into a 5-category variable (Non-smokers = 0, 1 to 10 cigarettes per day = 1, 11 to 20 = 2, 21 to 30 = 3, and 30+ = 4). Separate dummies for each category are used in the model.

Gender. Research has shown life expectancy and burden of disease gaps between females and males, with females living longer while experiencing a higher burden of chronic illnesses than males (Annandale 2009; Gorman and Read 2007). This variable is dichotomous (Males = 0, Females = 1).

Table 1A (below) shows the logistic regression output for the estimation of the probability of survival (i.e., the propensity score). The output of this regression is in essence uninteresting, however. The goal of this procedure is to assemble an informative model of survival agnostic to model-based interpretations.[39]

[39] For instance, research on propensity scores argues that it would be inconsequential to evaluate the predictive accuracy of the model, including the proportion of the explained variation in survival status. The main goal of this step is simply the estimation of a less-biased effect of mortality over participation (Harrell 2001; Heinze and Jüni 2011; Schemper and Henderson, 2000).

Table 1A. Logistic Regression Output for the Estimation of the Probability of Survival (Propensity Score)

Covariates	Coeff. (SE)
Black (white=1, black=2)	-.16
	(.25)
Perceived neighborhood quality (1 to 4)	-.25**
	(.11)
Physical health rate (=2)	.63**
	(.25)
Physical health rate (=3)	1.30***
	(.26)
Physical health rate (=4)	1.95***
	(.29)
Physical health rate (=5)	2.26***
	(.36)
Mental or emotional health (1 to 5)	-.21***
	(.06)
Self-rated general health (0 to 10)	.02
	(.04)
Sum chronic conditions (=1)	-.40**
	(.18)
Sum chronic conditions (=2)	-.16
	(.20)
Sum chronic conditions (=3)	-.53***
	(.199)
Sum chronic conditions (=4)	-.34
	(.23)
Sum chronic conditions (=5)	-.13
	(.28)
Sum chronic conditions (=6)	-.15
	(.31)
Sum chronic conditions (=7)	-.65**
	(.30)
Sum chronic conditions (=8)	-.46
	(.42)
Sum chronic conditions (=9)	-.49
	(.42)
Sum chronic conditions (=10)	-.87**
	(.34)
Alcohol consumption (=1)	.47**
	(.22)
Alcohol consumption (=2)	.34
	(.36)
Alcohol consumption (=3)	-.18
	(.45)
Alcohol consumption (=4)	.75
	(.85)

Alcohol consumption (=5)	.44
	(.76)
Waist-to-hip ratio (0.46 to 1.39)	-4.15***
	(.77)
Body mass index (14.4 to 64)	.038***
	(.01)
Hospital/physician visit index (0 to 4)	-.40***
	(.09)
Smoking cigarettes (=1)	.43
	(.78)
Smoking cigarettes (=2)	-1.06***
	(.24)
Smoking cigarettes (=3)	-.26
	(.38)
Smoking cigarettes (=4)	-.52**
	(.23)
Gender (male=0, female=1)	-.11
	(.16)
Smoking cigarettes (=1) x Gender (=1)	-1.20
	(.86)
Smoking cigarettes (=2) x Gender (=1)	.76**
	(.35)
Smoking cigarettes (=3) x Gender (=1)	-.27
	(.51)
Smoking cigarettes (=4) x Gender (=1)	-.18
	(.38)
Constant	6.1***
	(.85)
Observations	5,531

Statistical significance code: *** p<0.01, ** p<0.05, * p<0.1
Note 1: Dependent variable is survival status (0=deceased, 1=survivor).
Note 2: Robust standard errors correcting for non-normality and family membership clusters in parentheses.

Appendix References

Alberti, George, and Paul Zimmet. 1998. "Definition, Diagnosis and Classification of Diabetes Mellitus and its Complications. Part I: Diagnosis and Classification of Diabetes Mellitus Provisional Report of a WHO Consultation." *Diabetic Medicine* 15: 539-53.

Annandale, Ellen. 2009. *Women's Health and Social Change.* New York: Routledge.

Berkman, Lisa F., Thomas Glass, Ian Brissette, and Teresa Seeman. 2000. "From Social Integration to Health: Durkheim in the New Millennium." *Social Science & Medicine* 51: 843-57.

Doll, Richard, Richard Peto, Jilian Boreham, and Isabelle Sutherland. 2004. "Mortality in Relation to Smoking: 50 Years' Observations on Male British Doctors." *British Medical Journal.* doi:10.1136/bmj.38142.554479.AE

Emberson, Jonathan R., A. Gerald Shaper, S. Goya Wannamethee, Richard W. Morris, and Peter H. Whincup. 2005. "Alcohol Intake in Middle Age and Risk of Cardiovascular Disease and Mortality: Accounting for Intake Variation over Time." *American Journal of Epidemiology* 161 (9): 856-63.

Epel, Elissa S., Bruce McEwen, Teresa Seeman, Karen Matthews, Grace Castellazzo, Kelly D. Brownell, Jennifer Bell, and Jeannette R. Ickovics. 2000. "Stress and Body Shape: Stress-Induced Cortisol Secretion is Consistently Greater among Women with Central Fat." *Psychosomatic Medicine* 62: 623-32.

Finch, Brian Karl, Robert A. Hummer, Bohdan Kol, and William A. Vega. 2001. "Discrimination and Acculturative Stress in the Physical Health of Mexican-origin Adults." *Hispanic Journal of Behavioral Sciences* 23: 399-429.

Flegal, Katherine M., Barry I. Graubard, David F. Williamson, and Mitchell H. Gail. 2005. "Excess Deaths Associated with Underweight, Overweight, and Obesity." *The Journal of the American Medical Association* 293 (15): 1861-67.

Gorman, Bridget K., and Jen'nan Ghazal Read. 2007. "Why Men Die Younger than Women." *Geriatrics and Aging* 10 (3): 182-91.

Harrell, Frank E. 2001. *Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression and Survival Analysis*. New York: Springer Verlag.

Hayward, Mark D., Eileen M. Crimmins, Toni P. Miles, and Yu Yang. 2000. "The Significance of Socioeconomic Status in Explaining the Racial Gap in Chronic Health Conditions." *American Sociological Review* 65 (6): 910-30.

Heinze, Georg, and Peter Jüni. 2011. "An Overview of the Objectives and the Approaches to Propensity Score Analyses." *European Heart Journal* 32 (14): 1704-08.

Hunt, Kate, Gerry McCartney, Lamia Mahmood, Alastair H. Leyland, and David G. Batty. 2011. "Contribution of Smoking—and Alcohol-related Deaths to the Gender Gap in Mortality: Evidence from 30 European Countries." *British Medical Journal.*

Idler, Ellen L., and Yael Benyamini. 1997. "Self-rated Health and Mortality: A Review of Twenty-Seven Community Studies." *Journal of Health and Social Behavior* 38 (1): 21-37.

Jacobs, David R., Hisashi Adachi, Ina Mulder, Daan Kromhout, Alessandro Menotti, Aulikki Nissinen, and Henry Blackburn. 1999. "Cigarette Smoking and Mortality Risk." *Archives of Internal Medicine* 159: 733-40.

Katzmarzyk, Peter T., Cora L. Graig, and Claude Bouchard. 2001. "Original Article Underweight, Overweight, and Obesity: Relationships with Mortality in the 13-year Follow-up of the Canada Fitness Survey." *Journal of Clinical Epidemiology* 54: 916-20.

Kawachi, Ichiro and Lisa Berkman. 2003. *Neighborhoods and Health*. New York: Oxford University Press.

Liao, Youlian, Daniel L. McGee, Guichan Cao, and Richard S. Cooper. 2000. "Alcohol Intake and Mortality: Findings from the National Health Interview Surveys (1988 and 1990)." *American Journal of Epidemiology* 151 (7): 651-59.

Phillips, Anna C., Geoff Der, and Douglas Carroll. 2010. "Self-reported Health, Health-reported Fitness, and All-cause Mortality: Prospective Cohort Study." *British Journal of Health Psychology* 15 (2): 337-46.

Prince, Martin, Vikram Patel, Shekhar Saxena, Mario Maj, Joanna Maselko, Michael R. Phillips, and Atif Rahman. 2007. "No Health without Mental Health." *The Lancet* 370 (9590): 859-77.

Schemper Michael, and Robin Henderson. 2000. "Predictive Accuracy and Explained Variation in Cox Regression." *Biometrics*, 56 (1): 249-255

Schoenfeld, David E., Lynda C. Malmrose, Dan G. Blazer, Deborah T. Gold, and Teresa Seeman. 1993. "Self-rated Health and Mortality in the High-functioning Elderly—A Closer Look at Healthy Individuals: MacArthur Field Study of Successful Aging." *Journal of Gerontology* 49 (3): 109-15.

Seeman, Teresa, Sharon S. Merkin, Eileen Crimmins, Brandon Koretz, Susan Charette, and Arun Karlamangla. 2008. "Education, Income, and Ethnic Differences in Cumulative Biological Risk Profiles in a National Sample of US Adults: NHANES III (1988-1994)." *Social Science and Medicine*, 66: 72-87.

Visscher, Tommy L., Jacob C. Seidell, Alessandro Menotti, Henry Blackburn, Aulikki Nissinen, Edith J. Feskens, and Daan Kromhout. 2000. "Underweight and Overweight in Relation to Mortality among Men Aged 40-59 and 50-69 Years: The Seven Countries Study." *American Journal of Epidemiology* 151: 660-66.

Wen, Ming, Louise C. Hawkley, and John T. Cacioppo. 2006. "Objective and Perceived Neighborhood Environment, Individual SES and Psychosocial Factors, and Self-rated Health: An Analysis of Older Adults in Cook County, Illinois." *Social Science and Medicine* 63: 2575-90.

Willett, Walter. 1998. *Nutritional Epidemiology.* 2nd ed. New York: Oxford University Press.

Xu Jiaquan, Kenneth D. Kochanek, Sherry L. Murphy, and Betzaida Tejada-Vera. 2010. "Deaths: Final Data for 2007. National Vital Statistics Reports." Hyattsville, MD: National Center for Health Statistics.

