

The relationship between income and health: evidence from nationwide health exams in Korea*

Liran Einav, Amy Finkelstein, and Sehyun Hong[†]

October 1, 2026

Abstract

We examine how physical health, health behaviors, and mortality vary with income for prime-age adults. To do so, we leverage population-wide, individual-level cardiovascular health exams linked to income tax records and subsequent mortality for a representative sample of about 350,000 Koreans aged 47–57. For both men and women, there is a negative disease-income gradient for high blood pressure, high blood glucose, and high body mass index, but not for high cholesterol. Disease-income gradients are steeper for women than men, and tend to be concentrated in the top half of the income distribution; similar patterns hold for health behavior-income gradients. Subsequent mortality-income gradients are also concentrated in the top half of the income distribution, but are steeper for men than women. Although physical health and behavior measures are highly predictive of subsequent mortality, differences in these measures by income explain very little of the income-mortality gradient, even for cardiovascular mortality. Our findings suggest that something other than differences in disease burden at prime age drives most of the income-mortality gradient at older ages.

*We are grateful to Cheungeun Koo, Joonyoung Seo, Jong-sung You, and the Korean National Health Insurance Service for access to the National Health Insurance Service data (approval no. NHIS-2023-1-821), and to Adriana Lleras-Muney, Hannah Wunsch, and participants at the Junior-ish Health Economics Conference and the Opportunity Insights Conference on Economic Mobility for helpful comments. Jaeyoung Baak, Gayeon Kim, and Cheieun Im provided excellent research assistance. Hong gratefully acknowledges financial support from the French Agence Nationale de la Recherche (ANR-17-EURE-001).

[†]Einav: Stanford and NBER, leinav@stanford.edu; Finkelstein: MIT and NBER, afink@mit.edu; Hong: Paris School of Economics, sehyun.hong@psemail.eu.

1 Introduction

The study of population health has primarily meant the study of mortality. While of course an important outcome, mortality is only one measure of health, and arguably not well-suited for understanding the health of non-elderly adults. For them, mortality is both low and disproportionately from external causes (e.g., [Centers for Disease Control and Prevention 2010](#)). In this paper, therefore, we leverage novel, individual data linkages between population-wide cardiovascular health exams and income tax records to provide new evidence of the relationships between clinical health measures and income for prime-age adults, and explore their ability to explain subsequent mortality-income gradients.

Higher income individuals live longer. This robust pattern appears over time and across countries (e.g., [Cutler et al. 2006](#); [Costa 2015](#)), as well as across individuals within many high-income countries (e.g., [Mustard et al. 1997](#); [Chetty et al. 2016](#); [Mortensen et al. 2016](#); [Kinge et al. 2019](#); [Chen et al. 2022](#)). The nature and drivers of these gradients are naturally of great academic and policy interest. Yet the relative paucity of information about how non-mortality measures of health vary with income leaves open important questions about the relationship between income and health for prime-age adults, as well as how any such relationship relates to subsequent mortality-income gradients.

Data limitations have been the central challenge to analyzing non-mortality health-income gradients. Analysis requires data with reliable health measures for large, representative samples, linked to high-quality income measures. Existing work has looked mostly at socioeconomic gradients in self-reported health (e.g., [Case et al. 2002](#); [House et al. 2005](#); [Case and Deaton 2007](#); [Mackenbach et al. 2008](#); [Cutler et al. 2011](#)), in diagnoses recorded in administrative health data (e.g., [Kivimäki et al. 2020](#); [Danesh et al. 2024](#); [Pallesen et al. 2024](#)), or in biomarkers (e.g., [Kavanagh et al. 2010](#); [Davillas and Jones 2020](#); [Shui et al. 2024](#)). Such studies have provided valuable insights, but also suffer from important limitations. Self-reported health is non-specific and inherently subjective (potentially differentially so by socioeconomic status), diagnosis records suffer from selection in who seeks care and what is measured and recorded, even in the presence of universal coverage ([Song et al. 2010](#); [Baum et al. 2020](#)), and studies of biomarkers for representative populations have tended to be quite small (typically on the order of about 10,000 adults). In addition, existing studies have tended to rely on coarse measures of socioeconomic status such as education or self-reported, broad income bins.¹

In this paper, we seek to overcome these challenges by leveraging individual-level data from population-wide, cardiovascular health exams in Korea, which we have linked both to individual-

¹Two important recent exceptions are from work in the Netherlands. [Shui et al. \(2024\)](#) examine gradients in biomarkers by education in a representative sample of about 150,000 Dutch adults aged 18 to 80, a much larger sample than the prior biomarker studies. Likewise, [Danesh et al. \(2024\)](#) study the relationship between chronic disease – measured by prescription drug fills – and income for the universe of Dutch adults using rich measures of income in the Dutch tax data.

level tax records and to subsequent, cause-specific mortality. Since the mid-1990s, Korea has provided free annual or biennial health exams for all individuals aged 40 and over, with a primary focus on cardiovascular and cerebrovascular disease and its risk factors. The health exams record clinical data on blood pressure, fasting blood glucose, height and weight, and total cholesterol, as well as self-reported health behaviors (including smoking, drinking, and exercise), current medications, and basic demographics.

For a random eight percent of the population, we linked individual-level annual income tax records from 2006 through 2008 to in-person health exams from 2009 through 2012, and to cause-of-death through 2019. We focus our analysis on individuals who were 47–57 at the time of their health exam, an age range that was chosen to balance between a population young enough that mortality is a limited measure of health, but old enough that subsequent, eight-year cardiovascular-disease mortality is quantitatively meaningful. Our target population consists of almost 500,000 individuals with linked income-mortality data, of whom approximately 350,000 individuals completed a health exam in at least one year between 2009 and 2012.²

We have two main sets of findings. The first set of results concerns the nature of the disease-income and behavior-income gradients. For both men and women, the disease-income gradient is negative for rates of high blood pressure, high blood glucose, and high body mass index, but not for rates of high cholesterol. Disease-income gradients are steeper for women than men, and tend to be concentrated in the top half of the income distribution. We observe qualitatively similar income gradients for health behaviors, specifically smoking, drinking, and exercise: behaviors improve with income, especially in the top half of the income distribution and especially for women. Differences in successful use of medication to control high biomarkers for blood pressure and blood glucose do not seem to be a main contributor to those observed disease-income gradients.

Mortality-income gradients are also concentrated in the top half of the income distribution, but are steeper for men than women. Steeper mortality-income gradients for men have also been documented in several other high-income countries (Chetty et al. 2016; Mortensen et al. 2016; Kinge et al. 2019; Chen et al. 2022) but are not well-understood. One possible explanation has been that household income is less of a proxy for life-time socioeconomic status for women, especially in countries (including Korea) where female labor force participation is low. However, our finding that disease and behavior-income gradients are steeper for women than men suggests something other than measurement error may be behind these patterns.

Our second set of findings concerns the lack of a clear connection between the disease and behavior income gradients and the subsequent mortality income gradients. The fact that the former are steeper for women while the latter are steeper for men is already suggestive that the

²As we discuss in more detail, participation rates in the health exams tend to increase with income, raising concerns about potential selection bias. Specifically, because exam participants are, on average, healthier (lower subsequent mortality) than non-participants, rising participation rates with income could potentially bias downward estimates of improvement in clinical health with income. However, we present several pieces of evidence that suggest that non-participation is unlikely to drive our estimates of these health-income gradients.

subsequent mortality income gradient may not be tightly related to prior disease and behavior income gradients. More direct evidence on this point comes from examining predicted mortality based on the underlying biomarker and behavior measures. Although variation in the underlying disease and behavior measures is associated with substantial variation in predicted subsequent mortality, the magnitudes of the disease-income and behavior-income gradients are not sufficient to generate much variation in predicted mortality by income, even when focusing on cardiovascular mortality.

Taken together, our findings suggest that differences in cardiovascular health by socioeconomic status around age fifty are not the primary driver of subsequent mortality-income gradients. This is consistent with the findings of the Whitehall studies (Marmot et al. 1978, 1991), which measured cardiovascular health and subsequent cardiovascular disease and cardiovascular mortality for a cohort of British civil servants and found that one’s civil service grade was a much more important predictor of subsequent health outcomes than measures of cardiovascular health. Our findings are also consistent with the literature on the myriad pathways by which a lower position within a social hierarchy can affect cardiovascular health without involving traditional risk factors (Brondolo et al. 2017; Geronimus 2023; Lavetti et al. 2025). Our findings underscore the importance of looking at these other pathways for understanding socioeconomic gradients in cardiovascular mortality.

2 Setting

Korea is a high-income country with a population of about 50 million people. It has a progressive income tax with the individual as the unit of taxation (Hong et al. 2024). Universal national health insurance covers about two-thirds of national medical expenditures, and about three-quarters of Koreans have supplementary private insurance (Shin et al. 2022). Life expectancy at birth is high, about 83 years compared to 78 in the US (OECD 2025), and there is a substantial mortality-income gradient (e.g., Khang and Kim 2005; Khang et al. 2016; Shin et al. 2018; Kim et al. 2021). Survey data that measure biomarkers suggest that some cardiovascular disease risk factors (such as high blood pressure) are higher in Korea than in the US while others (such as high body mass index) are lower (Appendix Table A1).³

Since the mid-1990s, free health exams have been available for all individuals aged 40 and older. They are annual for manual workers (i.e., non-desk jobs) and biennial for everyone else.⁴ The exams focus primarily on cardiovascular and cerebrovascular health and its risk factors, as part of an effort to improve population health (Kang 2022a; Shin et al. 2022).

The Korean National Health Insurance Service (NHIS) automatically informs individuals when they are due for an exam and provides them with the results. Exams are legally mandatory

³Using these same survey data, Lawton and Cutler (2026) describe how risk factors have trended over time in the US compared to Korea and England. They also note that disparities by education level exist in all three countries.

⁴Examples of “manual workers” include, among others, drivers, flight attendants, doctors, chefs, and singers.

for employees, with financial penalties on the employer if their employee does not comply. All individuals (whether employed or not) face higher cost-sharing requirements associated with a cancer diagnosis if they are not in compliance (Shin et al. 2022; Ministry of Health and Welfare 2023). The compliance rate with the recommended health exam schedule has been gradually increasing over our 2009–2019 study period, from about 60% to about 75% (Kang 2022a; National Health Insurance Service 2022; Shin et al. 2022).

The health exam collects basic demographics, self-reported medical history and health behaviors, physical measurements, and blood samples. Specifically, individuals fill out a questionnaire with information about their own and their family’s past medical history, whether they currently take medication for certain conditions, including for hypertension and diabetes, and current and past health behaviors (including smoking, alcohol consumption, and exercise). They then undergo a series of measurements (blood pressure, waist circumference, weight, and height), as well as blood draws which are used to measure cholesterol and blood glucose. Conditional on participating in the health exam, there is virtually no missing data for the health measures we examine. Appendix A provides more detail on the Korean setting in general and the health exams in particular.

3 Data

In this section, we briefly describe the construction of the sample and the key variables, and then report summary statistics. Appendix B provides many additional details.

3.1 Sample construction

We obtained access to a random 8% of the universe of individual-level annual income tax records from 2006 through 2008 from the Korean National Tax Service. These individuals are then linked to their death records (if any) through 2019 (from Statistics Korea) and to NHIS data on any in-person health exams completed during 2009–2012.⁵

We focus on individuals born between 1955 and 1962 who are still alive in 2008. We refer to these individuals as the *target population*, and the subset of individuals in the target population for whom we observe at least one health exam (during our 2009–2012 exam window) as the *exam population*. We use the first health exam in this period. The exam population therefore spans the 47–57 age range at the time of their exam. We selected that age range to balance between a population young enough so that sample selection associated with survival is quantitatively less important, yet old enough so that subsequent, eight-year cardiovascular mortality rates are meaningful (Centers for

⁵Previous work has used the NHIS health exam data to study, among other things, the impact of job loss on health outcomes (Hwang et al. 2024), the impact of alcohol consumption on dementia (Jeon et al. 2023), and the prevalence of obesity (Jeong et al. 2024). Linked health exam and mortality data have been previously used to study the relationship between socioeconomic status, cardiovascular health, and mortality among male public servants, with income-related insurance contributions used to proxy for two to four income groups (Song and Byeon 2000; Khang et al. 2008; Kang 2022b).

Disease Control and Prevention 2010; Statistics Korea 2023). For younger adults, a larger share of deaths are from external causes that are unlikely to be affected by cardiovascular health measures (Appendix Table A2).

As we discuss below, exam participants are not a random subsample of the target population. However, we present several pieces of evidence that suggest that this selection is unlikely to substantially affect our key results.

3.2 Key variables

Income. For each individual in the sample, we closely follow Chetty et al. (2016) and convert their observed income into an income percentile. Specifically, we define household income as the sum of their taxable income and their spouse’s (if any).⁶ After excluding individuals with zero taxable (household) income, we compute each individual’s percentile rank for their yearly household income relative to the household income of all others with the same birth year and sex in that year’s tax data. To reduce noise, we then compute the average household income percentile over 2006, 2007, and 2008.

Thus, income percentiles are all conditional on positive (household) income. Throughout the analysis, we also report results separately for the approximately one-quarter of individuals who had zero (household) income in all three years (2006–2008). As we discuss in more detail in Appendix B.1, these individuals are likely very low income.

Biomarkers. We focus on four measures of physical health that are viewed as key predictors of cardiovascular and cerebrovascular diseases: blood pressure, (fasting) blood glucose, body mass index (BMI), and cholesterol (Wawrzyniak 2020; Powell-Wiley et al. 2021; Centers for Disease Control and Prevention 2024b; American Diabetes Association Professional Practice Committee 2025). In each case, we observe the underlying continuous biomarkers, and convert them to binary variables (“high” vs. normal) using standard cutoffs; throughout, we will use the terms ‘high biomarker’, ‘biomarker out of range’, and ‘disease’ interchangeably. For some of the analysis, we also define an intermediate, “elevated” category, again using standard cutoffs.⁷

Health behaviors. We observe two types of health behaviors. First, we observe three health behaviors that are important potential contributors to cardiovascular disease: smoking, alcohol consumption, and exercise. Based on participants’ (self-reported) answers and standard cutoffs, we define binary indicators for current smoking, heavy smoking, current drinking, heavy drinking,

⁶Taxable income includes labor income, dividend income, interest income, business income, pension income, and other income.

⁷Our blood pressure cutoffs are based on Ajay et al. (2007). American Diabetes Association (2024) provides cutoffs for blood glucose. For BMI, we use Centers for Disease Control and Prevention (2024a). Lastly, for cholesterol, we use cutoffs from Johns Hopkins Medicine (2024).

any exercise, and sufficient exercise.⁸ Second, we examine medication use for high blood pressure and for high blood glucose. Specifically, we observe whether each participant self-reports that they currently take prescribed medication for hypertension and whether they currently take prescribed medication for diabetes. We further link their medication use to their corresponding biomarker level in order to construct – among those who use medication – indicators for a “successfully controlled” and “unsuccessfully controlled” biomarker, which we interpret as proxies for adherence.

Mortality. We observe all-cause and cause-specific mortality (if any) through 2019. We focus on three mortality measures: all-cause mortality, the subset of all-cause mortality that is from non-external causes, and the subset of non-external mortality that is from cardiovascular disease. Our primary mortality variable is an indicator for whether the individual died within eight years after their health exam.

Predicted mortality. We generate predicted mortality using a linear probability model in which we use the predicted value from a regression of an eight-year mortality indicator on a fully saturated set of the four biomarkers and their three possible values (high, elevated, or normal) (that is, $3^4 = 81$ binary regressors which capture all possible biomarker combinations) as well as on a similar saturated model that also includes the three health behaviors (smoking, drinking, and exercising) and their three possible values (which implies $3^7 = 2,187$ binary regressors).

3.3 Summary statistics

Table 1 presents summary statistics for the exam population, which is our primary analysis sample. It contains about 350,000 people (70% of the target population) who participated in at least one health exam during 2009–2012. About 20% of men have high blood pressure, 11% have high blood glucose, 3% have high BMI, and 13% have high cholesterol. For women, the comparable rates are 13%, 5%, 4%, and 17%. Smoking and drinking rates are substantially higher among men (42% and 68%, respectively) than among women (4% and 23%, respectively) while exercise rates are more similar (71% for men and 66% for women).

For high blood pressure and high blood glucose, we can further distinguish between those who are and are not using medication. This yields four categories of individuals based on whether they are on medication (either with biomarkers in range (“successful control”) or biomarkers out of range (“unsuccessful control”)) and those who are not on medication (either with biomarkers out of range or with biomarkers in range). Women tend to exhibit better medication behavior: they

⁸Heavy smoking is defined as 20 or more cigarettes a day (Hirayama 1981), heavy drinking as 15 or more drinks a week for men and 8 or more drinks a week for women (National Institute on Alcohol Abuse and Alcoholism 2007), and sufficient exercise is defined as at least 150 minutes of moderate exercise or 75 minutes of vigorous exercise a week (World Health Organization 2010).

are more likely to take medication conditional on needing it, and that medication is more likely to control their condition successfully.⁹

The eight-year all-cause mortality rate in the exam population is 1.0% for women and 2.7% for men; for both sexes, about two-thirds of this mortality is associated with non-external causes, and about 10–15% of the mortality is from cardiovascular causes.

Individually, both high biomarkers and poor health behaviors are highly predictive of subsequent mortality for both sexes (Appendix Table A3). High blood pressure, high blood glucose, and high BMI are each associated with higher eight-year mortality (for all three mortality outcomes) but, consistent with other findings (e.g., Schatz et al. 2001; Ravnskov et al. 2016), high cholesterol is associated with lower mortality. Having high blood glucose, for example, roughly triples the eight-year cardiovascular mortality rate. Likewise, smoking is associated with substantially elevated mortality, roughly tripling the eight-year cardiovascular mortality rate, while exercise has an opposite-signed effect that is roughly similar in magnitude. Drinking is associated with lower mortality.