Supporting [Online] Material for

Infant Mortality and the President's Party

1. Description of general methods

1.1 Median cubic b-splines

The data was de-trended to disassociate long- from short-term effects. Residual variation could therefore be more appropriately linked to policies directly enacted by presidents while in office, independent from effects attributable to historical trends. These historical trends were determined by fitting a median cubic b-spline (*1-4*). Residuals were subsequently recovered and constituted the variables implemented in analyses thereafter. All splines were functions from time (i.e., year of measurement) to the respective dependent variable (e.g., IMR, education, health expenditure). Splines for all de-trended variables used knots fixed at minimum and maximum values of time range plus equidistant internal knots fixed at 33.3 and 66.7% of time range. For example, splines fitted for WIMR and BIMR (1965-2010) used fixed knots at 1965, 1980, 1995, and 2010 (Fig. S1); splines fitted for alcohol consumption (1965-2007) used fixed knots at 1965, 1979, 1993, and 2007, and so forth. This simple yet efficient at smoothing the data version of the spline was chosen; it helped to avoid overfitting and to diminish autocorrelation between the knots. Because a cubic spline is a linear combination of the cubic "plus-functions" derived at the knots (*3*), it is possible that trends would have been different depending on the number of knots and/or the localization of the knots throughout the x-axis variable range. Accordingly, alternative splines were fitted for each variable. For example, splines with fixed knots at minimum and maximum, 25, 50, and 75% of time range, and also by moving knots by 1 year-unit to the left and right of originally fixed knots (e.g., if a knot was originally fixed at 1980, the spline was re-run using 1979 and 1981 as alternative knots, one at a time). Trend

differences (and therefore differences in recovered residuals) across variables were trivial. The splines were fitted using STATA 11.

1.2 Medians and means comparison tests

To gauge the differences between Democratic and Republican presidents, differences in medians and means by president's party were tested for all variables included in analyses. Because median and mean comparisons can sometimes differ under small-sample circumstances, the probability that a random draw from any given variable during Democratic presidents was higher than a random draw from the same variable under Republican presidents was also estimated (*5*). The selected test for median differences between the parties was the Wilcoxon Rank-sum Test (*6*) also known as the Mann-Whitney Two-sample Statistic (*7*), which produces identical results as the K-sample Kruskal-Wallis Test for equality of medians with two groups (*8*). Comparison of the means was done using Student's T-tests (via Satterthwaite's approximation formula (*9*)).

Parametric cross-validation of differences in infant mortality were run using DW-Nominate scores instead of president's party through quantile regression at the median and through ordinary linear regression with robust correction of standard errors (Table S1). If infant mortality variation is related to the policy platforms imparted by the party that controls the presidency, then a measure of the president's ideological position should reproduce these findings. Explicitly, the quantile [median] regressions fitted what is known as Least-absolute-value models and Minimum L1-norm models to the data (*5, 10*). Both types of regressions attempt to correct for outlier sensitivity. However, because of small sample conditions, it is possible that our robust estimations are still influenced by one or a few special cases (*11*). For this reason, influence Delta-beta analysis was carried out (*12*). For racial IMRs with 46 year

data-points, the models were respectively re-estimated 46 times, each time removing one year data-point from the model fitting. The difference between the president's party coefficient estimated with complete data (n=46) and the coefficient estimated without one year data-point (n=45) (i.e., the "delta-beta") was computed every time. Differences in the *p*-values were also estimated (Table S2). All analyses were run in STATA 11.

1.3 Least Angle Regression

Least Angle Regression (LAR) is a model-building algorithm that provides all possible LASSO (l_1 penalized regression) (*13*) estimates while efficiently using a forward stepwise mechanism. The variable selection procedure works by entering variables in small steps into the model. It starts with all coefficients equal to 0, then the algorithm chooses the independent variable with the highest correlation with the dependent variable and adds it into the model. At that time the algorithm assigns a small value to the coefficient in the direction of the sign of the correlation, re-estimates the residuals of the response variable, and proceeds adding to the size of the coefficient until another predictor becomes equally correlated and earns its way into the model, then increasing both coefficients in their joint least squares direction until another predictor is chosen into the model, and so forth. The fine model-selection properties of LAR are elaborated somewhere else (*14, 15*). In this article LAR was implemented to explore the importance and robustness of the president's party at predicting infant mortality rates free of a priori model specifications. Importance was revealed by how early the president's party was selected into the model and by the size of its coefficient in the model that minimized Mallows' Cp statistic (Figs. 2 and S2; Tables S4a-S4c) (*16*). Coefficient paths were used as visual representations of robustness, for they show the size at which the president's party coefficient explains the most of IMR residual variance while other variables are chosen into the model and

the sizes of coefficients vary in direction of prediction accuracy. Because high multicollinearity could still generate unstable coefficients (*15*), and because there were several variables measuring the same phenomenon, sets of variables submitted to LAR were reduced to variables that helped diminish multicollinarity, yet all sets included variables representing most key factors known to affect IMRs (Fig. 2). Analyses were run using the STATA 11 package written by Adrian Mander, MRC Biostatistics Unit, Cambridge, UK.

1.4 Linear Seemingly Unrelated Regressions

Racial infant mortality rates are inbuilt components of national infant mortality rates. Yet in spite of being measures of the very same inherent phenomenon they seem to be subjected to different explanatory equations. Even when using explanatory equations comprised by the same predictors, it is expected that these predictors would have different independent effects on racial and national mortality rates. Under such circumstances, the explanatory equations are linked to each other through their joint error term distribution. In this article Linear Seemingly Unrelated Regressions (SUR) (*17, 18*) were used to investigate these three relationships at the same time; respective parameters of the president's party on IMR, WIMR, and BIMR were estimated while exploiting the correlation between the cross-equation contemporaneous disturbances to obtain efficient estimates, even more so than through OLS or maximum likelihood estimation. The nice properties of SUR are also shown to be advantageous when the sample size is moderate or small (*19-21*). The analyses were run using STATA 11.

1.5 Sensitivity Analyses on Structural Equation Models

Structural equation modeling (SEM) offers a convenient framework to analyze mediation dynamics (*22*). SEM allowed for the organization of the infant mortality predicting mechanism in a way that mediation effects were readily estimated. This was propitious for the purposes of

this article because the effects that the president's party exert over infant mortality are necessarily mediated by indicators affected by policy. In other words, it is mainly through the analysis of the variation in the proximal determinants of health imparted by policy that we can understand the connections between infant mortality and the president's party. Accordingly, a four-variable model with two mediators was proposed (Fig. 3a). This parsimonious model suited our small *n* circumstances without overly compromising the reliability of estimations. The contemporaneous error terms of mediators were permitted to correlate thus enhancing the stability of parameter estimation given instability linked to both small *n* and multicollinearity. All structural models were estimated via maximum likelihood. Despite that maximum likelihood (ML) is regarded as a "large-sample" method of estimation, there is ample simulation research that shows that mediated effects can be meaningfully assessed in simple structural models via ML (without latent factors) when the sample size is small (*23*). All models successfully converged.

Considering small *n* limitations and that the president's party could affect infant mortality via various indicators, and even in different manners depending on different measures of the same phenomenon, sensitivity analyses were run on the baseline structural model. The technique implemented herein is an adaptation of Leamer's extreme-bound analysis (*24*) and Levine and Renelt's sensitivity analysis (*25*) originally applied in regression analysis. The technique consisted in allowing mediators to "rotate" one at a time and in sequence. In this manner, the indirect and direct effects of the president's party over IMRs were assessed in all possible mediation combination models (1,218 in total, 406 models for IMR, WIMR, and BIMR respectively). Sensitivity analyses were run three times, each time using IMR, WIMR, and BIMR, respectively, as the dependent variable, and the president's party as the independent

variable of interest. Roughly the "rotating" mechanism worked as follows: Out of 29 possible, literature-identified mediators (R_1, R_2… R_{29}), a first structural model was assembled by inserting R_1 and R_2 as mediators (Fig. 3a), then the model was run and parameter estimates were recovered. Subsequently, a second model was assembled using R_1 and R_3 (i.e., switching mediator R_2 for R_3), then a third one using R_1 and R_4, and so forth until all of the 28 [different from R_1] mediators shared the model with R_1. At that time, a new model was assembled, now beginning with R_2 and R_3, and then another with R_2 and R_4, and so forth until exhausting all possible mediation-model combinations.

These sensitivity analyses represented a way to incorporate into the analysis multiple, competing theoretical expectations that are the focus of current research. Inferences on mediation dynamics became less model-dependent, too. To test the hypothesis advanced herein by exposing the relationship between infant mortality and the president's party to multiple models increased its level of refutability. By altering the literature-driven set of informative mediators via "rotation" we were able to examine the decrease of direct *vs.* the increase of indirect effects thus identifying those indicators that reliably mediate the relationship between the president's party and infant mortality rates. Finally, this statistical application also offered a simple yet inclusive opportunity to situate such unexplored relationship within the extensive framework of current literature. Analyses were run using the "R Interface to EQS" (*26*) on EQS6.2.[40]

2. Description of variables and data sources

All variables implemented in this article are listed in Table 1 of the manuscript. There are variables for all races, and when available or applicable, these variables were stratified by race

[40] The code for the R Interface to EQS was written in conjunction with Analyst Eric Wu, Department of Statistics, UCLA.

(white and black). Variables by race were respectively implemented in models where the dependent variable was racial IMR. There is some redundancy across indicators, however (e.g., there is a poverty indicator for *females* and another for *females without husband*, and so forth). But so is the nature of the literature. Different associations have been found using different variables assessing the same phenomenon (e.g., different measures of income and income inequality are shown to have positive and negative associations with indicators of public health (*27*)). It is not the aim of this article to become part of this fruitful debate; rather, it is to expose the relationship between infant mortality and the president's party to possible confounders and mechanisms already identified in the literature. The set of variables implemented in this study are by no means exhaustive; it attempts to be, however, representative of what is commonly found in current research. A description of the variables and their respective data sources follows.

2.1 President's party (partisan control)

This is the independent variable of interest. It was lagged by one year because policies enacted by presidents do not produce immediate effects. This lag is in accordance with current literature (see manuscript). The variable was coded as 1 if partisan control belonged to a Republican president and 0 if partisan control belonged to a Democratic president. The period of partisan control expands from the year after inauguration-year (i.e., the second year the president was in office) to the next inauguration-year. For example, partisan control for Clinton initiated in 1994 and ended in 2001; for Bush Jr. initiated in 2002 and ended in 2009, and so forth.

2.2 Infant Mortality Rates

These are the dependent variables of interest (for all races, whites, and blacks). Infant mortality rate (IMR) is defined as the number of deaths of infants under 1 year of age per 1,000

live births. The data are for both sexes. The data are from the National Vital Statistics Reports (*28-31*). Data for both whites and blacks include individuals of Hispanic origin. The data are used in this format to avoid dealing with several definitions of the infant mortality measurement, mainly because information on Hispanic origin is unavailable for all years of the period under observation. The data, however, do not differ drastically; for example, IMR for (total) whites in 2010 was 5.19 and for non-Hispanic whites was 5.09 (data for 2009 and 2008 were 5.30 *vs.* 5.25 and 5.54 *vs.* 5.52 respectively). Racial disparities in IMR are therefore slightly underestimated.

2.3 Low Birthweight Rate and Preterm Births Rate

Low birthweight rate (LBW) is defined as the percent of infants weighting less than 2,500 grams at birth. Preterm births rate (PB) is defined as the percent of infants born at less than 37 weeks of gestation. The data are for both sexes and all races. These two variables are widely identified to be crucial predictors of IMR (*32, 33*) and are among the leading determinants of racial disparities in IMRs (*34*). The data are from the National Vital Statistics Reports (e.g., *35, 36*) and the Monthly Vital Statistics Reports, Natality Statistics (e.g., *37*), various years.

2.4 Recessions

The negative and "protective" effects of economic recessions on health indicators like infant mortality rates have been widely discussed (*38-41*). This variable was not de-trended. The data are from the National Bureau of Economic Research (NBER) and it is available to the public online at http://www.nber.org/cycles.html The variable was coded 1 for recession-year and 0 otherwise. For example, the last economic contraction dated from December 2007 through June 2009, thus 2008 and 2009 were coded as 1; the economic contraction of 2001 dated from March 2001 through November 2001, thus the year 2001 was coded 1, and so forth. A description of how the NBER determines the dating of a recession can be accessed online at

http://www.nber.org/cycles/dec2008.pdf for the peak in economic activity and at http://www.nber.org/cycles/sept2010.pdf for the end of the recession.

2.5 Income inequality

The negative and positive associations between income inequality—measured in different forms and through different indexes, nationally and cross-nationally—have produced a wide range of research in the social determinants of health literature, including its association to infant mortality (*27, 42-45*). Variables included in this study were:

a. Gini for family income (for all races, whites, and blacks). Data for "white alone" and "black alone" were chosen between 2002 and 2010 and for "white" and "black" between 1965 and 2001 instead of using Hispanic-origin data because these measures offer higher measurement consistency throughout the time series. The data come from the U.S. Census Bureau, Current Population Survey, Annual Social and Economic Supplements.
b. Income share (for all races, whites, and blacks respectively), and the share ratio (for all races) of those located at the top 5% and the bottom 20% of the aggregate family income distribution. The data come from the U.S. Census Bureau, Current Population Survey, Annual Social and Economic Supplements.
c. Women's earnings as a percentage of men's earnings. The data are based on median earnings of full-time, year-round workers 15 years old and over. Earnings data between 1965 and 1988 include civilian workers only. The data come from the U.S. Census Bureau, Current Population Survey, Annual Social and Economic Supplements.

2.6 Poverty

Infant mortality is highly concentrated among the poor (*46*) and has been found to be strongly associated with racial differences in IMR (*47-49*). Particularly high infant mortality

rates are found among females who are poor and unmarried (*50*). Due to these strong associations, analyses included poverty rates (for all races, whites, and blacks), poverty rates for females, and poverty rates for families with female householder where the husband was not present. Other measures like the income share of the bottom 20% (see measures of income inequality above) and the mean income for families located at the bottom 20% of the family income distribution (for all races, whites, and blacks) could also be interpreted as poverty indicators. The mean incomes for families located at the bottom 20% also represent the income component that is mostly related to infant mortality. For data stratified by race, "white alone" and "black alone" were chosen between 2002 and 2010 and for "white" and "black" between 1965 and 2001 instead of using Hispanic-origin data because these measures offer higher measurement consistency throughout the period under observation. The data come from the U.S. Census Bureau, Current Population Survey, Annual Social and Economic Supplements.

2.7 The uninsured, and financial and economic indicators

Inflation, unemployment, and the cost of medical care have been widely used in public health research as either proxies for or confounders of macroeconomic indicators and/or access to medical services that are known to affect health outcomes including infant mortality. To have or not health insurance coverage is also known to determine access to health services key to infant mortality such as prenatal care. These various associations are, for example, due to a high percent of uninsured individuals among the unemployed or to a lower access to prenatal care among the uninsured and the poor (*51-58*). A brief description of these variables follows.

a. The uninsured. This variable was the percentage of individuals under 65 years of age without health insurance coverage. "Without health insurance coverage" means that the individual did not have any type of private or government-sponsored health plan,

Medicare or Medicaid, Children's health insurance, or military plan. The person will also be considered uninsured if s/he only had Indian Health Service coverage or a one-service type of private plan (e.g., dental care). The data are from the National Health Interview Survey (the National Health Statistics Reports and the National Center for Health Statistics (*59, 60*)).

b. For a historical estimate of inflation the Consumer Price Index Research Series for all urban consumers (CPI-U-RS) was used. For a historical estimate of the changing costs of medical care services the Price Change for Medical Care Services in the CPI was implemented. The data come from the Department of Labor, Bureau of Labor Statistics. The data for the CPI-U-RS can be accessed online at http://www.census.gov/hhes/www/income/data/incpovhlth/2010/CPI-U-RS-Index-2010.pdf Further information on the CPI-U-RS is available online at http://www.bls.gov/opub/mlr/1999/06/art4abs.htm Information on the Price Change for Medical Care Services in the CPI is available at http://www.bls.gov/cpi/cpifact4.htm/

c. Unemployment rate (for all races, whites, and blacks). The variable was the unemployed as percent of the civilian labor force (persons of 16 years of age and over). The data are from the Department of Labor, Bureau of Labor Statistics. Blacks between 1965 and 1971 defined as "black and other", between 1972 and 2003 defined as "black", and between 2004 and 2010 defined as "black or African American". Prior to 2003, individuals who selected more than one race were assigned the race they signaled as their main race. More information on classification methodology can be accessed online at http://www.bls.gov/cps/documentation.htm#comp Information on how the government measures unemployment can be found online at http://www.bls.gov/cps/cps_htgm.pdf

2.8 Education

It is well known that infant mortality declines as education increases. Considering that education is a crucial component of socioeconomic status, racial differences in education are reflected in racial differences in health, including differences in IMR (*61-64*). The variable implemented in the analyses was percent of non-institutionalized individuals 25 years and over who completed 4 years of high school education or more (for all races, whites and blacks (both sexes), and whites and blacks (females)). The data come from the U.S. Census Bureau, Current Population Survey, Annual Social and Economic Supplements. Data for blacks between 1965 and 2002 are for those individuals who self-identified as "blacks only". Data from 2003 to 2010 are for those individuals who indicated only one racial identity. Both blacks and whites include individuals of Hispanic origin. The data was preferred in this format because it offered consistency throughout the period under observation (data based on Hispanic-origin do not).

2.9 Health expenditure

Measures of health care expenditure are useful proxies for the amount and/or quality of the technology, surveillance, infrastructure, and human capital invested in the provision of health care goods and services and in disease control and prevention, thus impacting the health status of the population at large including infant mortality (*65*). Two types of health expenditures were implemented, each specified as percent of GDP and per capita (i.e., specified expenditure divided by the size of the population). The first was the total national health expenditures on maternal/child health. The second was the arithmetic sum of the following federal expenditures on *maternal/child* health items: Personal health care, hospital expenditures, physician and clinical expenditures, prescription drug expenditures, and durable medical equipment. The sources of the data are the Centers for Medicare & Medicaid Services, Office of the Actuary,

National Health Statistics Group; the U.S. Department of Commerce, Bureau of Economic Analysis; and the U.S. Bureau of the Census. The data are available to the public from the National Health Expenditure Accounts website at http://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData/NationalHealthAccountsHistorical.html

2.10 Alcohol and tobacco consumption

For a long time research has shown the adverse pregnancy outcomes associated with behaviors such as smoking (*66-68*) and alcohol consumption (*69-71*). The variable for tobacco consumption was the adult population (18 years and over) per capita yearly consumption of manufactured cigarettes in the United States. The data are from the Economic Research Service of the United States, Department of Agriculture, and are available to the public from the Centers for Disease Control and Prevention website at http://www.cdc.gov/tobacco/data_statistics/tables/economics/consumption/index.htm The variable for alcohol consumption was the apparent per capita yearly consumption of ethanol (measured in gallons of ethanol). The data are for individuals 15 years of age and older prior to 1970 and for individuals 14 years and older thereafter. The data are from the National Institutes of Health, Institute on Alcohol Abuse and Alcoholism.

2.11 Age of mother

Maternal age is one important variable in the literature of infant mortality and its proximal determinants (*33, 72*). For example, gestational hypertension increases notably for women age 40 and over; multiple births, which are associated with low birthweight, increase with older maternal age (*35*). Research also shows that maternal age patterns in infant mortality vary between racial groups (*73, 74*). The variable used in the analyses was the median age of the

mother at the first live-birth (for all races, whites, and blacks). Race goes by race of the child from 1965 to 1979, and by race of the mother thereafter. Data from 1970 onward exclude births of nonresidents of the U.S. The data are from the Centers for Disease Control and Prevention, National Center for Health Statistics. The data are available online at http://www.cdc.gov/nchs/data/statab/natfinal2003.annvol1_05.pdf

2.12 Rural population

Health outcomes differ by place of residence; health in general is better in urban than rural settings (*75*). Metropolitan areas, for example, are inhabited by larger populations who enjoy a higher level of economic and social integration (*76*) than those who reside in rural areas. Access to technology, health services, personnel, and infrastructure, among many other factors, mark a difference in health outcomes on the basis of place of residence, including infant mortality (*77-79*). The variable used in the analyses was the percent of the total population that live in rural areas (calculated as the difference between total population and urban population). The data are from the World Bank, World Development Indicators. The data are accessible to the public online at http://databank.worldbank.org/ddp/home.do

2.13 Abortion

Research has shown that legalized abortion has played an important role in the decline in infant mortality rates in the United States (*80, 81*). Its effects were also found to be among the most important at explaining racial infant mortality rates (*81*). Unborn children would have had a higher probability of dying as infants, nevertheless (*82*). The present analyses implemented three abortion measures:

a. Abortion ratio: Number of abortions per 1,000 live births.

b. Abortion percent: Abortions as percentage of pregnancies excluding fetal deaths/miscarriages.

c. Abortion rate: Number of abortions per 1,000 women between 15 and 44 years of age.