Together, these measures generate substantial variation in predicted mortality (Appendix Table A4). When the four biomarker measures are combined to form predicted mortality, the standard deviation of predicted mortality is at least 40% of the mean of observed mortality. Adding the three health behaviors substantially increases this predictive power.

Sample selection. Importantly, the exam population is not a random subsample of the target population. Consistent with other evidence that individuals who select into following recommended health practices tend to be healthier (Einav et al. 2020; Oster 2020; Kowalski 2023), for both men and women, all-cause mortality rate is about 40% lower for the exam population than for the target population (Appendix Table A5).¹⁰ In addition, the probability of participation rises noticeably with income until about the 85th to 90th income percentile, and then falls; the participation-income gradient is particularly pronounced for men (Appendix Figure A1). This same selection pattern is evident in other studies examining the relationship between socioeconomic status and biomarkers (Marmot et al. 1991; Kavanagh et al. 2010; Davillas and Jones 2020; Shui et al. 2024).

The combination of non-random participation by health and differences in participation rates by income raises important concerns about selection into the exam across the income distribution. As participation rates increase, exam participants become less selected and therefore less healthy. This could lead us to under-estimate (in absolute value) the slope of the negative biomarker-income gradients up to about the 85th percentile – since exam participants are healthier and participation is increasing with income – and then over-estimate it after that for the reverse reason. However, in

⁹Survey data indicate that for those who need medication, rates of successful control are higher in the US than in Korea (Lawton and Cutler 2026).

¹⁰Interestingly, despite the different mortality rates, the share of deaths due to cardiovascular diseases is quite similar for the exam population and the target population.

Section 4.4 below, we present several pieces of evidence that suggest that, in practice, our gradient estimates are unlikely to be substantially affected by such selection.

4 Results

Throughout this section, we report outcomes by household income percentile, separately for men and women. To make these income gradients more comparable across genders and outcomes, we report mean outcomes in relative log points; that is, we report the (natural) log of the outcome measure at a given percentile of the income distribution minus the log of the overall (gender-specific) average outcome. As a result, the health measure that is reported at a given income percentile is essentially a multiplicative factor of the population average. We then smooth this relative outcome at each income percentile by taking a moving average over 5 percentiles (the one shown and two on either side); we therefore report results from the 3rd to the 98th income percentile. Recall that income percentiles are computed conditional on positive (household) income; in each plot we therefore also report the corresponding outcome for those individuals who have zero income in all years.¹¹

4.1 Disease-income gradients

Figure 1 shows disease-income gradients for each of the four biomarkers, separately for men (left panels) and women (right panels); Appendix Figure A2 provides the raw rates. While there is little indication of an income gradient in high cholesterol for either men or women, the qualitative patterns for all three other biomarkers – high blood pressure, high blood glucose, and high BMI – are quite similar.¹²

All three show pronounced income gradients in the expected direction, with higher income individuals associated with lower rates of high levels of each biomarker. All three income gradients are much steeper for women than for men, and the gradients – for both men and women – tend to be concentrated in the top half of the income distribution.

For instance, 11% of men have high blood glucose, but rates of high blood glucose at the 95th percentile of the income distribution are 26% lower compared to the 50th percentile, but the difference between the 50th and 5th percentile is only 10%. For women, rates of high blood glucose (which are 5% on average) are 41% lower at the 95th compared to the 50th percentile, but only 21% lower for the 50th percentile compared to the 5th percentile.

The concentration of the disease-income gradients in the top half of the income distribution is striking in light of prior evidence based on self-reported health, mortality, and diagnoses that in

¹¹For completeness, in the appendix we also report corresponding “raw” rates at each household income percentile as well as for those with zero household income, prior to taking logs and smoothing.

¹²These biomarker-income gradients are unaffected by adjusting for age at the time of the health exam (Appendix Figure A3).

other countries (such as the US, Sweden, and the Netherlands), health-income gradients are steeper at the bottom of the income distribution (Chen et al. 2022; Danesh et al. 2024; Lleras-Muney et al. 2025).

4.2 Behavior-income gradients

The health-income gradients in Figure 1 reflect some combination of differential rates of onset of high levels of biomarkers by income and differential behavior and disease management. To shed some light on these channels, we can look directly at income gradients in health behaviors.

Figure 2 shows behavior-income gradients for smoking, drinking, and exercise for men (left panels) and women (right panels); Appendix Figure A4 provides the raw rates. Like the disease-income gradients shown above, income gradients for health behavior also tend to be more pronounced for women than men, and steeper in the top half of the income distribution. Smoking rates in particular decline sharply with income. For example, for men, smoking rates average 42% but are 31% lower at the 95th compared to the 50th percentile, while they are almost identical for the 50th and 5th percentiles. Smoking rates are much lower for women (4%), but the gradient is much steeper; smoking rates are 60% lower at the 95th compared to the 50th percentile.

For both high blood pressure and high blood glucose, we can also look at how rates of medication used to control these conditions vary by income. Figure 3 summarizes the implications of the medication-income gradients by comparing the disease-income gradients (shown in blue) to those produced by two counterfactuals.¹³ The first (shown in red) considers what the biomarker-income gradient would look like if, counterfactually, medication were unavailable and thus anyone who currently is on medication and has biomarkers in range is now out of range. This substantially flattens the high blood pressure income gradient for men, but makes little difference for the other three gradients. The second counterfactual (shown in green) considers what the biomarker-income gradient would look like if all medication successfully controlled the disease; this makes very little difference for any of the gradients. Thus, it appears that for high blood pressure, higher use of medication by higher income men steepens the male disease-income gradient, but for high blood glucose for men and for both diseases for women, differential rates of medication usage appear to play little role in generating the disease-income gradient. This suggests that differential rates of onset are likely more important.

4.3 Mortality-income gradients

Figure 4 shows mortality-income gradients and predicted-mortality income gradients for all-cause mortality, mortality from non-external causes, and cardiovascular mortality, separately for men (left panels) and women (right panels); Appendix Figure A9 and Appendix Figure A10 provide the raw rates.

¹³Appendix Figures A5 through A8 show the underlying gradients.

Similar to the health-income gradient, mortality-income gradients are also concentrated in the top half of the income distribution. Yet, unlike the health-income and behavior-income gradients, mortality-income gradients are much more pronounced for men than for women. For example, cardiovascular mortality rate is 70% lower for men at the 95th percentile of the income distribution relative to the 50th percentile, but does not vary much with income in the lower half of the income distribution. Women exhibit a similar qualitative pattern.

Yet, predicted mortality varies much less with income. This holds whether we use only the four biomarkers for the prediction (lighter green) or additionally use the three health behaviors (darker green). For both sexes, when we use only biomarkers to predict mortality, the difference between predicted mortality rates at the 95th percentile of the income distribution and at the 50th percentile is about ten times lower than the difference in actual mortality rates; it is about three times lower when health behaviors are also included as predictors. For example, predicted cardiovascular mortality using biomarkers only declines by 7% from the 50th to the 95th percentile for both sexes; it declines by 27% (for men) and 11% (for women) when health behaviors are included for prediction, but this is still much smaller than the decline in actual cardiovascular mortality (of 70% for men and 56% for women).

This finding exists despite the quantitatively meaningful relationship between high biomarkers and mortality. For example, although having high blood glucose roughly triples the eight-year cardiovascular mortality rate (Appendix Table A3), the 41% lower rate of high blood glucose for higher income women only generates a 4% difference in predicted cardiovascular mortality, a small fraction of the 56% difference in actual cardiovascular mortality between the 95th and 50th percentile. This is because only 5% of women had high blood glucose, so higher income women had only a 2 percentage point lower rate of high blood glucose.

4.4 Sensitivity to sample selection

As noted at the end of Section 3, our disease-income gradient estimates may be biased by non-random participation: exam participants are associated with lower mortality (and therefore likely healthier) than the target population and participation rates vary with income. However, multiple pieces of evidence suggest that such bias is unlikely to be meaningful in practice.

First, the pattern of fairly monotone biomarker-income gradients, with the slope steepening at above median income levels (except for cholesterol where it is flat throughout), is very different from the non-monotonic participation rate patterns by income (Appendix Figure A1).

Second, we can gauge the likely extent of any selection bias by looking at mortality gradients since, unlike the health measures, mortality is observed for the entire target population. Appendix Figure A11 indicates that the slope of the mortality gradient is very similar for the exam population and the target population, which is suggestive that the health-income gradient may also be similar.

Third, biomarker-income gradients look similar for members of the exam population who par-

ticipate in more vs. fewer of their scheduled exams. Specifically, we focus on the 80% of the target population who are scheduled to have two exams over our 2009–2012 exam window and compare biomarker-income gradients for those who had two exams and those who had only one. Slightly over half of the exam population who were scheduled for two exams actually participated in both. The direction of selection for single-time vs two-time participants looks similar to that for non-participants vs. exam participants. In particular, just as exam participants have lower mortality than non-participants (Appendix Table A5), two-time participants have lower mortality than single-time participants (see Appendix Table A6); and just as the probability of exam participation is generally higher for higher income individuals (Appendix Figure A1), so too is the probability of participating twice rather than once (Appendix Figure A12). Reassuringly, however, Appendix Figures A13 and A14 show that the disease-income and behavior-income gradients look very similar for one-time and two-time participants.

5 Conclusion

Understanding and addressing socioeconomic gradients in health is of substantial academic and policy interest, yet, outside of mortality, we know relatively little about the nature of these gradients. In this study, we examined population-wide health exams for non-elderly adults linked to income tax data to obtain estimates of the relationship between income and a variety of different health measures. Specifically, we examined four biomarkers that are key sources of cardiovascular mortality and morbidity, several health behaviors, and subsequent cardiovascular and all-cause mortality.

The general pattern indicates negative disease-income gradients, mortality-income gradients, and poor health behavior-income gradients. These gradients are all concentrated in the top half of the income distribution. This is noteworthy in light of other work using self-reported health that has found steeper gradients at the bottom of the income distribution (Lleras-Muney et al. 2025).

Although both biomarkers and health behaviors are highly predictive of subsequent mortality for both sexes, disease and behavior gradients can explain very little of the mortality-income gradients. In addition, while disease and health behavior gradients are much steeper for women than for men, the opposite is true for mortality gradients. Taken together, our findings suggest that differences in cardiovascular health by socioeconomic status around age 50 are not the primary driver of subsequent mortality-income gradients, either for cardiovascular or for all-cause mortality. As noted in the Introduction, our findings, together with the existing literature, suggest the importance of looking for the role of other pathways besides traditional risk factors for understanding socioeconomic gradients in cardiovascular disease mortality.

At the same time, our findings also raise interesting questions about the drivers of the disease-income gradients. The gradients we observe reflect the net result of disease onset, diagnosis and awareness, and disease management, or lack thereof (Currie and Stabile 2003; Lawton and Cutler 2026). For high blood pressure and high blood sugar – where we could observe medication usage –

we found relatively little role for successful medication use in contributing to the observed health-income gradients. However, future work exploiting the panel nature of the data together with additional information on diagnosis history and medical care will hopefully be able to shed more light on the underlying drivers of these gradients.

More broadly, the lack of consistently and comprehensively measured data on health outcomes other than mortality has considerably hampered the field of health economics. The nation-wide Korean health exams offer a rare and rich opportunity to learn more about the causes and consequences of a broader range of health measures.

References

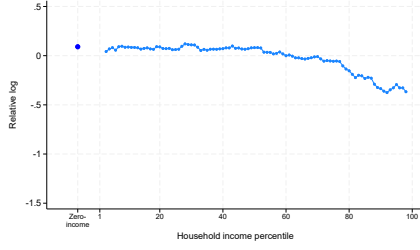
- Ajay, Vamadevan S., Dorairaj Prabhakaran, and Kolli Srinath Reddy, "Chapter 40 - Electrolyte Intake and Human Hypertension," in Gregory Y.H. Lip and John E. Hall, eds., *Comprehensive Hypertension*, Mosby, 2007, pp. 477–482.
- American Diabetes Association, "Understanding Diabetes Diagnosis," <https://diabetes.org/about-diabetes/diagnosis> 2024. Accessed: 2024-12-02.
- American Diabetes Association Professional Practice Committee, "10. Cardiovascular Disease and Risk Management: Standards of Care in Diabetes—2025," *Diabetes Care*, 2025, 48 (Supplement 1), S207–S238.
- Baum, Aaron, Juan Wisnivesky, Sanjay Basu, Albert L Siu, and Mark D Schwartz, "Association of geographic differences in prevalence of uncontrolled chronic conditions with changes in individuals' likelihood of uncontrolled chronic conditions," *Jama*, 2020, 324 (14), 1429–1438.
- Brondolo, Elizabeth, Kahaema Byer, Peter J Gianaros, Cindy Liu, Aric A Prather, Kamala Thomas, Cheryl L Woods-Giscombé, L Beatty, Patricia DiSandro, and G Keita, "Stress and health disparities: Contexts, mechanisms, and interventions among racial/ethnic minority and low socioeconomic status populations," *American Psychological Association (APA) Working Group Report*, 2017, 64.
- Case, Anne and Angus Deaton, "Broken Down by Work and Sex: How Our Health Declines," *Analyses in the Economics of Aging*, 2007, p. 185.
- , Darren Lubotsky, and Christina Paxson, "Economic status and health in childhood: The origins of the gradient," *American Economic Review*, 2002, 92 (5), 1308–1334.
- Centers for Disease Control and Prevention, "National Vital Statistics System, Mortality 2005-2006 on CDC WONDER On-line Database," 2010. Accessed: 2025-07-11.
- , "Adult BMI Categories," <https://www.cdc.gov/bmi/adult-calculator/bmi-categories.html> 2024. Accessed: 2024-12-02.
- , "Heart Disease Risk Factors," Heart Disease December 2024. Accessed August 6, 2025.
- Chen, Yiqun, Petra Persson, and Maria Polyakova, "The roots of health inequality and the value of intrafamily expertise," *American Economic Journal: Applied Economics*, 2022, 14 (3), 185–223.
- Chetty, Raj, Michael Stepner, Sarah Abraham, Shelby Lin, Benjamin Scuderi, Nicholas Turner, Augustin Bergeron, and David Cutler, "The association between income and life expectancy in the United States, 2001-2014," *Jama*, 2016, 315 (16), 1750–1766.
- Costa, Dora L, "Health and the Economy in the United States from 1750 to the Present," *Journal of Economic Literature*, 2015, 53 (3), 503–570.
- Currie, Janet and Mark Stabile, "Socioeconomic status and child health: why is the relationship stronger for older children?," *American Economic Review*, 2003, 93 (5), 1813–1823.
- Cutler, David, Angus Deaton, and Adriana Lleras-Muney, "The Determinants of Mortality," *Journal of Economic Perspectives*, 2006, 20 (3), 97–120.
- Cutler, David M., Adriana Lleras-Muney, and Tom Vogl, "Socioeconomic Status and Health: Dimensions and Mechanisms," in "The Oxford Handbook of Health Economics," Oxford University Press, 2011, pp. 124–163.
- Danesh, Kaveh, Jonathan T Kolstad, Johannes Spinnewijn, and William D Parker, "The Chronic Disease Index: Analyzing Health Inequalities Over the Lifecycle," Technical Report, National Bureau of Economic Research 2024.
- Davillas, Apostolos and Andrew M Jones, "Ex ante inequality of opportunity in health,

- decomposition and distributional analysis of biomarkers,” *Journal of health economics*, 2020, *69*, 102251.
- Einav, Liran, Amy Finkelstein, Tamar Oostrom, Abigail Osterker, and Heidi Williams**, “Screening and selection: The case of mammograms,” *American Economic Review*, 2020, *110* (12), 3836–3870.
- Geronimus, Arline**, *Weathering: The extraordinary stress of ordinary life on the body in an unjust society*, Hachette UK, 2023.
- Hirayama, Takeshi**, “Non-smoking wives of heavy smokers have a higher risk of lung cancer: a study from Japan,” *Br Med J (Clin Res Ed)*, 1981, *282* (6259), 183–185.
- Hong, Sehyun, Nak-Nyeon Kim, Zhexun Mo, and Li Yang**, “Income Inequality in South Korea, 1933-2022: Evidence from Distributional National Accounts,” *World Inequality Lab Working Paper 2024/03*, 2024.
- House, James S, Paula M Lantz, and Pamela Herd**, “Continuity and change in the social stratification of aging and health over the life course: evidence from a nationally representative longitudinal study from 1986 to 2001/2002 (Americans’ Changing Lives Study),” *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 2005, *60* (Special_Issue_2), S15–S26.
- Hwang, Jisoo, Hyuncheol Brian Kim, and Jungmin Lee**, “Is Job Loss Always Bad for Health? Evidence from Health Screening Data,” 2024.
- Jeon, Keun Hye, Kyungdo Han, Su-Min Jeong, Junhee Park, Jung Eun Yoo, Juhwan Yoo, Jinkook Lee, SangYun Kim, and Dong Wook Shin**, “Changes in alcohol consumption and risk of dementia in a nationwide cohort in South Korea,” *JAMA network open*, 2023, *6* (2), e2254771–e2254771.
- Jeong, Su-Min, Jin-Hyung Jung, Ye Seul Yang, Wonsock Kim, In Young Cho, You-Bin Lee, Kye-Yeung Park, Ga Eun Nam, and Kyungdo Han**, “2023 Obesity Fact Sheet: Prevalence of Obesity and Abdominal Obesity in Adults, Adolescents, and Children in Korea from 2012 to 2021,” *Journal of Obesity & Metabolic Syndrome*, 2024, *33* (1), 27.
- Johns Hopkins Medicine**, “Lipid Panel,” <https://www.hopkinsmedicine.org/health/treatment-tests-and-therapies/lipid-panel> 2024. Accessed: 2024-12-02.
- Kang, Hee-Taik**, “Current status of the national health screening programs in South Korea,” *Korean Journal of Family Medicine*, 2022, *43* (3), 168.
- Kang, Hee-Yeon**, “The Contribution of Repeatedly Measured Behavioral and Bioclinical Factors to Socioeconomic Mortality Inequalities in Korea,” 2022. Accessed: December 8, 2025.
- Kavanagh, Anne, Rebecca J Bentley, Gavin Turrell, Jonathan Shaw, David Dunstan, and SV Subramanian**, “Socioeconomic position, gender, health behaviours and biomarkers of cardiovascular disease and diabetes,” *Social science & medicine*, 2010, *71* (6), 1150–1160.
- Khang, Young-Ho and Hye Ryun Kim**, “Relationship of education, occupation, and income with mortality in a representative longitudinal study of South Korea,” *European journal of epidemiology*, 2005, *20*, 217–220.
- , **Jinwook Bahk, Nari Yi, and Sung-Cheol Yun**, “Age-and cause-specific contributions to income difference in life expectancy at birth: findings from nationally representative data on one million South Koreans,” *The European Journal of Public Health*, 2016, *26* (2), 242–248.
- , **JW Lynch, Kyunghye Jung-Choi, and Hong-Jun Cho**, “Explaining age-specific inequalities in mortality from all causes, cardiovascular disease and ischaemic heart disease among South Korean male public servants: relative and absolute perspectives,” *Heart*, 2008, *94* (1), 75–82.
- Kim, Minhye, Suzin You, Jong sung You, Seung-Yun Kim, and Jong Heon Park**,