The data were compiled by Wm. Robert Johnston from various sources, including the Centers for Disease Control and Prevention and the Government Statistical Service among others, and are available online at http://www.johnstonsarchive.net/policy/abortion/ab-unitedstates.html

2.14 DW-Nominate scores

DW-Nominate scores measure the ideological position of a politician by using their complete voting records. To locate politicians in a liberal-conservative spectrum, scores are computed by using information on the frequency with which politicians voted in conjunction with other politicians throughout their recorded votes (*83, 84*). DW-Nominate scores therefore offer a more punctual assessment of the president's ideology than using the party he belongs to. Indeed, DW-Nominate scores vary within presidents of the same party; for example, the DW-Nominate for Nixon was 0.49 while for Bush Jr. was 0.90 (-1 being extremely liberal and 1 being extremely conservative). Methodological descriptions and the data (updated on February 3, 2011) are available online at http://voteview.com/dwnomin.htm (*85, 86*). Used DW-Nominate scores were for the "first dimension", which refers to the liberal-conservative spectrum in the modern era.

3. Supplementary Figures

Fig S1. Infant Mortality Race Rate Trends by Party (1965–2010)

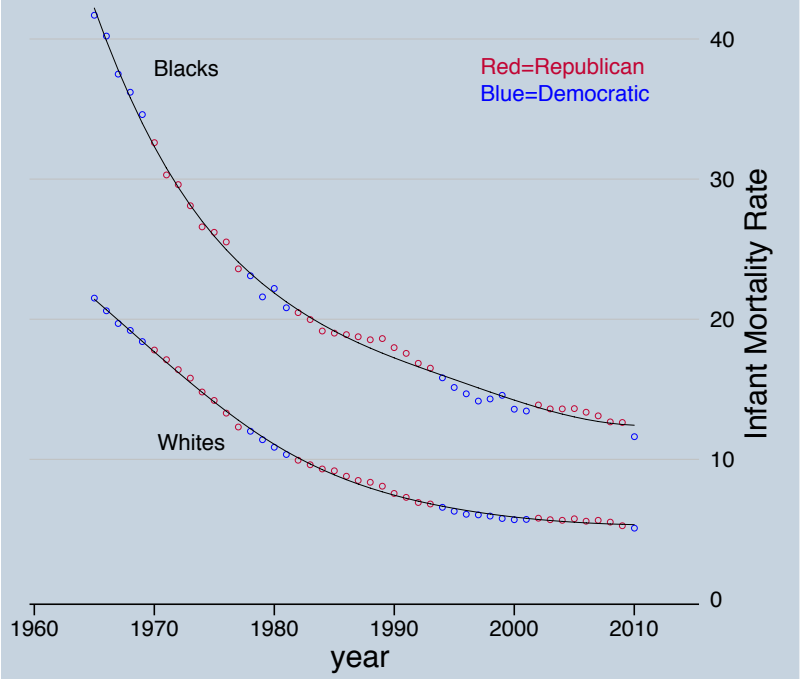
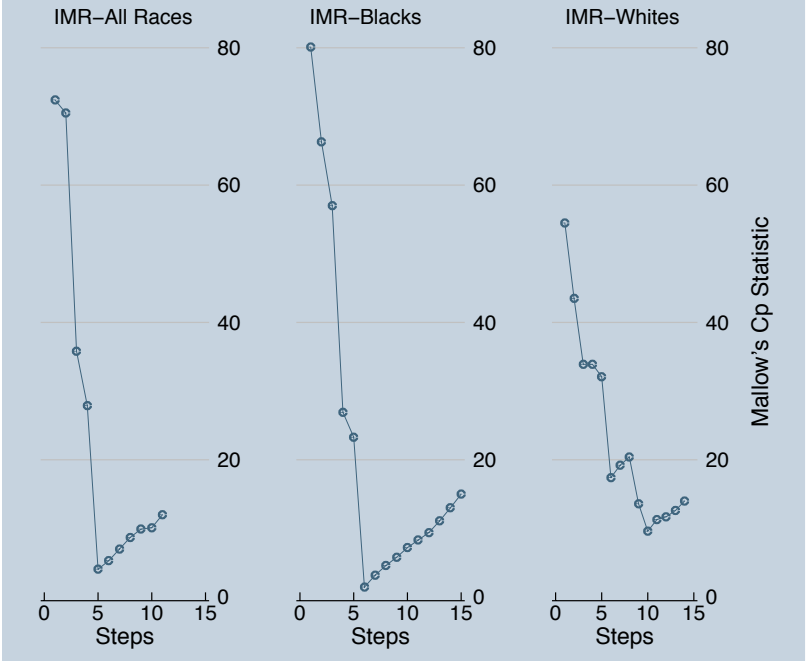


Fig S2. Least Angle Regression Mallow's Cp Estimates



4. Supplementary Tables

Table S1. Parameter estimates of quantile and OLS regressions for IMRs using president's party and DW-Nominate scores (1965-2010)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Quantile	OLS	Quantile	OLS	Quantile	OLS	Quantile	OLS	Quantile	OLS	Quantile	OLS
VARIABLES	IMR	IMR	IMR	IMR	WIMR	WIMR	WIMR	WIMR	BIMR	BIMR	BIMR	BIMR
President's party	.31***	.29***			.30***	.25***			.63***	.50***		
	(.07)	(.05)			(.04)	(.04)			(.14)	(.13)		
DW-Nominate			.23***	.23***			.21***	.20***			.48***	.40***
			(.05)	(.04)			(.05)	(.03)			(.14)	(.10)
Constant	-.19***	-.17***	-.07*	-.05*	-.19***	-.15***	-.06*	-.04*	-.43***	-.31***	-.08	-.08
	(.05)	(.04)	(.03)	(.03)	(.03)	(.02)	(.03)	(.02)	(.11)	(.11)	(.09)	(.07)
Observations	46	46	46	46	46	46	46	46	46	46	46	46
R-squared		.39		.40		.37		.39		.26		.28

Table S1. A DW-Nominate score is a measure of the president's ideological position where -1 stands for extremely liberal and 1 for extremely conservative. Accordingly, average scores for Democratic presidents is -.56 and for Republican presidents is .70 (difference of the means = 1.26 units). By multiplying the score's coefficient by 1.26 it is possible to see that results using scores are practically identical to results using the president's party. Standard errors in parentheses for quantile regressions. Standard errors in parentheses for OLS regresisons were estimated using robust estimation. Statistical significant code: *** p<0.01, ** p<0.05, * p<0.1

Table S2. Delta-beta analysis for the effect of the president's party on racial IMRs (1965-2010)

	Statistic	Observations	Mean	Std. Dev.	Min	Max
All Presidents	Coefficient for BIMR	46	0.65	0.02	0.63	0.69
	Delta-beta for BIMR	46	0.02	0.02	0.00	0.06
	p-values for BIMR	46	0.00	0.00	0.00	0.00
	Coefficient for WIMR	46	0.30	0.00	0.30	0.30
	Delta-beta for WIMR	46	0.00	0.00	0.00	0.01
	p-values WIMR	46	0.00	0.00	0.00	0.00
Democratic Presidents	Coefficient for BIMR	18	0.64	0.03	0.63	0.69
	Delta-beta for BIMR	18	0.02	0.03	0.00	0.06
	p-values for BIMR	18	0.00	0.00	0.00	0.00
	Coefficient for WIMR	18	0.30	0.00	0.30	0.30
	Delta-beta for WIMR	18	0.00	0.00	0.00	0.01
	p-values WIMR	18	0.00	0.00	0.00	0.00
Republican Presidents	Coefficient for BIMR	28	0.65	0.02	0.63	0.69
	Delta-beta for BIMR	28	0.02	0.02	0.00	0.06
	p-values for BIMR	28	0.00	0.00	0.00	0.00
	Coefficient for WIMR	28	0.30	0.00	0.30	0.30
	Delta-beta for WIMR	28	0.00	0.00	0.00	0.01
	p-values WIMR	28	0.00	0.00	0.00	0.00

Table S2. Delta-beta analysis run using quantile regressions at the median. The results are for all possible bivariate models between 1965 and 2010, taking away one president-year at a time. "Observations" refer to the total number of models considered. The dependent variables were BIMR and WIMR. For instance, the first row states that looking at all president years for BIMR, there were 46 total models from which coefficients were estimated with an average of .65, SD=.02, and with minimum and maximum values of .63 and .69, respectively. "Delta-beta" refers to the absolute value of the diference between the estimated coefficient (without one president-year) and the coefficient estimated using all president-years. The "mean" column reports the average absolute deviation between the coefficients. For example, the second row shows reports that the average deviation of the coefficients for BIMR was .02, with a SD of .02, a minimum value of 0 (meaning no deviation) and a maximum deviation of .06. These results show that the relationship between the president's party and the BIMR is not significantly influenced by any specific president-year. "P-values" refers to the variation of the coefficient's statistical significance. For example, the third row shows no variation of the statistical significance level throughout the analysis, meaning that we can be sure that the statistical significance of the relationship between the president's party and BIMR is not an artifact of the influence of any given president-year. Results are also separated in Democratic-president-years and Republican-president-years. For example, the first row under "Democratic Presidents" shows that the effect of the president's party over BIMR was not significantly influenced by any given Democratic-president-year. Results were practically identical using ordinary linear regression.