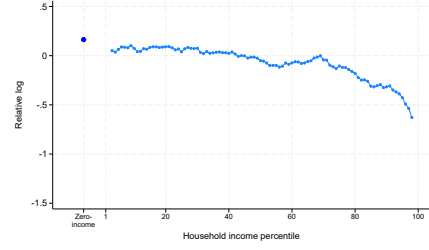
- “Income-Related Mortality Inequalities and Its Social Factors among Middle-Aged and Older Adults at the District Level in Aging Seoul: An Ecological Study Using Administrative Big Data,” *International Journal of Environmental Research and Public Health*, 2021, 19 (1), 383.
- Kinge, Jonas Minet, Jørgen Heibø Modalsli, Simon Øverland, Håkon Kristian Gjessing, Mette Christophersen Tollånes, Ann Kristin Knudsen, Vegard Skirbekk, Bjørn Heine Strand, Siri Eldevik Håberg, and Stein Emil Vollset**, “Association of household income with life expectancy and cause-specific mortality in Norway, 2005-2015,” *Jama*, 2019, 321 (19), 1916–1925.
- Kivimäki, Mika, G David Batty, Jaana Pentti, Martin J Shipley, Pyry N Sipilä, Solja T Nyberg, Sakari B Suominen, Tuula Oksanen, Sari Stenholm, Marianna Virtanen et al.**, “Association between socioeconomic status and the development of mental and physical health conditions in adulthood: a multi-cohort study,” *The Lancet Public Health*, 2020, 5 (3), e140–e149.
- Kowalski, Amanda E**, “Behaviour within a clinical trial and implications for mammography guidelines,” *The Review of economic studies*, 2023, 90 (1), 432–462.
- Lavetti, Kurt J, Long Hong, Jonathan A Holmes, and Trevon D Logan**, “Workplace stratification and racial health disparities,” Technical Report, National Bureau of Economic Research 2025.
- Lawton, Ralph and David M. Cutler**, “Cardiovascular Risk Factor Prevalence and Control by Education Strata in Older Adults in 3 Countries,” *JAMA Network Open*, 2026, 9(7), e2625410.
- Lleras-Muney, Adriana, Hannes Schwandt, and Laura R Wherry**, “Poverty and health,” *Annual Review of Economics*, 2025, 17 (1), 31–56.
- Mackenbach, Johan P, Irina Stirbu, Albert-Jan R Roskam, Maartje M Schaap, Gwenn Menvielle, Mall Leinsalu, and Anton E Kunst**, “Socioeconomic inequalities in health in 22 European countries,” *New England journal of medicine*, 2008, 358 (23), 2468–2481.
- Marmot, Michael G, Abraham M Adelman, Nicola Robinson, and Geoffrey A Rose**, “Changing social-class distribution of heart disease,” *Br Med J*, 1978, 2 (6145), 1109–1112.
- , **Manolis Kogevinas, and Mary Ann Elston**, “Socioeconomic status and disease,” *WHO Reg Publ Eur Ser*, 1991, 37, 113–146.
- Ministry of Health and Welfare**, “Special Assistance for Cancer Patients,” 2023.
- Mortensen, Laust H, Johan Rehnberg, Espen Dahl, Finn Diderichsen, Jon Ivar Elstad, Pekka Martikainen, David Rehkopf, Lasse Tarkiainen, and Johan Fritzell**, “Shape of the association between income and mortality: a cohort study of Denmark, Finland, Norway and Sweden in 1995 and 2003,” *BMJ open*, 2016, 6 (12), e010974.
- Mustard, Cameron A, Shelley Derksen, Jean-Marie Berthelot, Michael Wolfson, and Leslie L Roos**, “Age-specific education and income gradients in morbidity and mortality in a Canadian province,” *Social science & medicine*, 1997, 45 (3), 383–397.
- National Health Insurance Service**, “2008-2022 Annual NHIS Health Exam Statistics,” 2022.
- National Institute on Alcohol Abuse and Alcoholism**, *Helping Patients Who Drink Too Much, A Clinician’s Guide*, US Department of Health and Human Services, 2007.
- OECD**, “Life Expectancy at Birth,” 2025. Accessed July 21, 2025.
- Oster, Emily**, “Health recommendations and selection in health behaviors,” *American Economic Review: Insights*, 2020, 2 (2), 143–160.
- Pallesen, Anna Vera Jørring, Jochen O Mierau, Frederik Kølby Christensen, and Laust Hvas Mortensen**, “Educational and income inequalities across diseases in Denmark: a register-based cohort study,” *The Lancet Public Health*, 2024, 9 (11), e916–e924.

- Powell-Wiley, Tiffany M., Paul Poirier, Lora E. Burke et al.**, “Obesity and Cardiovascular Disease: A Scientific Statement From the American Heart Association,” *Circulation*, 2021, 143 (21).
- Ravnskov, Uffe, David M Diamond, Rokura Hama, Tomohito Hamazaki, Björn Hammar skjöld, Niamh Hynes, Malcolm Kendrick, Peter H Langsjoen, Aseem Malhotra, Luca Mascitelli et al.**, “Lack of an association or an inverse association between low-density-lipoprotein cholesterol and mortality in the elderly: a systematic review,” *BMJ open*, 2016, 6 (6), e010401.
- Schatz, Irwin J, Kamal Masaki, Katsuhiko Yano, Randi Chen, Beatriz L Rodriguez, and J David Curb**, “Cholesterol and all-cause mortality in elderly people from the Honolulu Heart Program: a cohort study,” *The lancet*, 2001, 358 (9279), 351–355.
- Shin, Dong Wook, Juhee Cho, Jae Hyun Park, and BeLong Cho**, “National General Health Screening Program in Korea: history, current status, and future direction,” *Precision and Future Medicine*, 2022, 6 (1), 9–31.
- Shin, WY, HC Kim, T Lee, D-H Jeon, KH Ha, DJ Kim, and H-J Chang**, “Combined effects of diabetes and low household income on mortality: a 12-year follow-up study of 505 677 Korean adults,” *Diabetic Medicine*, 2018, 35 (10), 1345–1354.
- Shui, Ailun, Gerard J van den Berg, Jochen Mierau, and Laura Viluma**, “Lifetime Trajectories and Drivers of Socioeconomic Health Disparities: Evidence from Longitudinal Biomarkers in the Netherlands,” 2024.
- Song, Yun-Mi and Jai Jun Byeon**, “Excess mortality from avoidable and non-avoidable causes in men of low socioeconomic status: a prospective study in Korea,” *Journal of Epidemiology & Community Health*, 2000, 54 (3), 166–172.
- Song, Yunjie, Jonathan Skinner, Julie Bynum, Jason Sutherland, John E. Wennberg, and Elliott S. Fisher**, “Regional Variations in Diagnostic Practices,” *New England Journal of Medicine*, 2010, 363 (1), 45–53.
- Statistics Korea**, “Cause of Death Statistics,” 2023.
- Wawrzyniak, Andrew J.**, “Framingham Heart Study,” in “Encyclopedia of Behavioral Medicine,” New York, NY: Springer, 2020, pp. 1–4.
- World Health Organization**, “Global recommendations on physical activity for health,” in “Global recommendations on physical activity for health” 2010, pp. 60–60.

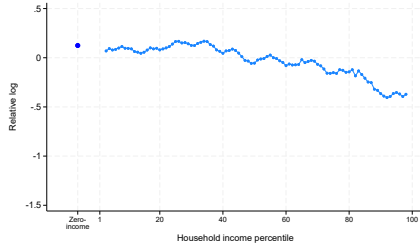
Figure 1: Biomarker-income gradients



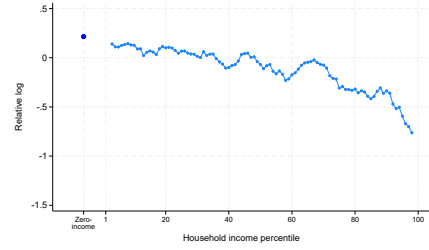
(a) high blood pressure - men



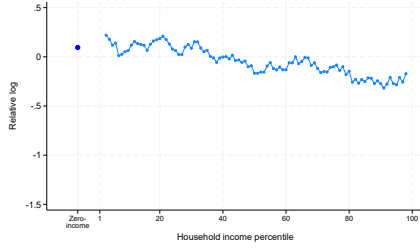
(b) high blood pressure - women



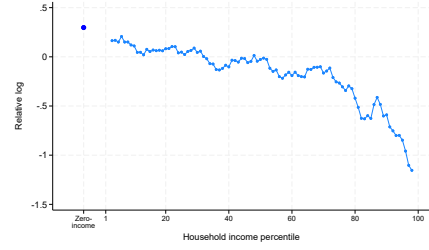
(c) high blood glucose - men



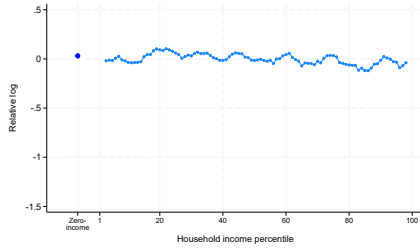
(d) high blood glucose - women



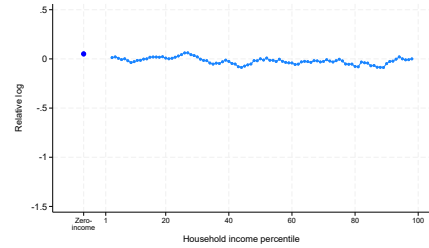
(e) high BMI - men



(f) high BMI - women



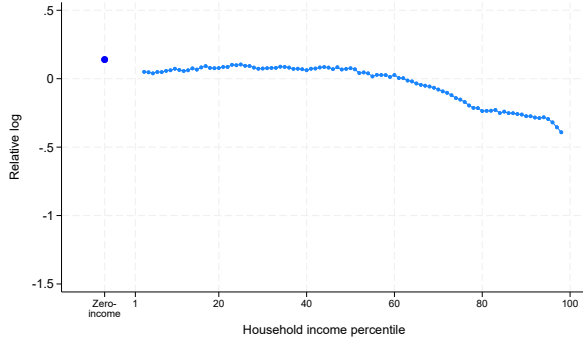
(g) high cholesterol - men



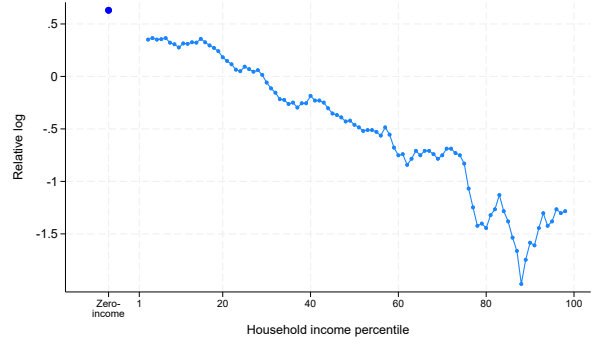
(h) high cholesterol - women

Notes: Figure shows biomarker-income gradients for 167,233 men (left panels) and 181,903 women (right panels). x-axis shows the household income percentile. y-axis reports the (natural) log of the a 5-percentile moving average of the disease rate at a given income percentile minus the log of the average disease rate across all income percentiles. 23% of men and 26% of women have zero taxable income and they also report the log of average disease rate among zero income earners minus the log of the average disease rate across all income percentiles.