Table S3a. Pairwise correlation matrix for variables included in sensitivity analysis using SEM for IMR (all races)

	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*	*19*	*20*	*21*	*22*	*23*	*24*	*25*	*26*	*27*	*28*	*29*	*30*	*31*
1	1.00																														
2	0.62	1.00																													
3	-0.01	0.14	1.00																												
4	0.54	0.29	-0.24	1.00																											
5	0.64	0.47	-0.14	0.64	1.00																										
6	0.07	0.25	-0.26	0.15	0.21	1.00																									
7	-0.18	0.03	-0.26	-0.05	0.06	0.82	1.00																								
8	0.45	0.20	0.20	0.36	0.41	-0.35	-0.48	1.00																							
9	-0.25	0.19	0.15	-0.11	0.06	0.52	0.39	-0.40	1.00																						
10	-0.25	0.17	0.15	-0.03	0.08	0.47	0.34	-0.32	0.98	1.00																					
11	-0.31	0.16	0.18	-0.14	0.02	0.41	0.34	-0.37	0.98	0.97	1.00																				
12	0.35	-0.01	-0.26	0.08	0.19	-0.13	0.03	0.04	-0.73	-0.80	-0.73	1.00																			
13	-0.29	0.18	0.17	-0.25	-0.09	0.28	0.26	-0.30	0.78	0.75	0.80	-0.55	1.00																		
14	0.27	0.36	0.15	0.12	0.04	0.11	-0.10	0.20	0.19	0.18	0.17	-0.18	0.16	1.00																	
15	-0.38	-0.25	-0.05	-0.35	-0.41	0.06	0.28	-0.43	0.20	0.11	0.23	-0.04	0.26	-0.30	1.00																
16	-0.08	-0.24	-0.16	-0.29	-0.22	-0.19	-0.04	-0.27	-0.30	-0.33	-0.30	0.30	-0.17	-0.42	0.22	1.00															
17	-0.09	-0.19	-0.19	0.22	0.04	-0.06	-0.05	0.35	0.00	0.05	0.03	-0.28	0.14	0.02	0.04	0.17	1.00														
18	-0.07	-0.17	-0.22	0.27	0.11	-0.13	-0.05	0.34	-0.09	-0.01	-0.07	-0.19	0.05	0.01	-0.05	0.21	0.82	1.00													
19	0.09	0.00	-0.18	0.01	-0.07	0.03	-0.01	-0.19	-0.05	-0.04	-0.09	-0.04	-0.01	-0.05	-0.09	0.66	0.30	0.54	1.00												
20	0.14	0.24	0.36	0.08	0.17	0.04	-0.07	-0.01	0.43	0.50	0.44	-0.52	0.03	0.24	-0.15	-0.13	-0.19	-0.03	0.18	1.00											
21	0.18	0.04	0.40	0.03	0.42	-0.36	-0.35	0.69	-0.11	-0.05	-0.07	-0.14	-0.06	0.10	-0.42	-0.24	0.20	0.17	-0.24	0.15	1.00										
22	-0.08	-0.11	-0.32	0.10	0.07	0.63	0.84	-0.26	-0.03	-0.05	-0.08	0.27	-0.13	-0.13	0.15	-0.03	0.02	0.05	-0.01	-0.21	-0.30	1.00									
23	-0.20	-0.40	0.21	-0.12	-0.28	-0.90	-0.60	0.36	-0.53	-0.47	-0.41	0.16	-0.26	-0.21	0.09	0.15	0.22	0.26	-0.11	-0.19	0.36	-0.35	1.00								
24	0.04	-0.39	-0.02	-0.08	-0.18	-0.63	-0.39	0.25	-0.92	-0.91	-0.88	0.73	-0.71	-0.26	0.00	0.35	-0.09	0.00	-0.04	-0.44	0.12	0.01	0.68	1.00							
25	0.28	0.10	0.26	-0.21	-0.04	-0.13	-0.28	0.10	-0.05	-0.05	-0.11	-0.07	-0.09	0.18	-0.23	0.16	-0.26	-0.07	0.37	0.42	0.26	-0.27	-0.03	0.04	1.00						
26	-0.19	-0.11	0.06	0.17	0.03	-0.07	0.05	0.21	-0.02	0.06	0.05	-0.08	0.03	-0.21	0.00	-0.39	0.01	0.05	-0.48	-0.20	0.17	0.14	0.25	0.02	-0.37	1.00					
27	-0.03	0.03	0.05	-0.09	0.08	-0.18	-0.08	0.35	-0.10	-0.13	-0.01	0.16	0.08	0.03	0.10	-0.20	0.16	-0.12	-0.63	-0.47	0.20	-0.01	0.30	0.18	-0.53	0.41	1.00				
28	-0.21	-0.08	0.14	-0.05	-0.28	-0.05	0.02	0.00	-0.04	-0.08	0.04	0.08	0.11	0.07	0.28	-0.10	0.12	-0.25	-0.54	-0.44	-0.14	0.04	0.18	0.12	-0.62	0.25	0.71	1.00			
29	-0.07	0.03	-0.01	-0.07	0.06	-0.05	0.08	0.19	-0.02	-0.05	0.08	0.18	0.15	-0.03	0.28	-0.17	0.13	-0.11	-0.62	-0.48	0.00	0.08	0.22	0.09	-0.59	0.49	0.93	0.75	1.00		
30	-0.02	0.21	-0.25	0.21	0.26	0.38	0.30	-0.04	0.40	0.41	0.39	-0.19	0.40	0.41	0.10	-0.23	0.17	0.31	-0.06	0.10	0.01	0.24	-0.33	-0.39	-0.03	0.08	-0.10	0.11	0.08	1.00	
31	0.05	-0.15	-0.23	-0.14	-0.20	0.07	0.08	-0.38	-0.38	-0.45	-0.45	0.50	-0.47	0.02	-0.11	0.44	-0.35	-0.10	0.47	-0.06	-0.40	0.25	-0.19	0.39	0.36	-0.31	-0.50	-0.41	-0.50	-0.25	1.00

Table S3a. Variables included are:

1=IMR **2**=President's party **3**=Recessions **4**=LBW **5**= PB **6**=Inc. share ratio top 5% / bottom 20%
7=Family Income Gini **8**=% Rural pop. **9**=% Poverty **10**=% Poverty females **11**=% Under 1.25 poverty threshold **12**=Mean income
13=Unemployment rate **14**=Income ratio women/men **15**=% High school degree **16**=Total Mat/Child Expend. Per cap **17**=Fed.Mat/Ch exp. Per cap.
18=Federal Mat/Ch exp.%GDP **19**=Total Mat/Ch exp. %GDP **20**=CPI **21**=Change in CPI of Medical Care **22**=Income share of top 5%
23=Income share bottom 20% **24**=Mean income of bottom 20% **25**=Alcohol **26**=Tobacco **27**=Abortion rate **28**=Abortion ratio
29=Abortion percent **30**=% Uninsured **31**=Median age of mothers without husband at 1st birth

Table S3b. Pairwise correlation matrix for variables included in sensitivity analysis using SEM for WIMR

	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*	*19*	*20*	*21*	*22*	*23*	*24*	*25*	*26*	*27*	*28*	*29*	*30*	*31*
1	1.00																														
2	0.61	1.00																													
3	-0.05	0.14	1.00																												
4	0.42	0.29	-0.24	1.00																											
5	0.58	0.47	-0.14	0.64	1.00																										
6	0.18	0.25	-0.26	0.15	0.21	1.00																									
7	-0.08	-0.05	-0.25	-0.08	0.01	0.77	1.00																								
8	0.20	0.20	0.20	0.36	0.41	-0.35	-0.53	1.00																							
9	-0.18	0.14	0.16	-0.18	-0.06	0.45	0.40	-0.50	1.00																						
10	0.11	0.24	0.09	0.18	0.23	0.42	0.25	-0.31	0.76	1.00																					
11	0.33	0.00	-0.28	0.09	0.21	-0.12	0.06	0.03	-0.68	-0.41	1.00																				
12	-0.27	0.18	0.19	-0.25	-0.09	0.25	0.23	-0.29	0.80	0.54	-0.53	1.00																			
13	0.24	0.36	0.15	0.12	0.04	0.11	-0.10	0.20	0.16	0.29	-0.18	0.15	1.00																		
14	-0.42	-0.26	-0.04	-0.31	-0.38	-0.03	0.15	-0.32	0.22	0.01	-0.12	0.24	-0.28	1.00																	
15	-0.17	-0.13	0.02	-0.27	-0.24	-0.13	-0.02	-0.25	0.22	0.09	-0.14	0.12	-0.21	0.86	1.00																
16	-0.09	-0.24	-0.16	-0.29	-0.22	-0.19	-0.02	-0.27	-0.24	-0.41	0.30	-0.15	-0.42	0.15	0.11	1.00															
17	-0.30	-0.19	-0.19	0.22	0.04	-0.06	-0.08	0.35	-0.07	-0.15	-0.28	0.14	0.02	0.13	0.03	0.17	1.00														
18	-0.24	-0.17	-0.22	0.27	0.11	-0.13	-0.06	0.34	-0.10	-0.18	-0.18	0.05	0.01	0.01	-0.01	0.21	0.82	1.00													
19	0.07	0.00	-0.18	0.01	-0.07	0.03	0.00	-0.19	-0.01	-0.13	-0.03	-0.01	-0.05	-0.15	-0.04	0.66	0.30	0.54	1.00												
20	0.22	0.24	0.36	0.08	0.17	0.04	-0.10	-0.01	0.41	0.30	-0.53	0.03	0.24	-0.10	0.14	-0.13	-0.19	-0.03	0.18	1.00											
21	0.01	0.04	0.40	0.03	0.42	-0.36	-0.39	0.69	-0.19	-0.17	-0.14	-0.04	0.10	-0.28	-0.21	-0.24	0.20	0.17	-0.24	0.15	1.00										
22	0.05	-0.08	-0.11	0.16	0.07	0.40	0.61	-0.14	0.06	0.06	0.20	0.12	-0.09	0.09	-0.07	0.02	0.09	0.13	-0.05	-0.26	-0.33	1.00									
23	-0.37	-0.36	0.19	-0.04	-0.30	-0.82	-0.67	0.44	-0.51	-0.40	0.07	-0.31	-0.19	0.14	0.10	0.08	0.18	0.18	-0.18	-0.19	0.37	-0.31	1.00								
24	0.01	-0.32	-0.03	0.01	-0.08	-0.62	-0.42	0.36	-0.91	-0.74	0.71	-0.74	-0.25	-0.01	-0.02	0.31	-0.04	0.02	-0.07	-0.42	0.19	-0.08	0.66	1.00							
25	0.36	0.10	0.26	-0.21	-0.04	-0.13	-0.24	0.10	-0.01	-0.09	-0.08	-0.09	0.18	-0.29	-0.11	0.16	-0.26	-0.07	0.37	0.42	0.26	-0.21	-0.06	0.02	1.00						
26	-0.19	-0.11	0.06	0.17	0.03	-0.07	0.07	0.21	-0.01	0.17	-0.07	0.02	-0.21	-0.02	-0.11	-0.39	0.01	0.05	-0.48	-0.20	0.17	0.22	0.32	0.00	-0.37	1.00					
27	-0.15	0.03	0.05	-0.09	0.08	-0.18	-0.12	0.35	-0.17	-0.02	0.16	0.10	0.03	0.17	0.02	-0.20	0.16	-0.12	-0.63	-0.47	0.20	0.13	0.33	0.23	-0.53	0.41	1.00				
28	-0.31	-0.08	0.14	-0.05	-0.28	-0.05	0.01	0.00	-0.08	0.02	0.08	0.13	0.07	0.30	0.07	-0.10	0.12	-0.25	-0.54	-0.44	-0.14	0.20	0.25	0.13	-0.62	0.25	0.71	1.00			
29	-0.14	0.03	-0.01	-0.07	0.06	-0.05	0.05	0.19	-0.05	0.11	0.18	0.17	-0.03	0.30	0.13	-0.17	0.13	-0.11	-0.62	-0.48	0.00	0.25	0.22	0.11	-0.59	0.49	0.93	0.75	1.00		
30	-0.06	0.21	-0.25	0.21	0.26	0.38	0.33	-0.04	0.34	0.42	-0.17	0.39	0.41	0.04	-0.02	-0.23	0.17	0.31	-0.06	0.10	0.01	0.26	-0.35	-0.44	-0.03	0.08	-0.10	0.11	0.08	1.00	
31	-0.02	-0.03	-0.44	0.01	-0.20	0.44	0.30	-0.51	0.12	0.11	0.01	0.02	-0.07	0.15	0.13	0.32	0.14	0.16	0.45	-0.08	-0.62	0.03	-0.41	-0.21	-0.24	-0.20	-0.34	-0.12	-0.20	-0.03	1.00