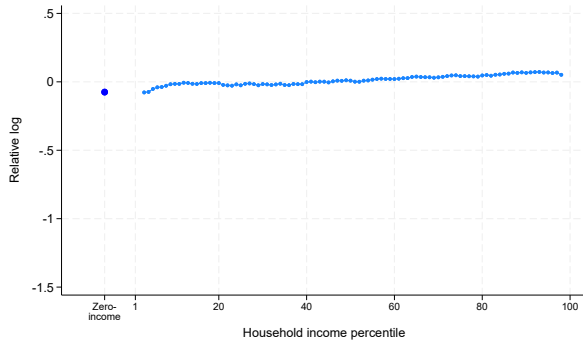
Figure 2: Behavior-income gradients



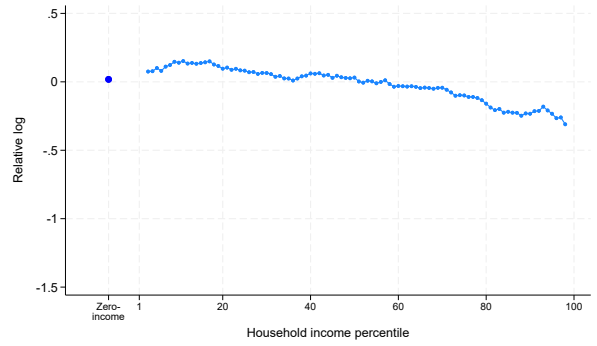
(a) smoking - men



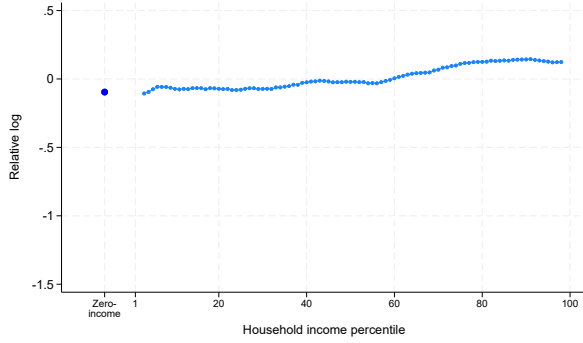
(b) smoking - women



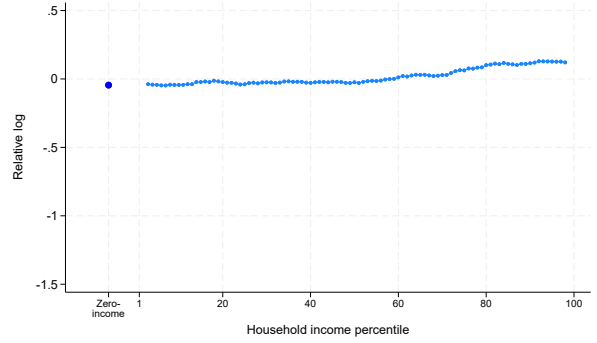
(c) drinking - men



(d) drinking - women



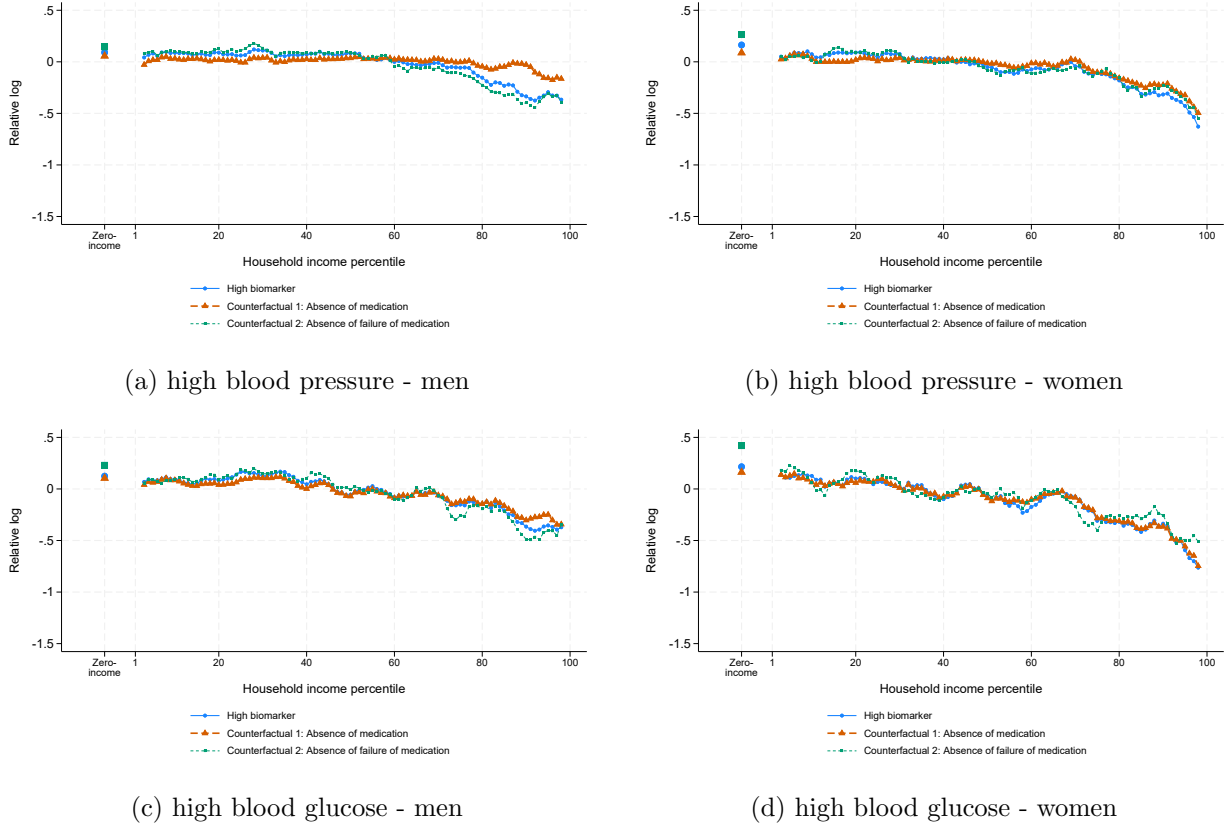
(e) exercise - men



(f) exercise - women

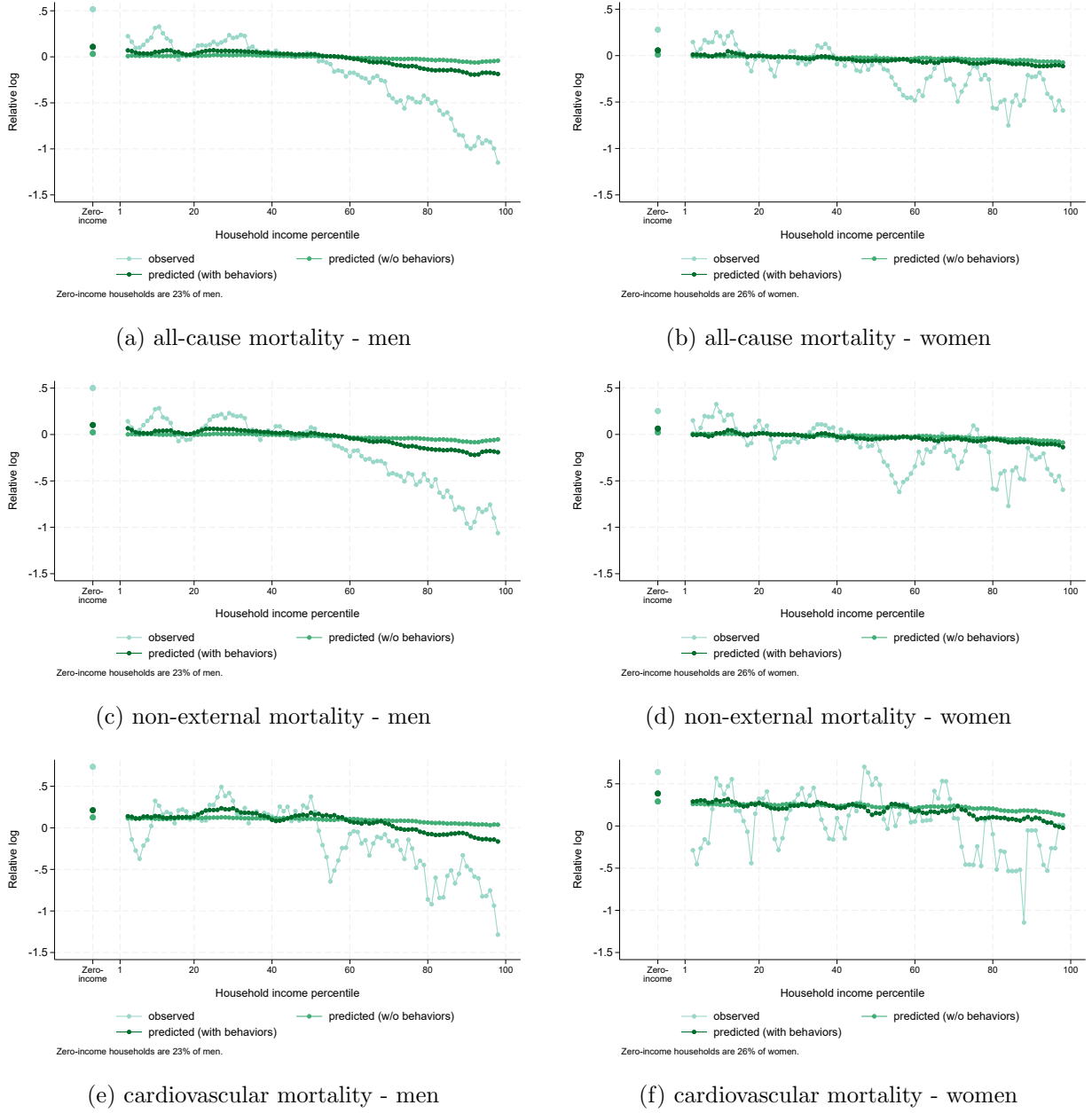
Notes: Figure shows behavior-income gradients for smoking, drinking and exercise. x-axis shows the household income percentile. y-axis reports the (natural) log of the 5-percentile moving average of the behavior rate at a given income percentile minus the log of the average rate across all income percentiles. Among men, the average rates of smoking, drinking, and exercise are 0.421, 0.679, and 0.705, respectively; among women, they are 0.039, 0.228, and 0.663. 23% of men and 26% of women have zero taxable income.

Figure 3: Biomarker-income gradients medication counterfactuals



Notes: Figure shows biomarker-income gradients for 167,233 men (left panels) and 181,903 women (right panels). x-axis shows the household income percentile. y-axis reports the (natural) log of the a 5-percentile moving average of the disease rate at a given income percentile minus the log of the average disease rate across all income percentiles. 23% of men and 26% of women have zero taxable income and they also report the log of average disease rate among zero income earners minus the log of the average disease rate across all income percentiles. We additionally plot two counterfactual gradients in the absence of medication (counterfactual 1; shown in red) and in the absence of failure in medication (counterfactual 2; shown in green).

Figure 4: Mortality-income gradients: Actual vs. Predicted



Notes: Figure shows mortality-income gradients and predicted-mortality income gradients for 167,233 men (left panels) and 181,903 women (right panels) separately for overall mortality, mortality from non-external causes, and cardiovascular mortality. x-axis shows the household income percentile. y-axis reports the (natural) log of the 5-percentile moving average of the mortality rate (or the predicted mortality rate) at a given income percentile minus the log of the average rate across all income percentiles. 23% of men and 26% of women have zero taxable income (light and dark green dots).

Table 1: Average health risks and 8-year mortality rate

	Men	Women
Biomarkers		
High Blood pressure	0.196	0.130
High Blood glucose	0.112	0.049
High BMI	0.030	0.038
High Cholesterol	0.131	0.171
Health behaviors		
Any Smoking	0.421	0.039
Any Drinking	0.679	0.228
Any Exercise	0.705	0.663
Medication		
<i>Blood pressure</i>		
Successfully controlled	0.132	0.110
Unsuccessfully controlled	0.074	0.054
Not on medication, High Blood pressure	0.122	0.076
Not on medication, Not High Blood pressure	0.672	0.760
<i>Blood glucose</i>		
Successfully controlled	0.035	0.019
Unsuccessfully controlled	0.054	0.026
Not on medication, High blood glucose	0.058	0.023
Not on medication, Not High blood glucose	0.853	0.932
8-year Mortality rate		
All-cause	0.027	0.010
Non-external	0.018	0.007
Cardiovascular	0.004	0.001

Note: Exam population (N = 167,233 men, and 181,903 women).

SUPPLEMENTARY MATERIAL

The relationship between income and health: evidence from nationwide health exams in Korea

Einav, Finkelstein, and Hong

A Setting

A.1 Health insurance coverage in Korea

Korea has had universal health insurance coverage since 1989. Since 2000, this has taken the form of a single, national, public health insurance plan (the National Health Insurance Service, or NHIS) which provides a uniform benefits package. The package includes cost-sharing for covered services, and an income-based cap on cumulative out-of-pocket services.¹⁴ Approximately one-third of total medical costs are not covered by the national insurance (either because the service is not covered or because of cost-sharing) and as a result, about three-quarters of Koreans have supplementary private insurance (Shin et al. 2022).

A.2 Health exams¹⁵

Program Structure. Korea began a national health screening program for industry workers in the early 1950s. The target population, target diseases, and health screening items have been gradually expanded ever since. Since the mid-1990s and continuing through our study period (2009-2019),¹⁶ free, biennial health exams have been available for all individuals aged 40 and over, and free, annual exams have been available for “manual” workers aged 40 and over. “Manual” workers include anyone who has a non-desk job; this includes drivers, flight attendants, chefs, doctors, pharmacists, correspondents, singers, actors, producers etc.

The program is motivated by a general desire to improve population health, but has been criticized for, among other things, lack of clear rationales for the chosen screening items and any evaluation of the program’s impact. This prompted a redesign of the health screening in 2009 which added various health measures and revised the questionnaire on health behaviors. We analyze health exams under this redesigned approach (i.e., 2009-2012). The chief focus is on cardiovascular and cerebrovascular disease and its risk factors (high blood pressure, high blood glucose, high BMI, high cholesterol, and smoking). In addition, individuals are screened for depression every 10 years, and those over 65 are screened regularly for cognitive function.

Individuals born in even-numbered years are offered the exam in even years, and individuals born in odd-numbered years are offered the exam in odd-numbered years (National Health Insurance Service 2022b). The NHIS identifies the individuals eligible for the health exam and sends them an invitation. Subjects are instructed to fast overnight and refrain from working night shifts or drinking

¹⁴For the non-indigent, cost-sharing ranges from 20% to 60% depending on the service and whether the region is urban or rural. For example, most inpatient services have a 20% coinsurance, while coinsurance for outpatient services in urban areas is 30% for clinics, 40% for urban pharmacy services, and 35% for rural pharmacy services.

¹⁵Information in this section draws heavily on Shin et al. (2022) and Kang (2022).

¹⁶In 2019, the provision of free, biennial general health screenings was extended to individuals aged 20 and over. Prior to that, only employees or self-employed individuals who paid insurance premia were eligible regardless of age.

alcohol the night before the exam. During the health exam, they are given a questionnaire about their past medical history and their health behaviors (including smoking, alcohol and exercise), and then undergo a series of measurements (including blood pressure, waist circumference, weight and height), and have blood samples taken for laboratory testing.¹⁷ Within about two weeks of the screening, the participants are mailed a report of the results of the test which include not only detailed results from the examination, but a general health classification (no high biomarker, boundary – e.g., elevated blood pressure or elevated blood glucose – or suspected disease) as well as a “health traffic light” used to visualize the patient’s risk for cardiovascular disease (low, medium, or high). The report also contains recommendations regarding health behaviors as well as referrals for e.g., smoking cessation medication or an alcohol rehabilitation clinic. Individuals with abnormal results (such as blood pressure outside of the normal range) are recommended to get confirmatory tests for which the typical co-payments are waived.

The NHIS encourages participation in the health exam through public campaigns (e.g., media and outdoor advertising). The exam may be done by any certified medical provider (public or private).

Enforcement and compliance. It is useful to distinguish between three types of individuals aged 40 and over in terms of the free health exam regime they face: (1) manual employees: legally required to have a health exam every year; (2) non-manual employees: legally required to have a health exam every other year; and (3) not employed (i.e., unemployed, self-employed, and out of the labor force): not legally required to complete the health exam.

The health exams are legally mandatory for all employees (National Assembly of South Korea 2023). If an employee does not comply with her mandatory medical check-up, her employer has to pay a fine, which increases from 100,000 KRW per person (approximately \$70 US) for a first-time violation, to 200,000 KRW (second-time), to 300,000 KRW (third-time), up to a maximum penalty of 10 million KRW (about \$7,000 US).¹⁸ In addition, starting in 2005, if the employee has not gotten their mandatory medical check-up and is diagnosed with cancer, they are not eligible for the special benefit from the Ministry of Health and Welfare of a cash transfer of up to 2 million KRW per year (for up to 3 years) to reimburse out-of-pocket medical expenses related to their cancer treatment (Ministry of Health and Welfare 2023).

Those not employed are not legally required to complete their (free) health exam, but they are still automatically assigned to do one and receive all of the notifications (mails, emails, and text messages) that employees do reminding them to get their health exam; failure to comply also results in the loss of the special benefit in the event that they are diagnosed with cancer.

Compliance with the recommended health exam schedule has been gradually increasing over the 2009-2019 period from about three-fifths to about three-quarters of the eligible population (National Health Insurance Service 2022a).¹⁹

¹⁷Shin et al. (2022) Table 2 provides a useful summary of the measures collected and how they have varied with age and over time.

¹⁸According to the Occupational Safety and Health Act Article 175 and the Enforcement Decree of the Occupational Safety and Health Act Article 119, if the employer can demonstrate that she did her best to enforce the required exam, then the employee pays the fine. In this case, the fine starts at 50,000 KRW, and increases to 100,000 KRW (second-time) and 150,000 KRW (third-time), with a maximum penalty of 3 million KRW. Please see Cho and Lee (2011) and National Assembly of South Korea (2023) for more detail.

¹⁹In 2008, on average 61.2% of the eligible population complied with NHIS’s recommended health exam schedule. This share varied by age from 48.7% (age 75-79) to 69% (age 55-59). By 2019, 77.9% of the eligible people complied

A.3 Income tax system and labor force participation

Korea has a progressive income tax system, with the individual as the unit of taxation (Lim and Hyun 2009; Hong et al. 2024). Korea has two types of personal income taxation: withholding taxation and global income taxation. Most incomes are taxed when they are paid (withholding taxation).²⁰ Some incomes above certain thresholds, however, are taxed again through global income taxation. This global income taxation requires taxpayers to self-assess and report their income. Less than 20% of taxpayers in Korea are subject to this additional global taxation (National Tax Service 2023; Hong et al. 2024).

Labor force participation in Korea is distinctive relative to other OECD countries in two respects. First, labor force participation among older Koreans is among the highest in the OECD. In 2017, for example, almost three-quarters of individuals 55-59 were employed, as were one-third of those aged 70-74 (OECD 2018). Second, female labor force participation rates are much lower (about 25 percentage points) than for men (Lee 2017) with this gap driven almost entirely by married women with children relative to their husbands (Cho and Cho 2015; Dynan et al. 2022; Lee et al. 2008).

B Data

We have individual-level data from the Korean National Health Insurance Service (NHIS) on health exams linked to individual-level tax records from the National Tax Service and mortality (if any) from Statistics Korea. The income tax data include information on wage, business, dividend, interest, pension and other income; people who do not file taxes in a given year appear in the data that year with zero taxable income. The mortality data include all-cause mortality and cause-specific mortality.

For this project we use individual-level tax records from 2006-2008, health exams from 2009-2012, and mortality through 2019. We use data from a random 8% of the population.²¹

B.1 Study population

To construct our analysis sample, we start with all individuals born between 1955 and 1962 who are still alive in 2008. We refer to this sample as our **“target population”**; it includes 249,325 men and 243,964 women.

We observe at least one health exam over the 2009-2012 period for approximately two-thirds of men (167,233) and three-quarters of women (181,903) in the target population. We refer to the subset of the target population for whom we observe at least one health exam between 2009 and 2012 as the **“exam population”**. If they have more than one health exam during that period, we use the first health exam that we observe. Individuals are therefore 47 to 57 at the time of their first health exam. Conditional on having a health exam, the health measures are present for 99.9%

with the recommended exam schedule, with the share varying from 69.7% (age 75-79) to 81.2% (age 65-69).

²⁰This includes wages, (some) business income, interest, dividend, pension, and other incomes.

²¹More specifically, the NHIS selected 4% of the population (with replacement) in 2002 and 2007 with stratified random sampling (sex and age of household head). For any individual who is selected, the NHIS links the individual to their current household members; for any selected individuals and household members, NHIS constructs a panel data set of them from 2002-2022 on their income tax data, mortality data, and health exam if available.

of the sample.²² We refer to the complement subset of the target population as the “**non-exam population**”; it contains 82,092 men and 62,061 women.

B.2 Variable definitions

Income and income percentiles. For each individual, we compute their household income as the sum of the person’s taxable income and her spouse’s taxable income if they have a spouse in their household. Taxable income includes labor income, dividend income, interest income, business income, pension income, and other income. Our approach to defining income percentiles follows Chetty et al. (2016) closely. After excluding individuals with zero taxable income, we compute each individual’s percentile rank for their yearly household income relative to the household income of all others in their birth cohort and sex in our full population. To reduce noise, we then compute the average household income percentile over 2006 through 2008.²³ Within an individual, income percentiles were highly correlated across years (Table A7).

Figure A15 shows the log income distribution by income percentile, separately for men and women; as is typical for income distributions (Champernowne 1953; Pareto 1896), it is a bit less compressed at both tails.

As noted, about one-quarter of individuals have zero household income in all three years. We report the outcomes for these individuals separately. They include people with no earnings who are receiving cash welfare through the Basic Livelihood Program (about 3% of the population (You et al. 2021)); daily wage workers, whose tax data were not shared with the NHIS (about 8% of the population aged 40-59 (National Tax Service 2023; You et al. 2021)); and farmers who do not pay income tax by Income Tax Act Article 12 (also about 7% of the population age 45-54) (Statistics Korea 2009b; National Assembly of South Korea 2025). In addition, some self-employed individuals may not report earnings (Sung 2008; National Tax Service 2023).²⁴ These individuals are likely low income.²⁵

Type of worker. The NHIS systematically collects data on individuals’ employment status, industry of employment, and occupational categories, as health insurance premia are determined by employment classification. Our sample is categorized into one of three employment statuses: manual employee, non-manual employee, or not employed. Employment classification was determined using the modal value across observations from 2009 to 2012 for each individual.

Mortality. We observe all-cause and cause-specific mortality (if any) through 2019. The cause of death is classified based on the 10th International Classification of Diseases (ICD-10). We focus on three mortality measures: all-cause mortality, the subset of all-cause mortality that is from non-external causes,²⁶ and cardiovascular disease mortality (causes from any cardiovascular diseases;

²²We exclude the few individuals who are missing any health measures from the exam population.

²³99.7% of our sample shows up in the tax data for all three years. For the 0.2% who show up for 2 years (or the 0.1% who show up for one year) we define their income percentile based on 2 (or 1) measures.

²⁴Retirees without pension benefits would also not file taxes, although given the high rates of labor force participation among Koreans in this age range, this seems less likely.

²⁵The majority of daily wage workers are located at the bottom 20 percent of the income distribution of the full population (Kim 2017; Statistics Korea 2009a). Most farm households earn less than the mean income of urban households (Institute 2004) and fall in the lowest income decile (Statistics Korea 2007).

²⁶External causes include accidents (V01-X29, X50-59), suicides (X60-X84), homicides (X85-Y09), drug overdose (X40-49), natural disasters (X30-39), and sequelae of other external causes (Y10-Y98).

I00-I99).

For our analyses that are limited to the exam population, we measure mortality for the eight years following their exam, which is the longest period before 2020 over which we can observe mortality for everyone. For some analyses, however, we will want to compare mortality between the exam population and the target population; for these analyses we measure 8-year mortality for both the target and the exam population based on the first year we could have observed them in the health exam data after they turn 50.²⁷

Health measures. The health exam data include data from the questionnaires (regarding self-reported health behaviors and prior medical history), and biomarkers based on physical measurements, blood samples, and screens for depression (every 10 years) and cognition (for those 65 and over). For this paper, we focus on four key biomarkers of physical health that are viewed as key predictors of cardiovascular and cerebrovascular disease: high blood pressure, high blood glucose, high BMI and high cholesterol ([American Diabetes Association Professional Practice Committee 2024](#); [Centers for Disease Control and Prevention 2024](#); [O’Donnell and Elosua 2008](#); [Powell-Wiley et al. 2021](#); [Wawrzyniak 2020](#)).

In each case we observe the underlying biomarkers (diastolic and systolic blood pressure, fasting blood glucose, height and weight, and total cholesterol). We convert each to binary measures of high biomarkers using standard thresholds. Specifically, we code high blood pressure as systolic blood pressure above 139 mmHg or diastolic blood pressure above 89 mmHg, high BMI as body mass index (BMI) of 30 kg/m² or above, high blood glucose as fasting blood glucose higher than 125 mg/dL, and high cholesterol as total cholesterol of 240 mg/dL or above. We likewise create binary versions of elevated (but not yet “high”) biomarkers using standard thresholds: we code elevated blood pressure as systolic blood pressure between 120 mmHg and 139 mmHg or diastolic blood pressure between 80 mmHg and 89 mmHg, elevated blood glucose as fasting blood glucose between 100 mg/dL and 125 mg/dL, elevated BMI as BMI over 25 kg/m² and less than 30 kg/m², and elevated cholesterol as total cholesterol over 200 mg/dL and less than 240 mg/dL.²⁸

Medication. For medication, we observe whether participants report currently taking prescribed medications for hypertension and diabetes. This information is self-reported by participants during the health exam interview. Using the current medication and biomarker information, we further define “successfully controlled” high blood pressure (and analogously “successfully controlled” high blood sugar) if, for the relevant biomarker, the individual reports being on medication and their biomarker is in the clinically normal range. Likewise we label the person as having an “unsuccessfully controlled” biomarker if they report being on medication but their biomarker is out of the clinically normal range.

Health behaviors At the health exam, a doctor asks every individual about their smoking, drinking, and exercise behavior, including status and intensity. Two questions are asked for each

²⁷Thus, for birth cohorts 1955-1959, we measure 8-year mortality from 2009 — i.e., died in 2009–2016 or earlier. For birth cohort 1960, however we measure 8-year mortality from 2010 (the year they turn 50) — i.e., died in 2010–2017; likewise mortality for birth cohorts 1961 and 1962 is measured starting in 2011 and 2012 respectively.

²⁸Our high blood pressure and elevated blood pressure cut-offs are based on [Ajay et al. \(2007\)](#). [American Diabetes Association \(2024\)](#) provides cutoffs for high blood glucose and elevated blood glucose. For high BMI and elevated BMI, we are using [Centers for Disease Control and Prevention \(2024\)](#). Lastly, for high cholesterol and elevated cholesterol, we use cutoffs from [Johns Hopkins Medicine \(2024\)](#).

measure.

For smoking, the individual is first asked “have you ever smoked?” Options are no, yes but not currently, and yes and currently smoking. Second, if they report currently smoking, they are also asked “How many cigarettes do you smoke a day?” Appendix Figure A16 shows the distribution of answers to each question in our analysis sample. Based on the answers, we define a binary indicator for current smokers, and a binary indicator for a heavy smoker as someone who currently smokes 20 or more cigarettes a day (Hirayama 1981).

For drinking, again two questions are asked. First, “how many days do you drink a week?” And second, “how many glasses do you drink at one time?” Multiplying the two answers gives weekly consumption. Appendix Figures A17 and A18 show the distribution of answers to each question in our analysis sample, and Appendix Figure A19 shows the distribution of weekly consumption. Based on this information, we define a binary indicator for drinkers based on whether the individual consumes any drinks per week. We also code a binary indicator for heavy drinkers, defined as consuming 15 or more drinks a week for men and 8 or more drinks a week for women (National Institute on Alcohol Abuse and Alcoholism 2007).

Finally for exercise, individuals are asked “How many days do you do a vigorous workout more than 20 minutes a week?” and “How many days do you do a moderate workout more than 30 minutes a week?” Multiplying each answer by its duration threshold gives weekly minutes of vigorous and moderate exercise. Appendix Figures A20 and A21 show the distribution of answers to the questions. Based on this calculation, we define a binary indicator for any exercise if the individual reports any exercise each week, and a binary indicator for sufficient exercise if the individual reports doing at least 150 minutes of moderate exercise or 75 minutes of vigorous exercise a week (World Health Organization 2010).

B.3 Statistical analyses

Gradients. We report mean health or mortality outcomes by household income percentile, separately for men and women. For our main analyses, we report gradients of outcomes in relative terms to make them more comparable across populations and outcomes. Specifically, we report mean outcomes in relative log points; that is, we report the (natural) log of the outcome measure at a given percentile of the income distribution minus the log of the average outcome across all income percentiles. As a result, the health measure that is reported at a given income percentile is essentially a multiplicative factor of the population average. We smooth this relative outcome at each income percentile by taking a moving average over 5 percentiles (the one shown and two on either side). We report results from the 3rd to the 98th income percentile.

In addition, for completeness, for the gradients we show in the main text, we also report the “raw” gradients, i.e., the (un-smoothed) absolute rates at each household income percentile.

Predicted mortality. To generate predicted mortality we use information on the biomarker status j which we define as a series of categorical variables by $h^j \in \{\text{high, elevated, normal}\}$ for four measures of j : blood pressure, blood glucose, BMI, and total cholesterol. We therefore characterize overall health status h to be the full set of interactions across the four biomarkers and their three possible values, and let $h_i \in \{1, \dots, 81\}$ denote which of these 81 ($=3^4$) cells individual i falls into, with a coefficient γ_h for each cell. We estimate the following equation separately for men and for women:

$$y_i = \sum_{h=1}^{81} \gamma_h \mathbf{1}\{h_i = h\} + \epsilon_i, \quad (1)$$

where y_i is an indicator for whether the individual died during the subsequent 8 years, or died from a specific cause (e.g., non external mortality or cardiovascular mortality) during the subsequent 8 years. Based on this regression we can form a predicted mortality measure $\hat{y}_i$ for each individual in the sample.

When we expand to also include the three health behaviors (drinking, smoking and exercise) each of which can also take on three possible values (heavy, any, none for smoking and drinking, and none, any, sufficient for exercise) h_i instead indexes 2,187 ($=3^7$) cells and the sum in equation (1) runs to 2,187 rather than 81.

C Results

C.1 Summary statistics

Demographics. Table A8 shows summary statistics by sex for the target population and separately for the exam population and the non-exam population. Average age in 2009 is 50 for the target population and for both sub-populations.²⁹ Not surprisingly given that exams are mandatory for manual and non-manual workers but not for the non-employed, exam participants are more likely to be employed than the target population. Among exam participants, 17% of women and one-third of men are manual workers, who are required to have an annual exam. The remainder are non-manual workers (about 14% of women and 45% of men), or not employed (about 70% of women and about 20% of men). Each worker type can be found throughout the income distribution (Figure A22).

About one-quarter of the target population has zero income in all three years; this number is lower in the exam population. Recall that income percentiles are averaged over three years (2006-2008); Table A7 indicates that average income percentiles are highly correlated over time (with correlations between adjacent years of 0.9 or higher).

Mortality. All-cause mortality is quite low in Korea. Table A2 reports official statistics on annual mortality in 2006 in Korea and in the United States. At ages 45-54, annual mortality rates for men are about 9% lower than in the US and annual mortality rates for women are over 50% lower. The share of mortality due to external causes falls rapidly with age, and is similar within an age $\times$ sex group between the US and Korea. The share of mortality due to cardiovascular disease is lower in Korea than in the US.

Table A5 shows 8-year mortality rates (measured from the first year we could have observed them participating in a 2009-2012 health exam data after they turned 50) by sex separately for the target population, exam population, and the non-exam population. In the target population, the all-cause 8-year mortality rate is 4.5% for men and 1.6% for women. Of this, mortality from non-external causes accounts for about two-thirds of male mortality and three-quarters of female mortality, while mortality from cardiovascular disease comprises about 15% of male mortality and 13% of female mortality. Mortality rates are substantially lower for exam participants. Specifically,

²⁹The average age at the time of the health exam is 51 for men and for women.

compared to the target population, 8-year mortality rates for the exam participants were about 40-45% lower for men (0.026 compared to 0.045) and for women (0.009 compared to 0.016). However, the share of deaths due to non-external causes or due to cardiovascular disease is quite similar for the exam population and the target population.

Exam population: ‘high biomarker’ rates. Table 1 shows the average rates of ‘high biomarkers’ for exam participants, as well as their eight year mortality rate following their health exam. On average, 20% of men have high blood pressure, 11% have high blood glucose, 3% have high BMI and 13% have high cholesterol; for women the comparable numbers are 13%, 5%, 4% and 17%, respectively. 8-year all-cause mortality rates are 2.7% for men and 1.0% for women; mortality rates from non-external causes are 1.8% for men and 0.7% for women, and mortality rates for cardiovascular disease are 0.4% for men and 0.1% for women.

Exam population: Relationship between ‘high biomarker’ and mortality. Table A3 shows the bi-variate relationship between each ‘high biomarker’ and 8-year mortality — all-cause, for non-external causes, and for cardiovascular disease — separately for men and women. The relationships are generally as expected. All three mortality measures are increasing in high blood pressure, high blood glucose, and high BMI for men, and mostly for women too (with the exception of cardiovascular disease mortality and high BMI for women), with many of these relationships statistically significant. The relationship between high cholesterol and mortality however tends to be ‘wrong signed’.

To get a sense of the predictive power of the fully interacted set of health measures from estimating equation (1), Table A4 shows the distribution of predicted mortality rates. The standard deviation of the predicted mortality rate is non-trivial. For example, the standard deviation of predicted all-cause mortality is about 40 percent of the mean of observed all-cause mortality for both men and women; these numbers are even higher for cardiovascular mortality.

C.2 Disease and mortality income gradients, exam participants

Figure A2 shows the raw disease-income gradients corresponding to the figures in the main text and Figure A9 and Figure A10 show the raw predicted mortality income gradients corresponding to the figures in the main text.

C.3 Robustness

Variation in Age by Income Percentile. Figure A23 shows that the average age at first health exam varies slightly by income. However, Figure A3 shows that all of the disease-income gradients in the main text look similar if the outcomes are first residualized for dummies for age at the time of the health exam.

Variation in Exam Participation Rates by Income Percentile. The exam population is not a random sub-sample of the target population. They are disproportionately female (74.6% of women are exam participants, compared to 67.1% of men). They are also healthier (recall Table A5).

Exam participation rates also vary substantially – albeit non-monotonically – with income (Figure A1).³⁰ This is particularly pronounced for men where participation rates rise steadily from about 55% at the bottom of the income distribution to about 90% at the 90th percentile, before plunging down to about 50% over the top decile. The same non-monotonic pattern holds for women, but the gradient is much shallower: participation rates rise from about 70% at the bottom of the income distribution to about 80% at the 85th percentile, before falling to about 50% over the top decile.

If selection into the health exam population differs by income, this could bias our measurement of health gradients with respect to income. In particular, any such selection effects on the gradient would likely lead us to under-estimate (in absolute value) the slope of the negative disease-income gradients up to about the 85th percentile – since exam participants are healthier and participation is increasing with income – and then over-estimate it after that for the reverse reason.

In practice, there are several pieces of evidence that suggest our gradient estimates are unlikely to be substantially affected by such selection. First, the pattern of fairly monotone disease-income gradients, with the slope steepening at above median income levels (except for cholesterol where it is flat throughout), is very different from the non-monotonic participation rate patterns by income. Second, disease-income gradients are larger for women than men, but the compliance-income gradient is much shallower for women than for men.

Third, we can gauge the likely extent of any selection bias by looking at mortality gradients since, unlike the health measures, mortality is observed for both exam participants and non-participants. Figure A11 indicates that the slope of the mortality gradient is very similar for the exam population and the target population, which is suggestive of the disease-income gradient also being similar.

And finally, we examine gradients separately for those who participate in the health exam once vs twice over the 2009-2012 health exam window. Figure A13 shows that the disease-income gradients look very similar for these two groups. Specifically, we exclude manual workers so that everyone is supposed to get an exam twice during that window.

Appendix References

Ajay, Vamadevan S., Dorairaj Prabhakaran, and Kolli Srinath Reddy, “Chapter 40 - Electrolyte Intake and Human Hypertension,” in Gregory Y.H. Lip and John E. Hall, eds., *Comprehensive Hypertension*, Mosby, 2007, pp. 477–482.

American Diabetes Association, “Understanding Diabetes Diagnosis,” <https://diabetes.org/about-diabetes/diagnosis> 2024. Accessed: 2024-12-02.

American Diabetes Association Professional Practice Committee, “10. Cardiovascular Disease and Risk Management: Standards of Care in Diabetes—2025,” *Diabetes Care*, 12 2024, 48 (Supplement 1), S207–S238.

belong Cho and Cheol Min Lee, “Current situation of national health screening systems in Korea,” *J Korean Med Assoc*, 2011, 54 (7), 666–669.

Centers for Disease Control and Prevention, “National Vital Statistics System, Mortality 2005-2006 on CDC WONDER On-line Database,” 2010. Accessed: 2025-07-11.

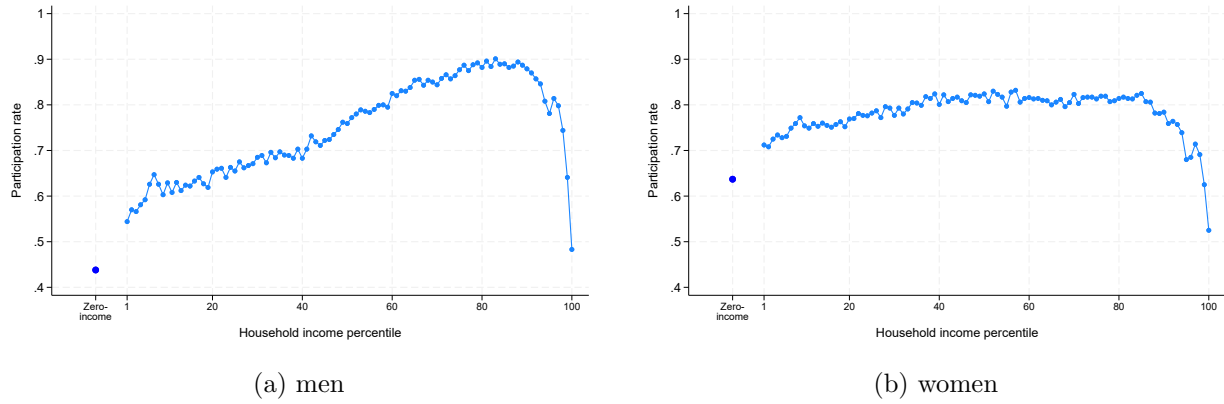
³⁰We define the exam participation rate as the share of the target population that is in the exam population; the exam participation rate is higher than the exam compliance rate since individuals have 2 or 4 recommended health exams during our 2009-2012 exam window.