Table S3b. Variables included are:

1=WIMR **2**=President's party **3**=Recessions **4**=LBW **5**= PB **6**=Inc. share ratio top 5% / bottom 20%
7=White Fam. Inc. Gini **8**=% Rural pop. **9**=% White poverty **10**=%White Pov. Fem. Without husb. **11**=White mean income
12=White unemployment **13**=Inc. ratio women/men **14**=%White High School **15**=%White Fem. Hgh.Sch. **16**=Total Mat/Child Expend. Per cap
17=Fed.Mat/Ch exp. Per capita **18**=Fed.Mat/Ch exp.%GDP **19**=Total Mat/Ch exp. %GDP **20**=CPI **21**=Change in CPI of Medical Care
22=White inc. share top 5% **23**=White inc. share bottom 20% **24**=White mean inc. of bottom 20% **25**=Alcohol **26**=Tobacco
27=Abortion rate **28**=Abortion ratio **29**=Abortion percent **30**=% Uninsured **31**=Median age of white mothers without husband at 1st birth

Table S3c. Pairwise correlation matrix for variables included in sensitivity analysis using SEM for BIMR

	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*	*19*	*20*	*21*	*22*	*23*	*24*	*25*	*26*	*27*	*28*	*29*	*30*	*31*
1	1.00																														
2	0.51	1.00																													
3	0.11	0.14	1.00																												
4	0.66	0.29	-0.24	1.00																											
5	0.56	0.47	-0.14	0.64	1.00																										
6	-0.06	0.25	-0.26	0.15	0.21	1.00																									
7	0.31	0.43	-0.24	0.30	0.44	0.59	1.00																								
8	0.58	0.20	0.20	0.36	0.41	-0.35	0.10	1.00																							
9	0.08	0.37	0.06	0.18	0.38	0.41	0.51	0.07	1.00																						
10	0.21	0.40	0.06	0.29	0.41	0.33	0.50	-0.02	0.85	1.00																					
11	0.07	-0.07	-0.28	-0.02	0.13	-0.14	-0.02	0.01	-0.58	-0.50	1.00																				
12	-0.19	0.17	0.05	-0.18	-0.04	0.37	0.07	-0.34	0.58	0.43	-0.56	1.00																			
13	0.24	0.36	0.15	0.12	0.04	0.11	0.01	0.20	0.16	0.02	-0.23	0.13	1.00																		
14	-0.43	-0.29	-0.04	-0.32	-0.37	0.17	-0.16	-0.60	-0.27	-0.20	0.16	0.07	-0.32	1.00																	
15	-0.33	-0.09	-0.14	-0.17	-0.27	0.26	-0.07	-0.51	-0.15	-0.14	0.06	0.08	-0.21	0.83	1.00																
16	-0.07	-0.24	-0.16	-0.29	-0.22	-0.19	-0.13	-0.27	-0.50	-0.29	0.34	-0.21	-0.42	0.23	0.16	1.00															
17	0.16	-0.19	-0.19	0.22	0.04	-0.06	0.13	0.35	0.13	0.13	-0.23	0.14	0.02	-0.28	-0.32	0.17	1.00														
18	0.22	-0.17	-0.22	0.27	0.11	-0.13	0.04	0.34	-0.03	-0.01	-0.20	0.09	0.01	-0.19	-0.23	0.21	0.82	1.00													
19	0.24	0.00	-0.18	0.01	-0.07	0.03	0.06	-0.19	-0.32	-0.11	-0.03	0.00	-0.05	0.06	-0.01	0.66	0.30	0.54	1.00												
20	0.23	0.24	0.36	0.08	0.17	0.04	0.15	-0.01	0.38	0.43	-0.48	0.06	0.24	-0.17	-0.06	-0.13	-0.19	-0.03	0.18	1.00											
21	0.33	0.04	0.40	0.03	0.42	-0.36	0.05	0.69	0.16	0.09	-0.11	-0.13	0.10	-0.45	-0.54	-0.24	0.20	0.17	-0.24	0.15	1.00										
22	0.14	0.32	-0.26	0.16	0.22	0.48	0.88	0.00	0.30	0.29	0.11	0.01	-0.05	0.05	0.08	-0.10	0.09	0.06	0.04	0.07	-0.04	1.00									
23	-0.37	-0.49	0.12	-0.33	-0.47	-0.54	-0.73	-0.03	-0.43	-0.48	0.09	-0.04	-0.19	0.19	0.14	0.16	0.07	0.11	-0.18	-0.27	0.02	-0.47	1.00								
24	-0.24	-0.48	-0.02	-0.30	-0.39	-0.55	-0.57	-0.03	-0.75	-0.72	0.54	-0.42	-0.28	0.28	0.19	0.44	-0.02	0.09	0.03	-0.39	-0.05	-0.27	0.82	1.00							
25	0.24	0.10	0.26	-0.21	-0.04	-0.13	-0.14	0.10	-0.25	-0.28	-0.03	-0.11	0.18	-0.02	-0.07	0.16	-0.26	-0.07	0.37	0.42	0.26	-0.19	-0.15	-0.04	1.00						
26	-0.17	-0.11	0.06	0.17	0.03	-0.07	-0.11	0.21	0.14	0.05	-0.16	0.06	-0.21	0.03	0.13	-0.39	0.01	0.05	-0.48	-0.20	0.17	0.00	0.34	0.11	-0.37	1.00					
27	-0.17	0.03	0.05	-0.09	0.08	-0.18	-0.05	0.35	0.19	0.03	0.13	0.00	0.03	-0.25	-0.16	-0.20	0.16	-0.12	-0.63	-0.47	0.20	-0.03	0.34	0.20	-0.53	0.41	1.00				
28	-0.21	-0.08	0.14	-0.05	-0.28	-0.05	-0.08	0.00	0.09	0.10	0.06	0.02	0.07	0.00	0.07	-0.10	0.12	-0.25	-0.54	-0.44	-0.14	-0.05	0.30	0.19	-0.62	0.25	0.71	1.00			
29	-0.25	0.03	-0.01	-0.07	0.06	-0.05	-0.04	0.19	0.20	0.10	0.14	0.09	-0.03	-0.04	0.07	-0.17	0.13	-0.11	-0.62	-0.48	0.00	0.02	0.35	0.18	-0.59	0.49	0.93	0.75	1.00		
30	0.07	0.21	-0.25	0.21	0.26	0.38	0.13	-0.04	0.38	0.30	-0.16	0.43	0.41	-0.07	0.10	-0.23	0.17	0.31	-0.06	0.10	0.01	0.07	0.01	-0.13	-0.03	0.08	-0.10	0.11	0.08	1.00	
31	-0.01	-0.28	-0.18	-0.03	-0.30	-0.03	-0.12	-0.32	-0.29	-0.26	-0.04	-0.16	-0.14	0.17	0.10	0.25	-0.07	0.16	0.51	0.26	-0.23	-0.03	0.03	0.16	0.30	-0.17	-0.61	-0.44	-0.66	-0.06	1.00

Table S3c. Variables included are:

1=BIMR **2**=President's party **3**=Recessions **4**=LBW **5**= PB **6**=Inc. share ratio top 5% / bottom 20%
7=Black Fam. Inc. Gini **8**=% Rural pop. **9**=% Black poverty **10**=%Black Pov. Fem. Without husb. **11**=Black mean income
12=Black unemployment **13**=Inc. ratio women/men **14**=%Black High School **15**=%Black Fem. Hgh.Sch **16**=Total Mat/Child Expend. Per cap
17=Fed.Mat/Ch exp. Per capita **18**=Fed.Mat/Ch exp.%GDP **19**=Total Mat/Ch exp. %GDP **20**=CPI **21**=Change in CPI of Medical Care
22=Black inc. share top 5% **23**=Black inc. share bottom 20% **24**=Black mean inc. of bottom 20% **25**=Alcohol **26**=Tobacco
27=Abortion rate **28**=Abortion ratio **29**=Abortion percent **30**=% Uninsured **31**=Median age of black mothers without husband at 1st birth

Table S4a. Least Angle Regression output for IMR (all races)

Step	Mallows' Cp	R-squared	Action
1	72.4	.00	
2	70.5	.04	*plus* president's party
3	35.8	.37	*plus* preterm birth rate
4	35.8	.39	*plus* percent rural population
5	27.9	.48	*plus* low birthweight rate
6	4.0	.72	*plus* alcohol consumption
7	5.3	.72	*plus* income share top 5%
8	7.0	.73	*plus* abortion ratio
9	8.7	.73	*plus* percent with high school
10	9.9	.74	*plus* federal maternal/child exp. pc
11	10.1	.75	*plus* income ratio women/men
12	12.0	.75	*plus* income share bottom 20%

Coefficient values for final model (minimum Mallow's Cp) in order of importance

Variable	Unstandardized	Standardized
Pres. party	.17	.49
PB	.86	.28
Rural pop.	.22	.22
LBW	.40	.37
Alcohol	1.91	.37

Table S4b. Least Angle Regression output for IMR whites

Step	Mallows' Cp	R-squared	Action
1	54.5	.00	
2	43.5	.14	*plus* president's party
3	33.9	.27	*plus* preterm birth rate
4	33.9	.29	*plus* percent white with high school
5	32.1	.34	*plus* abortion ratio
6	17.4	.52	*plus* alcohol consumption
7	19.2	.52	*plus* white income share bottom 20%
8	20.4	.53	*plus* federal maternal/child exp. pc
9	13.6	.63	*plus* low birthweight rate
10	9.6	.70	*plus* white income share top 5%
11	11.3	.70	*plus* percent rural population
12	11.7	.72	*plus* income ratio women/men
13	12.6	.73	*plus* tobacco consumption
14	14.0	.73	*plus* total maternal/child exp. Pc

Coefficient values for final model (minimum Mallow's Cp) in order of importance

Variable	Unstandardized	Standardized
Pres. party	.15	.43
PB	.17	.22
% High school	-.08	-.12
Abortion ratio	-.00	-.10
Alcohol	1.61	.31
Inc. shr. bt.20%	-.10	-.06
Fed. mat/ch pc	-.24	-.17
LBW	.53	.22
Inc. shr. top5%	.03	.11

Table S4c. Least Angle Regression output for IMR blacks

Step	Mallows' Cp	R-squared	Action
1	80.1	.00	
2	66.3	.13	*plus* low birthweight rate
3	57.0	.23	*plus* percent rural population
4	26.9	.51	*plus* president's party
5	23.3	.55	*plus* abortion percent
6	1.4	.76	*plus* alcohol consumption
7	3.2	.76	*plus* black income share top 5%
8	4.6	.76	*plus* recessions
9	5.8	.77	*plus* federal maternal/child exp. %GDP
10	7.2	.78	*plus* preterm birth rate
11	8.3	.78	*plus* Consumer price index-URS
12	9.4	.79	*plus* income ratio women/men
13	11.1	.79	*plus* black income share bottom 20%
14	13.0	.80	*plus* black females without husband poverty
15	15.0	.80	*plus* percent black with high school

Coefficient values for final model (minimum Mallow's Cp) in order of importance

Variable	Unstandardized	Standardized
LBW	2.88	1.23
Rural pop.	2.06	1.14
Pres. party	.25	.74
Abortion percent	-.07	-.69
Alcohol	1.33	.26

Table S5a. Seemingly Unrelated Regression parameter estimates for IMRs (1969-2007) using LAR's output and simplified model

VARIABLES	IMR (all Races)	IMR Blacks	IMR Whites	IMR (all Races)	IMR Blacks	IMR Whites
	Models specified through LAR			Simplified model [Model 6 in Table 3]		
Pres. party (Rep=1, Dem=0)	0.18	0.28	0.16	0.18	0.29	0.17
	(0.05)	(0.09)	(0.05)	(0.05)	(0.10)	(0.05)
LBW	1.07	3.26	0.74	1.09	3.14	0.85
	(0.40)	(0.75)	(0.41)	(0.46)	(0.90)	(0.47)
PB	0.21		0.23	0.22	0.01	0.28
	(0.13)		(0.14)	(0.15)	(0.29)	(0.15)
Alcohol Consumption	2.49	2.29	2.28	2.25	1.77	3.17
	(0.72)	(1.82)	(0.75)	(1.02)	(1.98)	(1.03)
Percent Rural Population	0.55	2.30		0.48	2.37	-0.32
	(0.16)	(0.55)		(0.31)	(0.60)	(0.31)
Abortion Percent		-0.07		0.00	-0.09	0.01
		(0.03)		(0.02)	(0.04)	(0.02)
Abortion Ratio			-0.00			
			(0.00)			
Federal Mat/Ch Exp. pc			-0.21			
			(0.09)			
Inc. Share Bottom 20%			-0.01			
			(0.13)			
Inc. Share Top 5%			0.01			
			(0.02)			
Pct. Whites High School			-0.03			
			(0.04)			
Constant	-0.11	-0.18	-0.10	-0.11	-0.18	-0.11
	(0.04)	(0.07)	(0.04)	(0.04)	(0.08)	(0.04)
Observations	39	39	39	39	39	39
RMSE of Regression	0.13	0.25	0.13	0.13	0.26	0.13
F-Statistic of Regression	23.2	26.4	11.0	14.5	17.9	10.0
R-squared	0.73	0.77	0.69	0.73	0.77	0.65

(1) Standard errors (in parentheses) adjusted for small sample size. All variables detrended (except "President's Party")
(2) Variables specification:
- Preterm births refers to births that happened before the 37th week of gestation.
- Low birthweight refers to newborns weighting less than 2500 grs.
- Abortion Rate is number of abortions per 1000 women ages 15 to 44 years.
- Abortion Ratio is number of abortions per 1000 live births.
- Abortion Percent is the number of abortions as percentage of pregnancies (excluding fetal deaths and miscarriages).
- Per capita federal health expenditure on Maternal and Child Health care refers to the sum of the following components of federal expenditures on maternal and child health care: (a) Personal health care, (b) Hospital (a) Personal health care, (b) Hospital expenditures, (c) Physician and clinical expenditures, (d) Prescription drug expenditures, and (e) Durable medical equipment.
- Per capita gallons of ethanol consumption on population age 15 and older prior to 1970, and on population age 14 and older thereafter.

Table S5b. Correlation matrix of cross-equation disturbances of LAR models

		Residuals for		
		IMR	BIMR	WIMR
Residuals for	IMR	1.00		
	BIMR	0.46	1.00	
	WIMR	0.91	0.29	1.00

Breusch-Pagan test of independence: chi2(3) = 43.99, Pr = 0.0000

Table S5c. Correlation matrix of cross-equation disturbances of simplified models [Model 6 in Table 3]

		Residuals for		
		IMR	BIMR	WIMR
Residuals for	IMR	1.00		
	BIMR	0.45	1.00	
	WIMR	0.92	0.19	1.00

Breusch-Pagan test of independence: chi2(3) = 42.56, Pr = 0.0000

Note: The Breusch-Pagan test of independence tests for conditional heteroscedasticity by regressing the squared residuals on the independent variables (*87*). Results show that the cross-equation disturbances are not independent.

Table S6. Covariate contribution to fitting of SUR models

	Change in R-sqared		
Without	IMR	WIMR	BIMR
Pres. Party	-0.102	-0.114	-0.057
LBW	-0.047	-0.037	-0.088
PB	-0.066	-0.060	-0.025
Alcohol	-0.054	-0.108	-0.006
Rural Pop.	-0.021	-0.011	-0.113
Abortion	0.000	-0.004	-0.040

Table S6. The cells represent the change in the R-squared statistic due to excluding a given variable at a time. The negative sign shows reduction of the R-squared statistic. Findings corroborate the results from LAR, where the president's party was the most important covariate at predicting IMR and WIMR, and the third for BIMR.

5. Supporting Material References

1. C. De Boor, *A practical guide to splines*. Applied mathematical sciences (Springer-Verlag, New York, 1978), pp. xxiv, 392 p.

2. Z. Ziegler, One–sided L1–approximation by splines of an arbitrary degree. *Schoenberg IJ (ed.)*, (1969).

3. R. B. Newson, B-splines and Splines Parameterized by Their Values at Reference Points on the X-axis. *Stata Technical Bulletin* **10**, (2001).

4. I. J. Schoenberg, *Approximations, with Special Emphasis on Spline Functions: Proceedings of a Symposium Conducted by the Mathematics Research Center, United States Army, at the University of Wisconsin, Madison, May 5-7, 1969*. (Academic Press, 1969).

5. S. C. . *Stata Base Reference Manual: Release 11*. (Stata Press, 2009).

6. F. Wilcoxon, Individual comparisons by ranking methods. *Biometrics Bulletin* **1**, 80 (1945).

7. H. B. Mann, D. R. Whitney, On a test of whether one of two random variables is stochastically larger than the other. *The annals of mathematical statistics* **18**, 50 (1947).

8. W. H. Kruskal, W. A. Wallis, Use of ranks in one-criterion variance analysis. *Journal of the American statistical Association*, 583 (1952).

9. F. E. Satterthwaite, An approximate distribution of estimates of variance components. *Biometrics Bulletin* **2**, 110 (1946).

10. A. C. Cameron, P. K. Trivedi, *Microeconometrics using stata*. (Stata Press Texas, 2009), vol. 5.

11. P. J. Rousseeuw, A. M. Leroy, J. Wiley, *Robust regression and outlier detection*. (Wiley Online Library, 1987), vol. 3.

12. K. J. Rothman, S. Greenland, T. L. Lash, *Modern epidemiology*. (Lippincott Williams & Wilkins, 2008).

13. R. Tibshirani, Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 267 (1996).

14. B. Efron, T. Hastie, I. Johnstone, R. Tibshirani, Least angle regression. *The Annals of statistics* **32**, 407 (2004).

15. T. Hesterberg, N. H. Choi, L. Meier, C. Fraley, Least angle and ℓ1 penalized regression: A review. *Statistics Surveys* **2**, 61 (2008).

16. C. L. Mallows, Some comments on Cp. *Technometrics*, 661 (1973).

17. A. Zellner, An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. *Journal of the American statistical Association*, 348 (1962).

18. D. H. Felmlee, L. L. Hargens, Estimation and hypothesis testing for seemingly unrelated regressions: A sociological application. *Social Science Research* **17**, 384 (1988).

19. J. Kmenta, R. F. Gilbert, Small sample properties of alternative estimators of seemingly unrelated regressions. *Journal of the American statistical Association*, 1180 (1968).

20. A. Zellner, Estimators for seemingly unrelated regression equations: Some exact finite sample results. *Journal of the American statistical Association*, 977 (1963).

21. A. Zellner, D. S. Huang, Further properties of efficient estimators for seemingly unrelated regression equations. *International Economic Review* **3**, 300 (1962).

22. M. W. L. Cheung, Comparison of approaches to constructing confidence intervals for mediating effects using structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal* **14**, 227 (2007).

23. R. H. Hoyle, D. A. Kenny, Sample size, reliability, and tests of statistical mediation. *Statistical strategies for small sample research*, 195 (1999).

24. E. E. Leamer, Let's take the con out of econometrics. *The American Economic Review* **73**, 31 (1983).

25. R. Levine, D. Renelt, A sensitivity analysis of cross-country growth regressions. *The American Economic Review*, 942 (1992).

26. P. Mair, E. Wu, P. M. Bentler, An R Interface to EQS: The REQS Package. (2008).

27. R. G. Wilkinson, K. E. Pickett, Income inequality and population health: a review and explanation of the evidence. *Social science & medicine* **62**, 1768 (2006).