- , “Adult BMI Categories,” <https://www.cdc.gov/bmi/adult-calculator/bmi-categories.html> 2024. Accessed: 2024-12-02.
- Champernowne, David G**, “A model of income distribution,” *The Economic Journal*, 1953, 63 (250), 318–351.
- Chetty, Raj, Michael Stepner, Sarah Abraham, Shelby Lin, Benjamin Scuderi, Nicholas Turner, Augustin Bergeron, and David Cutler**, “The association between income and life expectancy in the United States, 2001–2014,” *Jama*, 2016, 315 (16), 1750–1766.
- Cho, Donghun and Joonmo Cho**, “Over-heated education and lower labor market participation of Korean females in other OECD countries,” in “Women’s Studies International Forum,” Vol. 48 Elsevier 2015, pp. 1–8.
- Dynan, Karen E, Jacob F Kirkegaard, and Anna Stansbury**, “Why gender disparities persist in South Korea’s labor market,” *Peterson Institute for International Economics Working Paper*, 2022, pp. 22–11.
- Hirayama, Takeshi**, “Non-smoking wives of heavy smokers have a higher risk of lung cancer: a study from Japan,” *Br Med J (Clin Res Ed)*, 1981, 282 (6259), 183–185.
- Hong, Sehyun, Nak-Nyeon Kim, Zhexun Mo, and Li Yang**, “Income Inequality in South Korea, 1933–2022: Evidence from Distributional National Accounts,” *World Inequality Lab Working Paper 2024/03*, 2024.
- Institute, Korea Rural Economic**, “A Study on the Releasing Income Gap between Urban and Rural Areas through the Balanced Regional Development Strategies,” 2004.
- Johns Hopkins Medicine**, “Lipid Panel,” <https://www.hopkinsmedicine.org/health/treatment-tests-and-therapies/lipid-panel> 2024. Accessed: 2024-12-02.
- Kang, Hee-Taik**, “Current status of the national health screening programs in South Korea,” *Korean Journal of Family Medicine*, 2022, 43 (3), 168.
- Kim, Jaejin**, “Daily Wage Worker 8 million Era: Issues and Revision of Taxation,” *Public Finance Forum*, 2017, 257, 6–37.
- Korea Centers for Disease Control and Prevention**, “Korea National Health and Nutrition Examination Survey (KNHANES IV–V), 2009–2012,” 2009–2012.
- Lee, Bun Song, Soomyung Jang, and Jayanta Sarkar**, “Women’s labor force participation and marriage: The case of Korea,” *Journal of Asian Economics*, 2008, 19 (2), 138–154.
- Lee, Jungmin**, “The labor market in South Korea, 2000–2016,” *IZA World of Labor*, 2017.
- Lim, Byung-In and Jin Kwon Hyun**, “What makes the income tax system so progressive?—the case of Korea,” *Applied Economics Letters*, 2009, 16 (7), 683–687.
- Ministry of Health and Welfare**, “Special Assistance for Cancer Patients,” 2023.
- National Assembly of South Korea**, “Occupational Safety and Health Act,” 2023. https://elaw.klri.re.kr/eng_mobile/viewer.do?hseq=63609&type=sogan&key=6.
- , “Income Tax Act,” 2025. <https://law.go.kr/LSW/lsInfoP.do?lsiSeq=199742&viewCls=engLsInfoR&urlMode=engLsInfoR#0000>.
- National Center for Health Statistics**, “National Health and Nutrition Examination Survey Data, 2009–2012,” 2009–2012.
- National Health Insurance Service**, “2008–2022 Annual NHIS Health Exam Statistics,” 2022.
- , “The National Screening Program Handbook,” 2022.
- National Institute on Alcohol Abuse and Alcoholism**, *Helping Patients Who Drink Too Much, A Clinician’s Guide*, US Department of Health and Human Services, 2007.

- National Tax Service**, “2023 Statistics Yearbook of National Tax,” 2023.
- O’Donnell, Christopher J and Roberto Elosua**, “Cardiovascular risk factors. Insights from framingham heart study,” *Revista Española de Cardiología (English Edition)*, 2008, 61 (3), 299–310.
- OECD**, *Working Better with Age*, OECD: Paris, France, 2018.
- Pareto, Vilfredo**, *Cours d’économie politique*, Vol. 1, Librairie Droz, 1896.
- Powell-Wiley, Tiffany M, Paul Poirier, Lora E Burke, Jean-Pierre Després, Penny Gordon-Larsen, Carl J Lavie, Scott A Lear, Chiadi E Ndumele, Ian J Neeland, Prashanthan Sanders et al.**, “Obesity and cardiovascular disease: a scientific statement from the American Heart Association,” *Circulation*, 2021, 143 (21), e984–e1010.
- Shin, Dong Wook, Juhee Cho, Jae Hyun Park, and BeLong Cho**, “National General Health Screening Program in Korea: history, current status, and future direction,” *Precision and Future Medicine*, 2022, 6 (1), 9–31.
- Statistics Korea**, “2006 Household Income and Expenditure Survey,” 2007.
- , “2008 Household Income and Expenditure Survey,” 2009.
- , “Census of Agriculture, Forestry and Fisheries,” 2009.
- , “Cause of Death Statistics,” 2023.
- Sung, Myungjae**, “Estimation of Business Income Reporting Rate and Related Tax Evasion,” *Study of Public Finance*, 2008, 1 (3), 155–188.
- Wawrzyniak, Andrew J.**, *Framingham Heart Study*, New York, NY: Springer New York,
- World Health Organization**, “Global recommendations on physical activity for health,” in “Global recommendations on physical activity for health” 2010, pp. 60–60.
- You, Jongsung, Minhye Kim, Seungyeon Kim, and Soojin You**, “The Merits and Limitations of the NHIS Big Data in the Study of Income Distribution,” *Korea Social Policy*, 2021, 28 (3), 75–105.

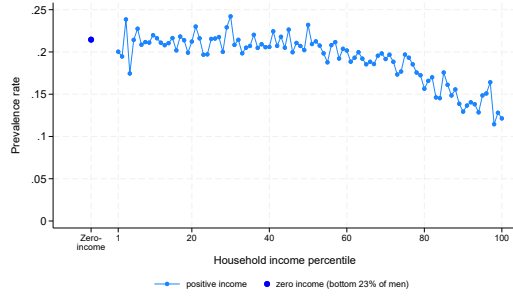
Appendix Figures

Figure A1: Mean participation rate by income percentile (2009-2012)

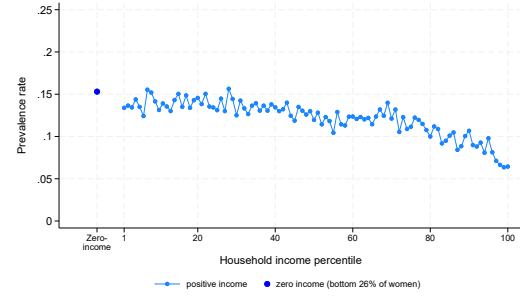


Notes: Participation rate is the share of the target population that is in the exam population. Average participation rates are 67.1% (men) and 74.6% (women). Sample size for target population is men: 249,325 and women: 243,964. 23% of men and 26% of women have zero taxable income.

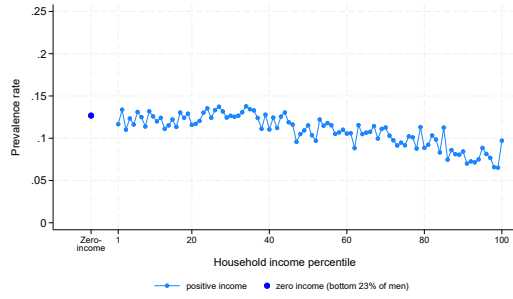
Figure A2: Biomarker-income gradients



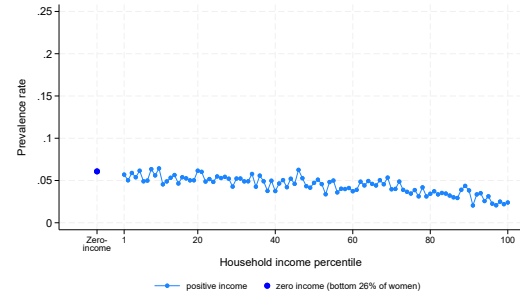
(a) high blood pressure - men



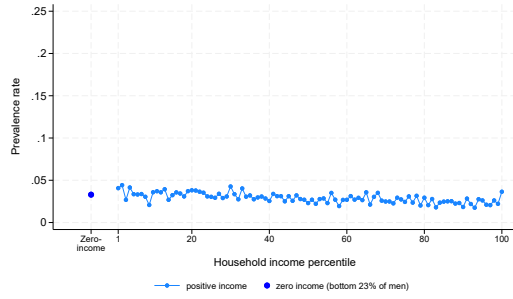
(b) high blood pressure - women



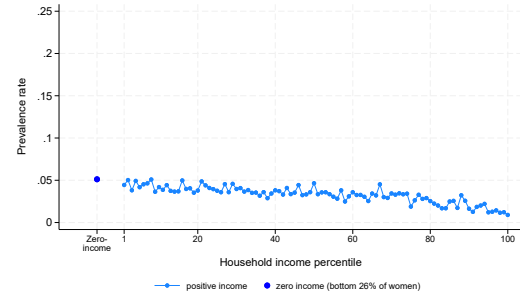
(c) high blood glucose - men



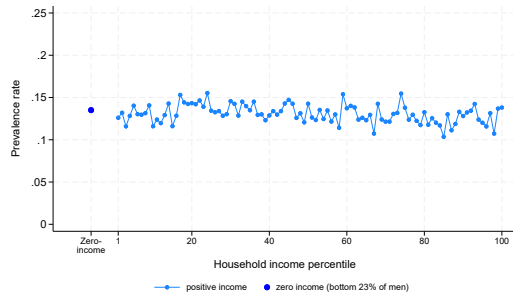
(d) high blood glucose - women



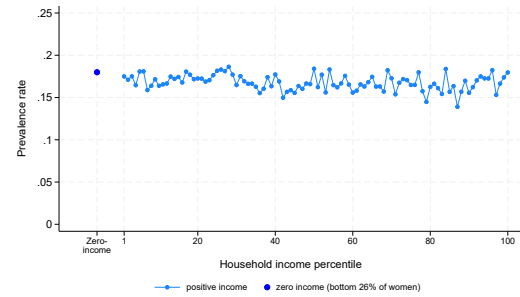
(e) high BMI - men



(f) high BMI - women



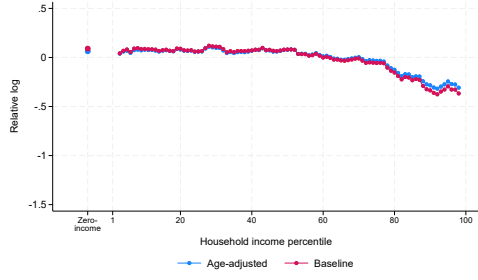
(g) high cholesterol - men



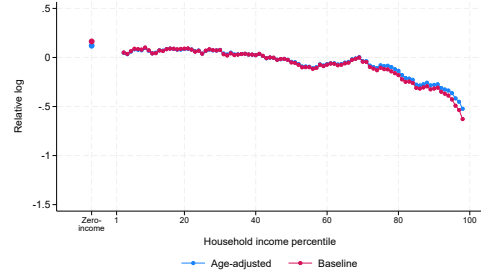
(h) high cholesterol - women

Notes: Sample is limited to exam population (men: 167,233 and women: 181,903.) x-axis shows the household income percentile. y-axis reports the average rate at each income percentile. 23% of men and 26% of women have zero taxable income.

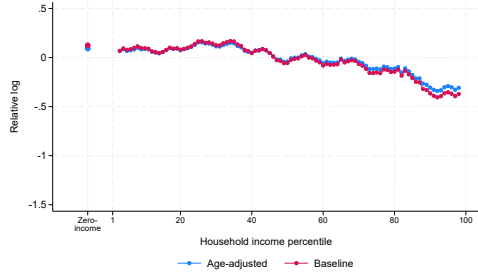
Figure A3: Age sensitivity check - biomarker-income gradient



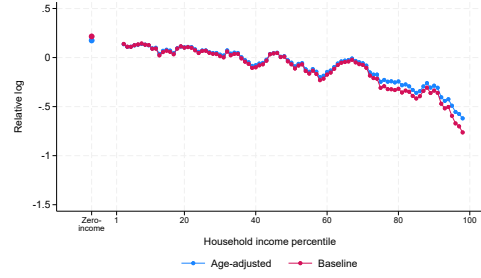
(a) high blood pressure - men



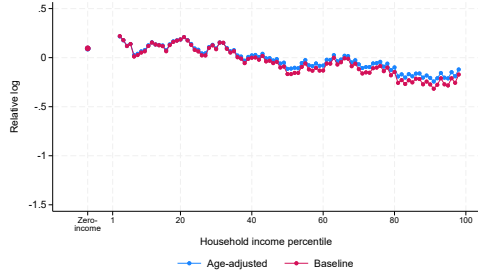
(b) high blood pressure - women



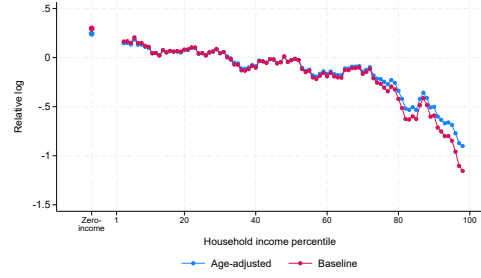
(c) high blood glucose - men



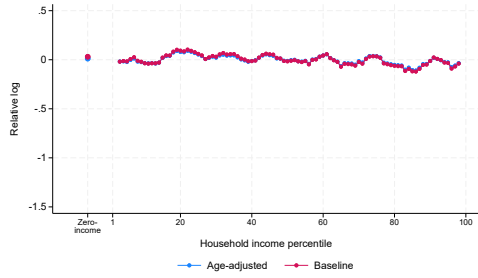
(d) high blood glucose - women



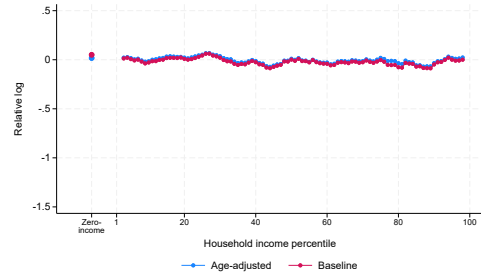
(e) high BMI - men



(f) high BMI - women



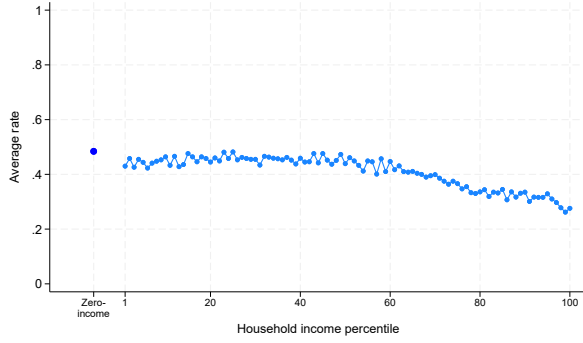
(g) high cholesterol - men



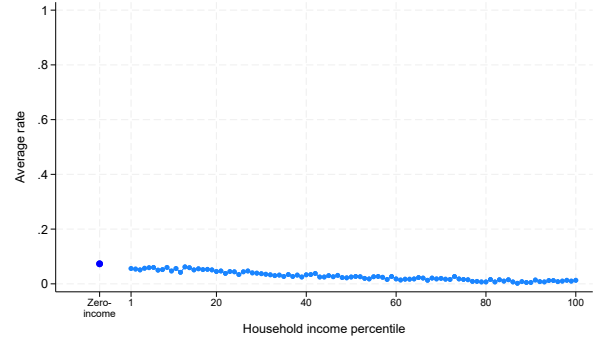
(h) high cholesterol - women

Notes: Sample is limited to the exam population (men: 167,233 and women: 181,903). x-axis shows the household income percentile. y-axis reports the (natural) log of the 5-percentile moving average of high biomarker rate at a given income percentile minus the log of the average high biomarker rate across all income percentiles. For the age-adjusted results, we first residualize the probability of having the high biomarker for age (using single year of age indicators) and then analyze the average residuals by income percentile. 23% of men and 26% of women have zero taxable income.

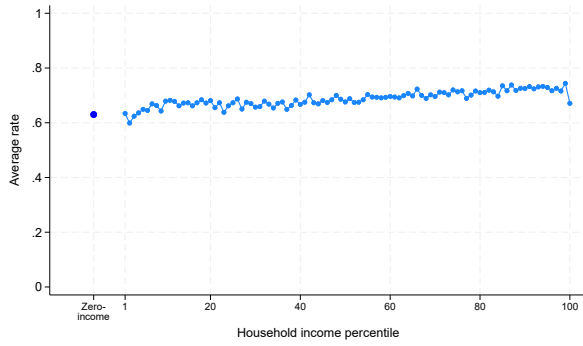
Figure A4: Behavior-income gradients



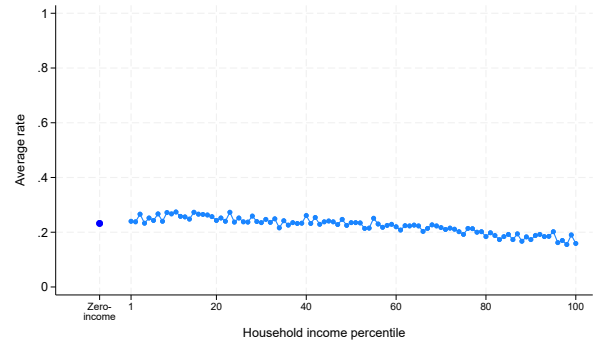
(a) smoking - men



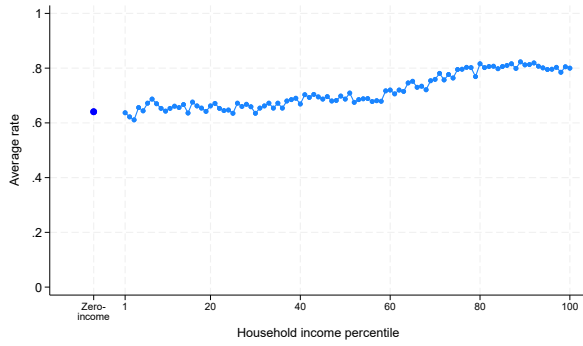
(b) smoking - women



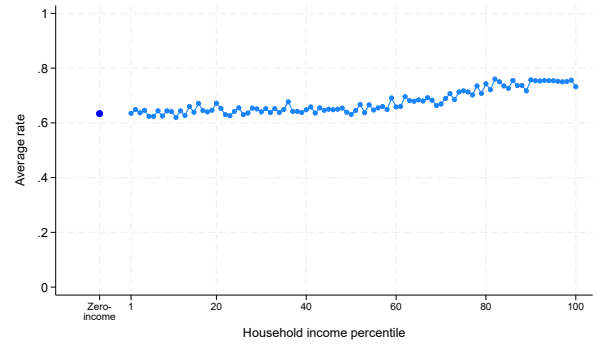
(c) drinking - men



(d) drinking - women



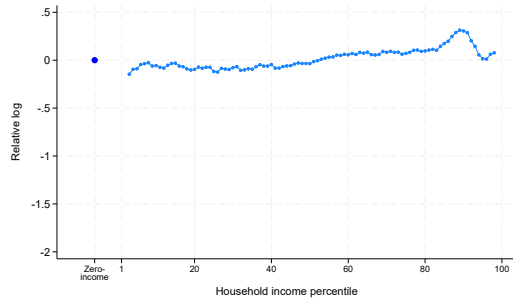
(e) exercise - men



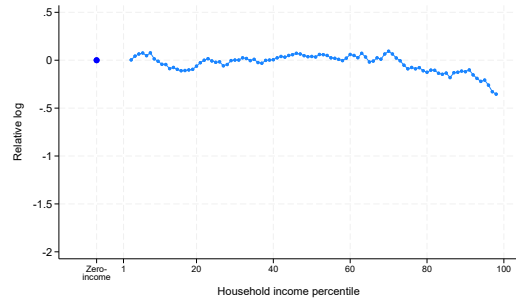
(f) exercise - women

Notes: Figure shows behavior-income gradients for smoking, drinking and exercise. x-axis shows the household income percentile. y-axis reports the average of rate at a given income.

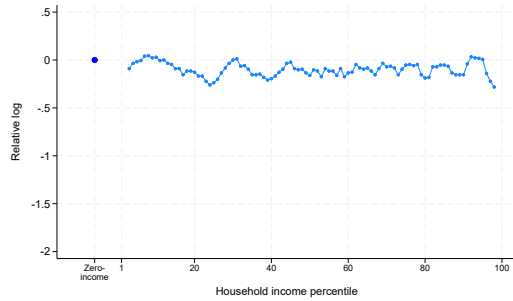
Figure A5: Income gradient in successfully controlled biomarkers



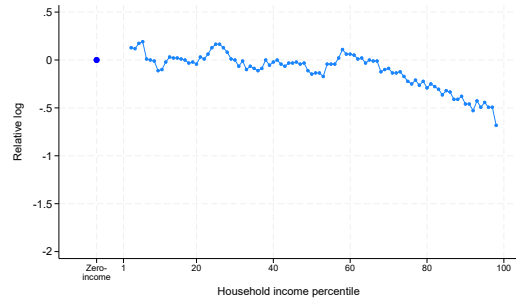
(a) high blood pressure - men



(b) high blood pressure - women



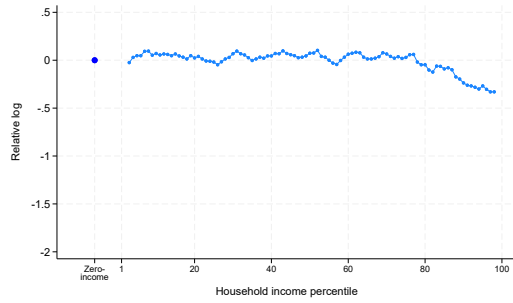
(c) high blood glucose - men



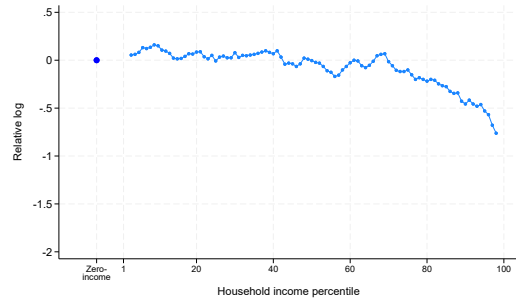
(d) high blood glucose - women

Notes: Sample is limited to exam population (men: 167,233 and women: 181,903.) x-axis shows the household income percentile. y-axis reports the (natural) log of the 5-percentile moving average rate at a given income percentile minus the log of the average rate across all income percentiles. The average rate of successfully controlled high blood pressure is 0.132 (men) and 0.110 (women). The average rate of successfully controlled high blood glucose is 0.035 (men) and 0.019 (women). 23% of men and 26% of women have zero taxable income.

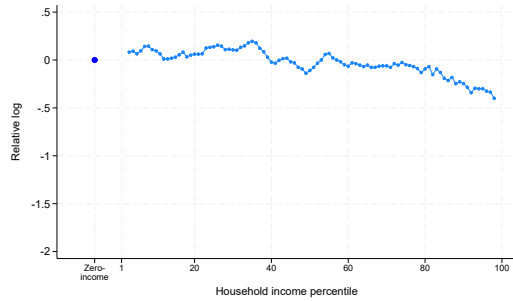
Figure A6: Income gradient in unsuccessfully controlled biomarkers



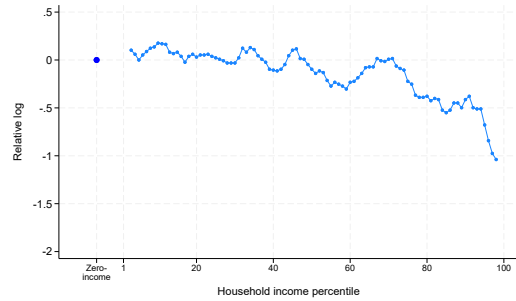
(a) high blood pressure - men



(b) high blood pressure - women



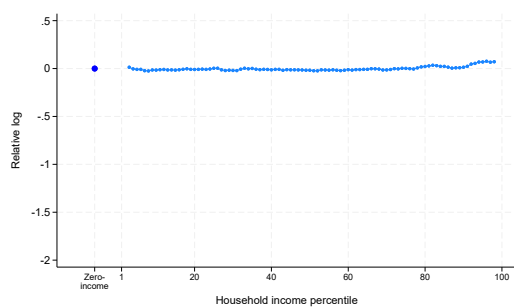
(c) high blood glucose - men



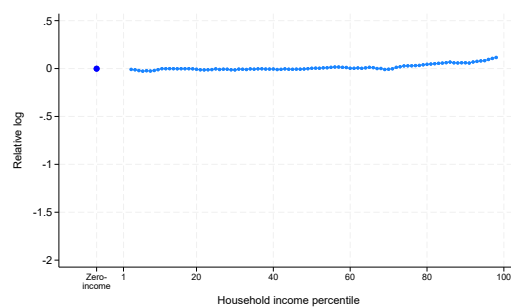
(d) high blood glucose - women

Notes: Sample is limited to exam population (men: 167,233 and women: 181,903.) x-axis shows the household income percentile. y-axis reports the (natural) log of the 5-percentile moving average rate at a given income percentile minus the log of the average rate across all income percentiles. The average rate of unsuccessfully controlled high blood pressure is 0.074 (men) and 0.054 (women). The average rate of unsuccessfully controlled high blood glucose is 0.054 (men) and 0.026 (women). 23% of men and 26% of women have zero taxable income.

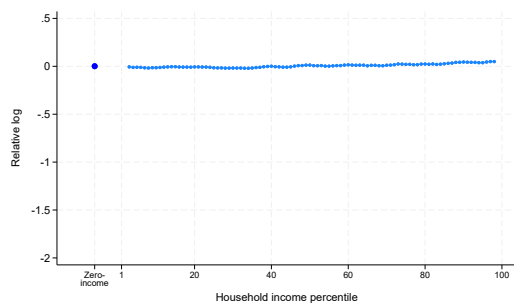
Figure A7: Income gradient in healthy biomarkers



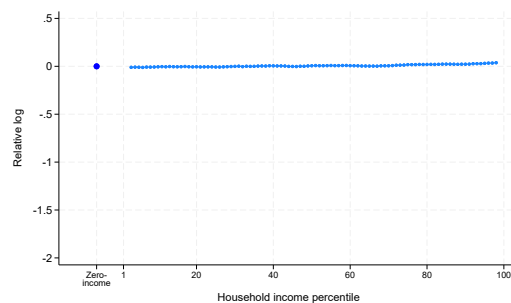
(a) high blood pressure - men



(b) high blood pressure - women



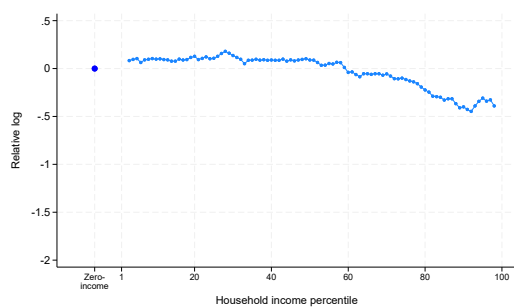
(c) high blood glucose - men



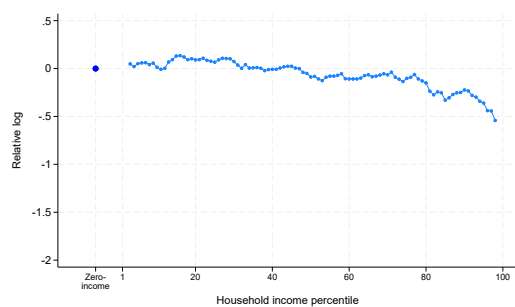
(d) high blood glucose - women

Notes: Sample is limited to exam population (men: 167,233 and women: 181,903.) x-axis shows the household income percentile. y-axis reports the (natural) log of the 5-percentile moving average rate at a given income percentile minus the log of the average rate across all income percentiles. The average rate of healthy blood pressure is 0.672 (men) and 0.759 (women). The average rate of healthy blood glucose is 0.854 (men) and 0.932 (women). 23% of men and 26% of women have zero taxable income.

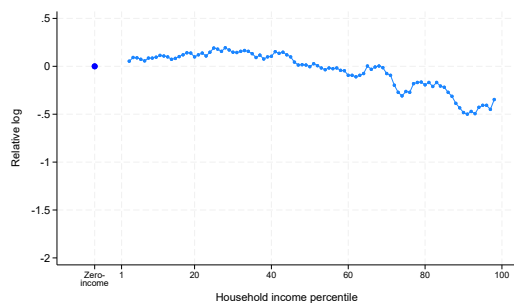
Figure A8: Income gradient in uncontrolled biomarkers



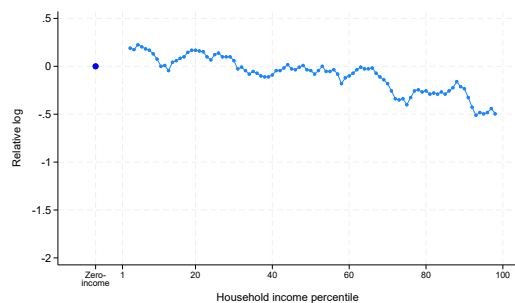
(a) high blood pressure - men



(b) high blood pressure - women



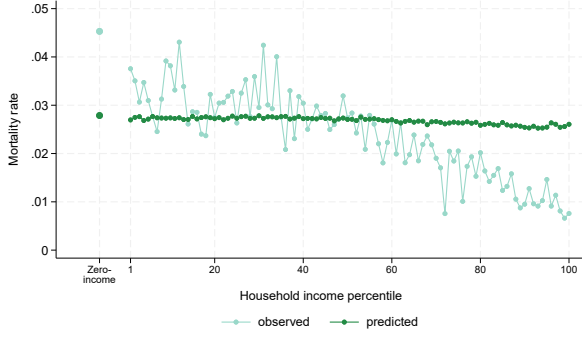
(c) high blood glucose - men



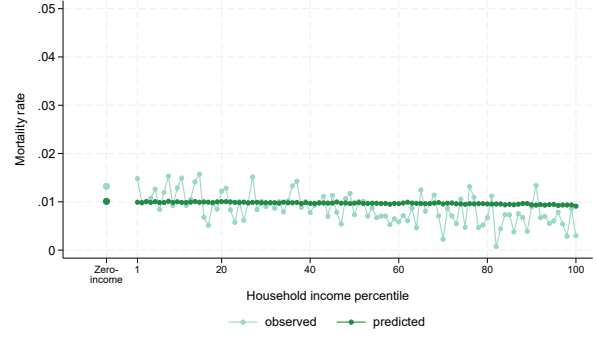
(d) high blood glucose - women

Notes: Sample is limited to exam population (men: 167,233 and women: 181,903.) x-axis shows the household income percentile. y-axis reports the (natural) log of the 5-percentile moving average rate at a given income percentile minus the log of the average rate across all income percentiles. The average rate of uncontrolled high blood pressure is 0.122 (men) and 0.076 (women). The average rate of uncontrolled high blood glucose is 0.058 (men) and 0.023 (women). 23% of men and 26% of women have zero taxable income.

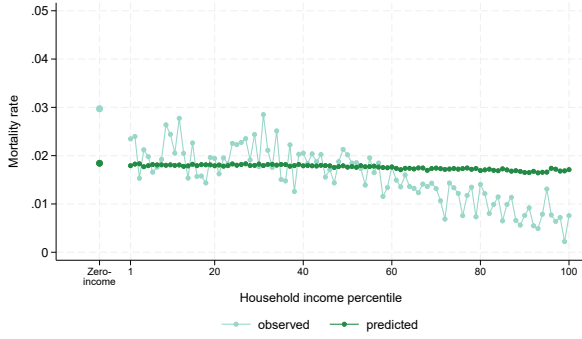
Figure A9: Mortality-income gradients: Actual vs. Predicted (without behaviors)



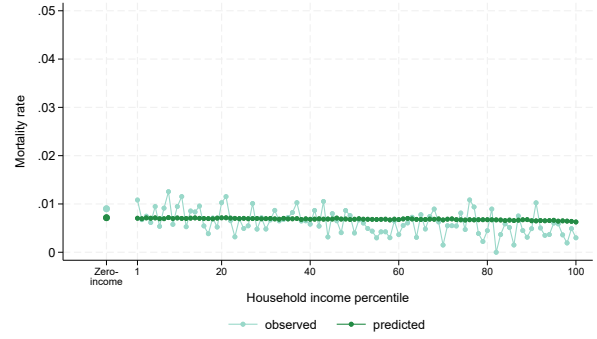
(a) all-cause mortality - men



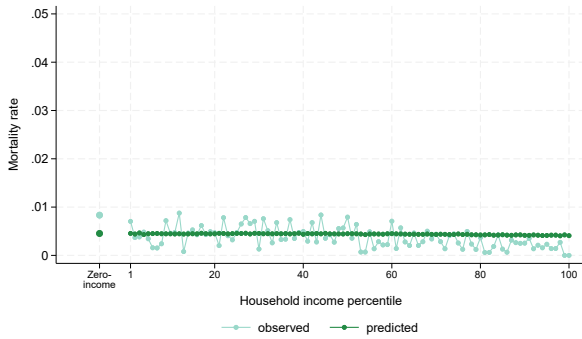
(b) all-cause mortality - women



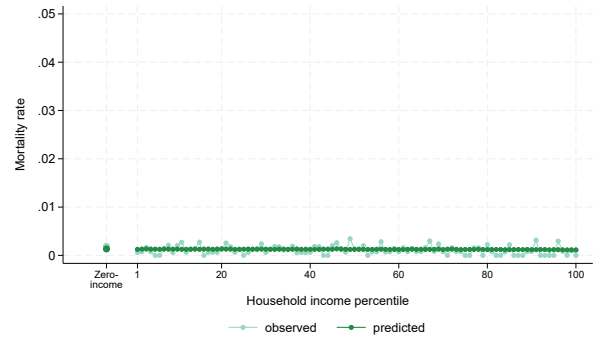
(c) non-external mortality - men



(d) non-external mortality - women



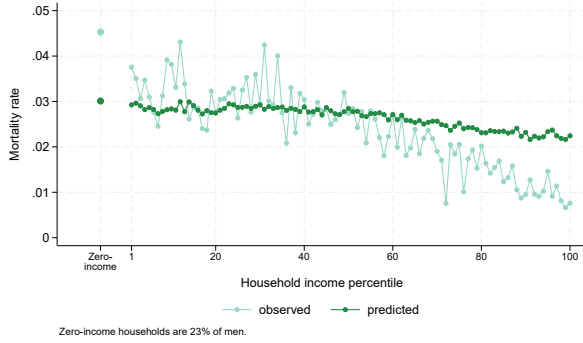
(e) cardiovascular mortality - men



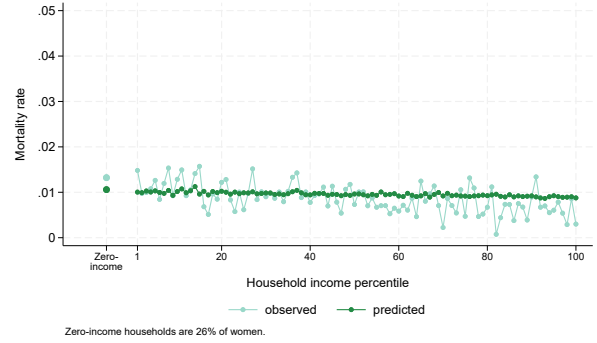
(f) cardiovascular mortality - women

Notes: Sample is limited to health participants. Figure shows mortality-income gradients and predicted-mortality income gradients for 167,233 men (left panels) and 181,903 women (right panels) separately for all-cause mortality, mortality from non-external causes, and cardiovascular mortality. x-axis shows the household income percentile. y-axis reports the average rate at each income percentile. 23% of men and 26% of women have zero taxable income (light and dark green dots).