28. S. L. Murphy, J. Xu, K. D. Kochanek, Deaths: Preliminary Data for 2010. *National Vital Statistics Reports* **60**, (2012).

29. J. Xu, M. Kenneth, D. Kochanek, M. Sherry, L. Murphy, Deaths: final data for 2007. (2010).

30. N. C. f. H. Statistics, Vital Statistics of the United States, 1990, Volume II, Mortality, Part A. Washington: Public Health Service., (1994).

31. K. D. X. Kochanek, J. Q.; Murphy, S.L.; Miniño, A.M.; Kung, H.C., Deaths: Preliminary data for 2009. National Vital Statistics Reports; Vol 59, No 4. Hyattsville, MD: National Center for Health Statistics. (2011).

32. M. S. Kramer, Determinants of low birth weight: methodological assessment and meta-analysis. *Bulletin of the World Health Organization* **65**, 663 (1987).

33. M. C. McCormick, The contribution of low birth weight to infant mortality and childhood morbidity. *New England journal of medicine* **312**, 82 (1985).

34. J. W. Collins, R. J. David, Racial disparity in low birth weight and infant mortality. *Clinics in perinatology* **36**, 63 (2009).

35. J. A. Martin *et al.*, Births: final data for 2009. *National vital statistics reports: from the Centers for Disease Control and Prevention, National Center for Health Statistics, National Vital Statistics System* **60**, 1 (2011).

36. S. Taffel, N. C. f. H. Statistics, *Factors associated with low birth weight, United States, 1976*. (US Department of Health, Education, and Welfare, Public Health Service, Office of Health Research, Statistics, and Technology, National Center for Health Statistics, 1980).

37. N. C. f. H. Statistics, Monthly Vital Statistics Report. Summary Report. Final Natality Statistics. 1970. U.S. Department of Health, Education, and Welfare, Public Health Services, Health Resources Administration. Vol 22, No 12. (1970).

38. M. H. Brenner, Fetal, infant, and maternal mortality during periods of economic instability. *International Journal of Health Services* **3**, 145 (1973).

39. E. Neumayer, Recessions lower (some) mortality rates::: evidence from Germany. *Social science & medicine* **58**, 1037 (2004).

40. A. S. Deaton, C. Paxson. (University of Chicago Press, 2004).

41. M. H. Brenner, Political economy and health. *Society and health*, 211 (1995).

42. R. J. Waldmann, Income distribution and infant mortality. *the Quarterly Journal of economics* **107**, 1283 (1992).

43. J. W. Lynch *et al.*, Income inequality and mortality in metropolitan areas of the United States. *American Journal of Public Health* **88**, 1074 (1998).

44. I. Wennemo, Infant mortality, public policy and inequality–a comparison of 18 industrialised countries 1950–85. *Sociology of Health & Illness* **15**, 429 (1993).

45. K. Judge, J. A. Mulligan, M. Benzeval, Income inequality and population health. *Social science & medicine* **46**, 567 (1998).

46. S. L. Gortmaker, Poverty and infant mortality in the United States. *American Sociological Review*, 280 (1979).

47. A. P. Polednak, Black-white differences in infant mortality in 38 standard metropolitan statistical areas. *American Journal of Public Health* **81**, 1480 (1991).

48. A. P. Polednak, Trends in US urban black infant mortality, by degree of residential segregation. *American Journal of Public Health* **86**, 723 (1996).

49. D. Black, *Inequalities in health: the Black report*. (Penguin Books, 1982).

50. T. Mathews, M. F. MacDorman, Infant mortality statistics from the 2006 period linked birth/infant death data set. *National vital statistics reports: from the Centers for Disease Control and Prevention, National Center for Health Statistics, National Vital Statistics System* **58**, 1 (2010).

51. J. Rogowski, Cost-effectiveness of care for very low birth weight infants. *Pediatrics* **102**, 35 (1998).

52. L. Shi, B. Starfield, The effect of primary care physician supply and income inequality on mortality among blacks and whites in US metropolitan areas. *Journal Information* **91**, (2001).

53. B. Lia-Hoagberg *et al.*, Barriers and motivators to prenatal care among low-income women. *Social science & medicine* **30**, 487 (1990).

54. L. Shi *et al.*, Primary care, infant mortality, and low birth weight in the states of the USA. *Journal of Epidemiology and Community Health* **58**, 374 (2004).

55. J. Currie, J. Gruber, Health insurance eligibility, utilization of medical care, and child health. *the Quarterly Journal of economics* **111**, 431 (1996).

56. M. Marmot, Social determinants of health inequalities. *The Lancet* **365**, 1099 (2005).

57. T. J. Joyce, N. H. Mocan, "Unemployment and Infant Health: Times-Series Evidence from the State of Tennessee" (National Bureau of Economic Research, 1993).

58. J. Hadley, Sicker and poorer—The consequences of being uninsured: A review of the research on the relationship between health insurance, medical care use, health, work, and income. *Medical Care Research and Review* **60**, 3S (2003).

59. R. A. Cohen, *Health insurance coverage trends, 1959-2007: estimates from the National Health Interview Survey*.

60. M. E. Martinez, R. Cohen, Health insurance coverage: early release of estimates from the National Health Interview Survey, January–June 2010. *National Center for Health Statistics. December*, (2011).

61. J. C. Kleinman, S. S. Kessel, Racial-Differences in Low-Birth-Weight - Trends and Risk-Factors. *New England journal of medicine* **317**, 749 (Sep 17, 1987).

62. T. J. Mathews, M. F. MacDorman, Infant mortality statistics from the 2006 period linked birth/infant death data set. *National vital statistics reports : from the Centers for Disease Control and Prevention, National Center for Health Statistics, National Vital Statistics System* **58**, 1 (Apr 30, 2010).

63. G. K. Singh, M. D. Kogan, Persistent socioeconomic disparities in infant, neonatal, and postneonatal mortality rates in the United States, 1969-2001. *Pediatrics* **119**, e928 (Apr, 2007).

64. D. R. Williams, C. Collins, Us Socioeconomic and Racial-Differences in Health - Patterns and Explanations. *Annu Rev Sociol* **21**, 349 (1995).

65. in *Health, United States, 2006: With Chartbook on Trends in the Health of Americans*. (Hyattsville (MD), 2006).

66. in *Health, United States, 2010: With Special Feature on Death and Dying*. (Hyattsville (MD), 2011).

67. J. C. Kleinman, M. B. Pierre, Jr., J. H. Madans, G. H. Land, W. F. Schramm, The effects of maternal smoking on fetal and infant mortality. *American journal of epidemiology* **127**, 274 (Feb, 1988).

68. A. L. Johansson, P. W. Dickman, M. S. Kramer, S. Cnattingius, Maternal smoking and infant mortality: does quitting smoking reduce the risk of infant death? *Epidemiology* **20**, 590 (Jul, 2009).

69. M. S. Kramer, Determinants of low birth weight: methodological assessment and meta-analysis. *Bull World Health Organ* **65**, 663 (1987).

70. R. Olegard *et al.*, Effects on the child of alcohol abuse during pregnancy. Retrospective and prospective studies. *Acta paediatrica Scandinavica. Supplement* **275**, 112 (1979).

71. K. Strandberg-Larsen, M. Gronboek, A. M. Andersen, P. K. Andersen, J. Olsen, Alcohol drinking pattern during pregnancy and risk of infant mortality. *Epidemiology* **20**, 884 (Nov, 2009).

72. T. B. Gage, F. Fang, E. O'Neill, H. Stratton, Maternal age and infant mortality: a test of the Wilcox-Russell hypothesis. *American journal of epidemiology* **169**, 294 (Feb 1, 2009).

73. A. T. Geronimus, The weathering hypothesis and the health of African-American women and infants: evidence and speculations. *Ethnicity & disease* **2**, 207 (Summer, 1992).

74. C. Love, R. J. David, K. M. Rankin, J. W. Collins, Jr., Exploring weathering: effects of lifelong economic environment and maternal age on low birth weight, small for gestational age, and preterm birth in African-American and white women. *American journal of epidemiology* **172**, 127 (Jul 15, 2010).

75. C. Dye, Health and urban living. *Science* **319**, 766 (Feb 8, 2008).

76. J. W. Lynch *et al.*, Income inequality and mortality in metropolitan areas of the United States. *Am J Public Health* **88**, 1074 (Jul, 1998).

77. Publication of Health, United States, 2001 with Urban and Rural Health Chartbook (Reprinted from MMWR, vol 50, pg 784, 2001). *Jama-J Am Med Assoc* **286**, 1832 (Oct 17, 2001).

78. M. S. Eberhardt, E. R. Pamuk, The importance of place of residence: Examining health in rural and nonrural areas. *American Journal of Public Health* **94**, 1682 (Oct, 2004).

79. L. Gamm, L. Hutchison, G. Bellamy, B. J. Dabney, Rural healthy people 2010: Identifying rural health priorities and models for practice. *Journal of Rural Health* **18**, 9 (Win, 2002).

80. M. Grossman, S. Jacobowitz, Variations in infant mortality rates among counties of the United States: the roles of public policies and programs. *Demography* **18**, 695 (Nov, 1981).

81. H. Corman, M. Grossman, Determinants of neonatal mortality rates in the U.S. A reduced form model. *Journal of health economics* **4**, 213 (Sep, 1985).

82. J. Gruber, P. Levine, D. Staiger, Abortion legalization and child living circumstances: Who is the "marginal child"? *Q J Econ* **114**, 263 (Feb, 1999).

83. N. M. McCarty, K. T. Poole, H. Rosenthal, *Polarized America : the dance of ideology and unequal riches*. The Walras-Pareto lectures (MIT Press, Cambridge, Mass., 2006), pp. xii, 240 p.

84. K. T. Poole, H. Rosenthal, D-NOMINATE after 10 years: A comparative update to congress: A political-economic history of roll-call voting. *Legis Stud Quart* **26**, 5 (Feb, 2001).

85. R. Carroll, J. B. Lewis, J. Lo, K. T. Poole, H. Rosenthal, Measuring Bias and Uncertainty in DW-NOMINATE Ideal Point Estimates via the Parametric Bootstrap. *Polit Anal* **17**, 261 (Sum, 2009).

86. J. B. Lewis, K. T. Poole, Measuring bias and uncertainty in ideal point estimates via the parametric bootstrap. *Polit Anal* **12**, 105 (Spr, 2004).

87. T. S. Breusch, A. R. Pagan, Simple Test for Heteroscedasticity and Random Coefficient Variation. *Econometrica* **47**, 1287 (1979).