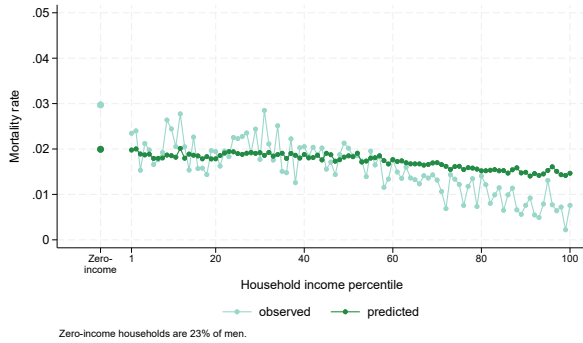
Figure A10: Mortality-income gradients: Actual vs. Predicted (with behaviors)



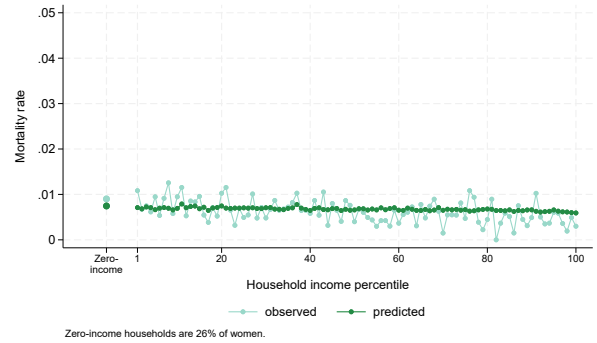
(a) all-cause mortality - men



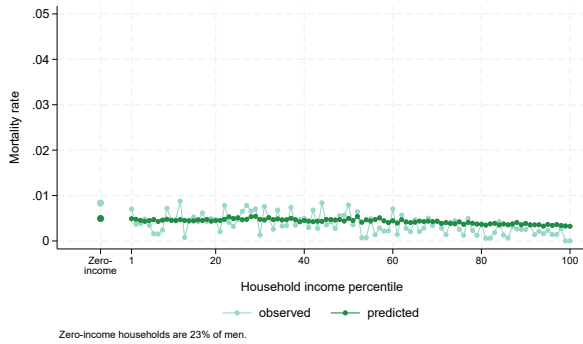
(b) all-cause mortality - women



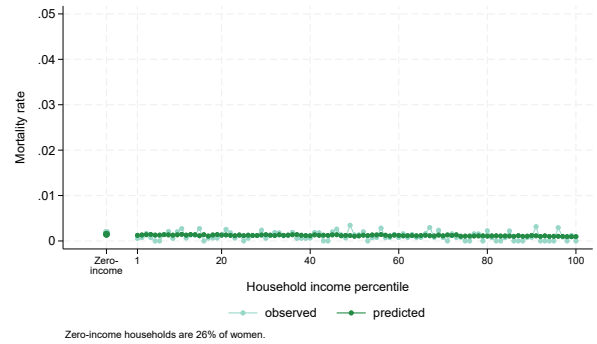
(c) non-external mortality - men



(d) non-external mortality - women



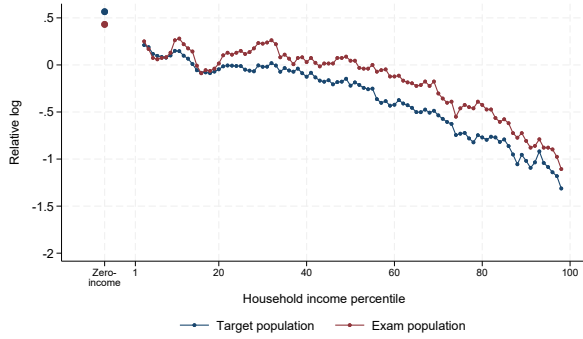
(e) cardiovascular mortality - men



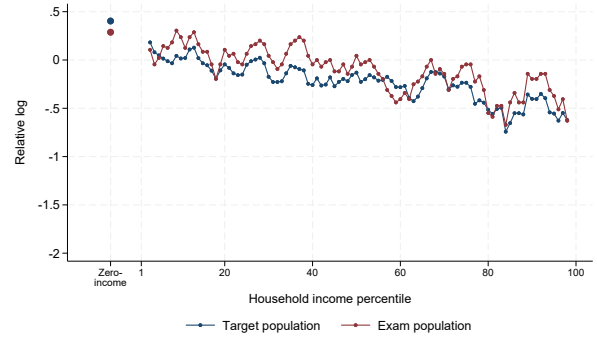
(f) cardiovascular mortality - women

Notes: Sample is limited to health participants. Figure shows mortality-income gradients and predicted-mortality income gradients for 167,233 men (left panels) and 181,903 women (right panels) separately for all-cause mortality, mortality from non-external causes, and cardiovascular mortality. x-axis shows the household income percentile. y-axis reports the average rate at each income percentile. 23% of men and 26% of women have zero taxable income (light and dark green dots).

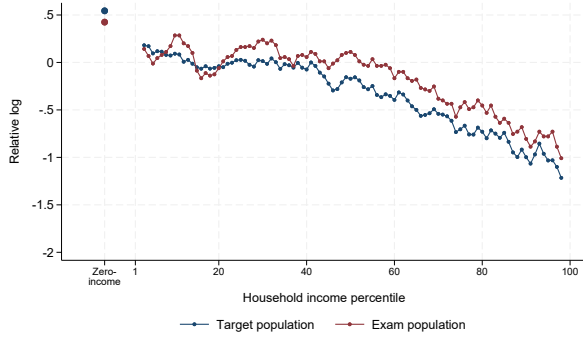
Figure A11: Mortality-income gradients - Target Population vs. Exam Population



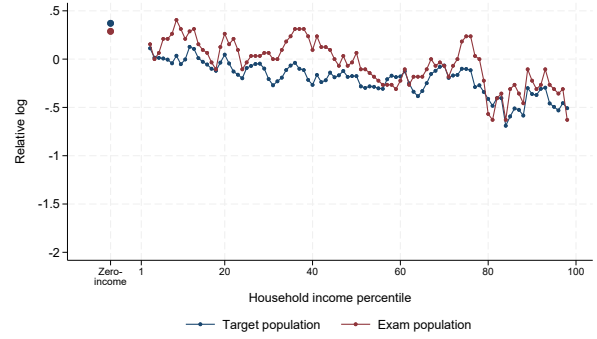
(a) all-cause mortality rate - men



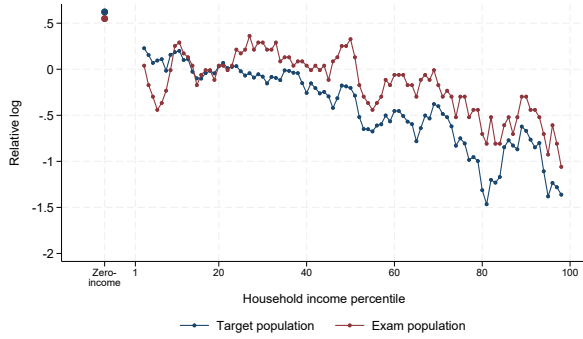
(b) all-cause mortality rate - women



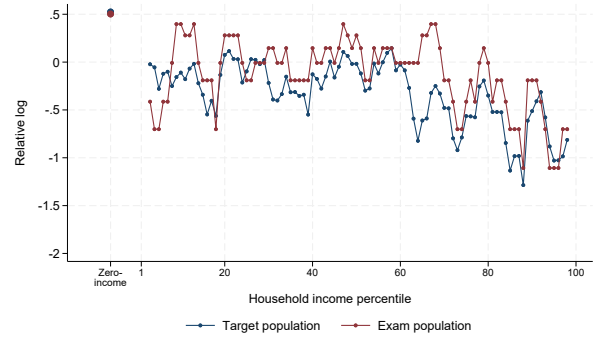
(c) non-external mortality rate - men



(d) non-external mortality rate - women



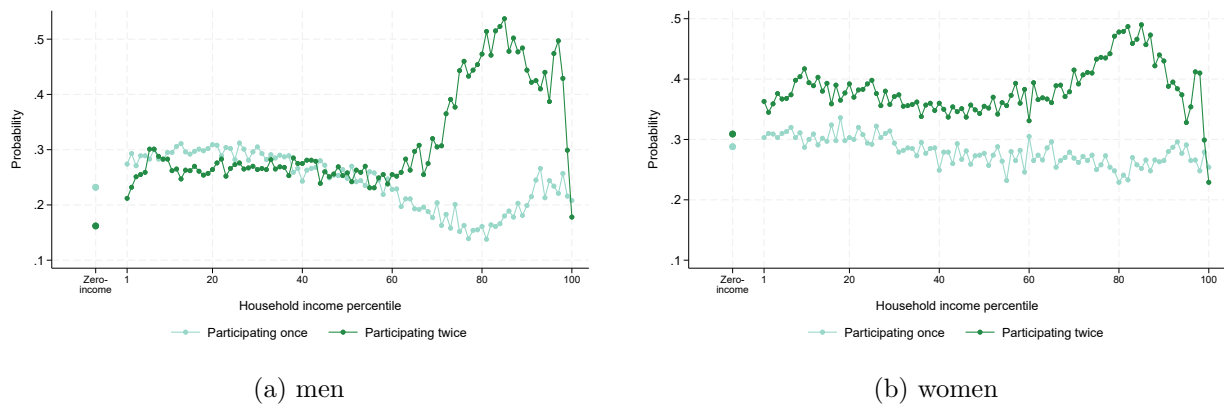
(e) cardiovascular mortality rate - men



(f) cardiovascular mortality rate - women

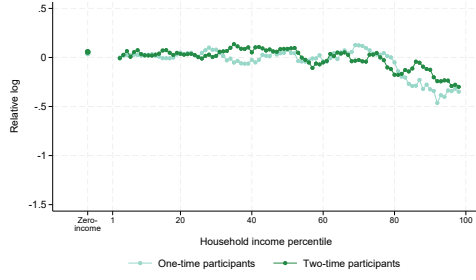
Notes: Mortality is measured as 8-year mortality from the first year we could have observed an individual in a health exam after they turned 50; mean mortality rates are given in Table A5. x-axis shows the household income percentile. y-axis reports the (natural) log of the a 5-percentile moving average of the mortality rate at a given income percentile minus the log of the average mortality rate across all income percentiles for that population. Sample size for target population is men: 249,325 and women: 243,964. Sample size for exam population is men: 167,233 and women: 181,903. 23% of men and 26% of women have zero taxable income.

Figure A12: Probability of participation by income percentile and participation frequency

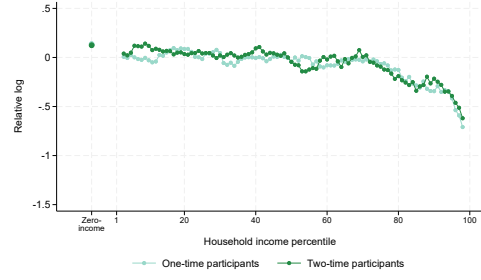


Notes: Sample is limited to the sub-sample of the target population who are not manual employees, and therefore are supposed to receive biennial health exams (men: 181,784 and women: 207,645). Average probability rates of participating once are 0.248 (men) and 0.284 (women) and of participating twice are 0.264 (men) and 0.362 (women). 23% of men and 26% of women have zero taxable income.

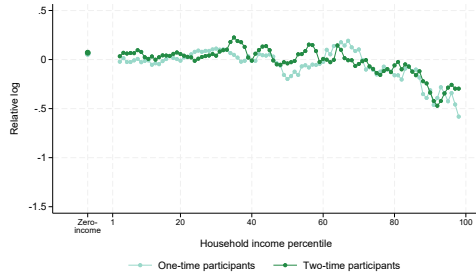
Figure A13: Biomarker-income gradient by participation frequency



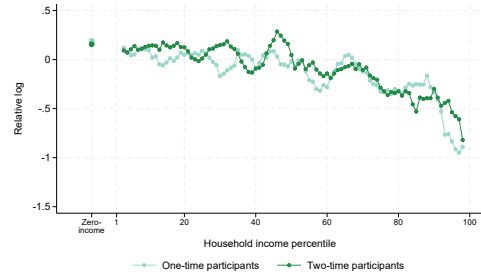
(a) high blood pressure - men



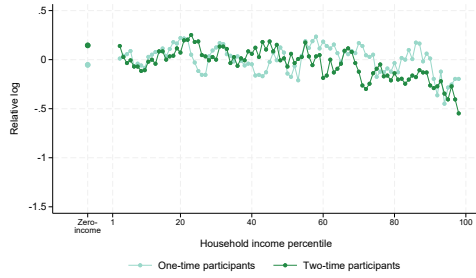
(b) high blood pressure - women



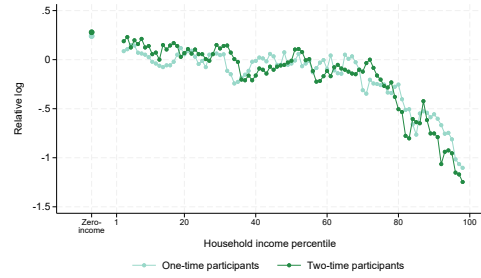
(c) high blood glucose - men



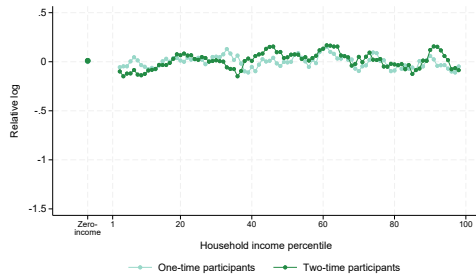
(d) high blood glucose - women



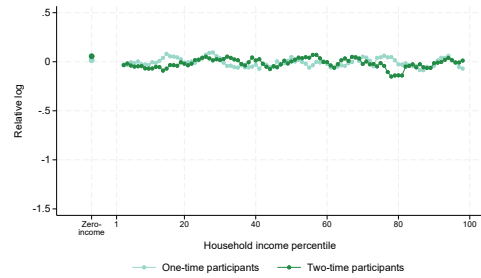
(e) high BMI - men



(f) high BMI - women



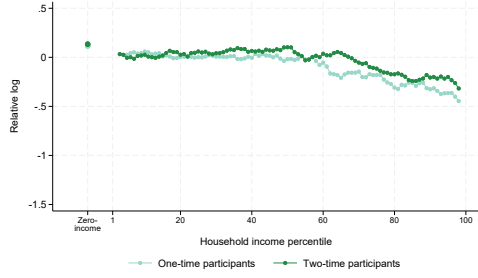
(g) high cholesterol - men



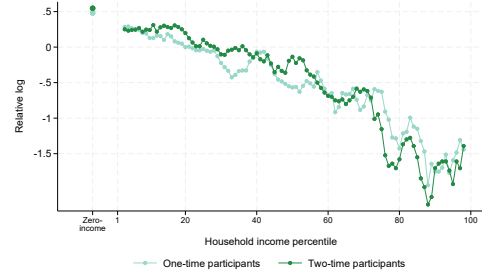
(h) high cholesterol - women

Notes: Sample is limited to the sub-sample of the exam population who are not manual employees, and therefore are supposed to receive biennial health exams (men: 93,078 and women: 134,179). We then show results separately for those who participate in only one health exam between 2009 and 2012 (One-time participants; men: 45,098 and women: 59,052) and those who participate in two health exams over this period (Two-time participants; men: 47,980 and women: 75,127). x-axis shows the household income percentile. y-axis reports the (natural) log of the 5-percentile moving average of high biomarker rate at a given income percentile minus the log of the average high biomarker rate across all income percentiles. 23% of men and 26% of women have zero taxable income.

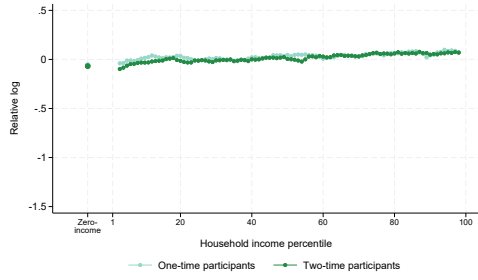
Figure A14: Behavior-income gradient by participation frequency



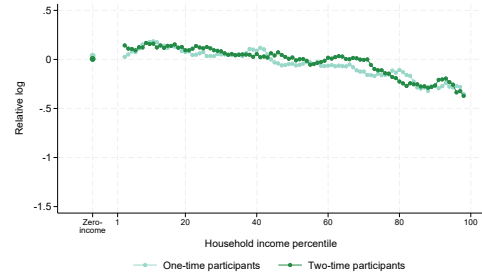
(a) smoking - men



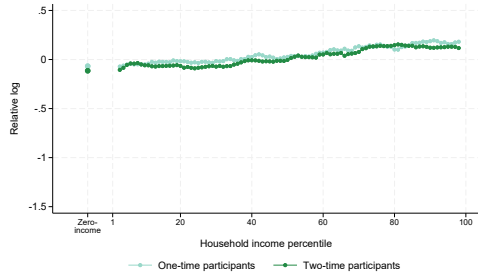
(b) smoking - women



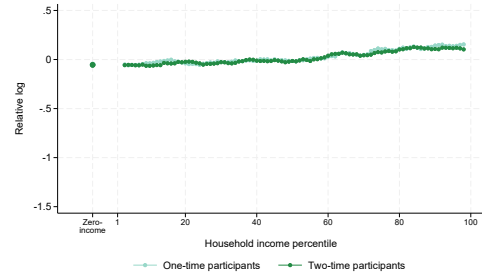
(c) drinking - men



(d) drinking - women



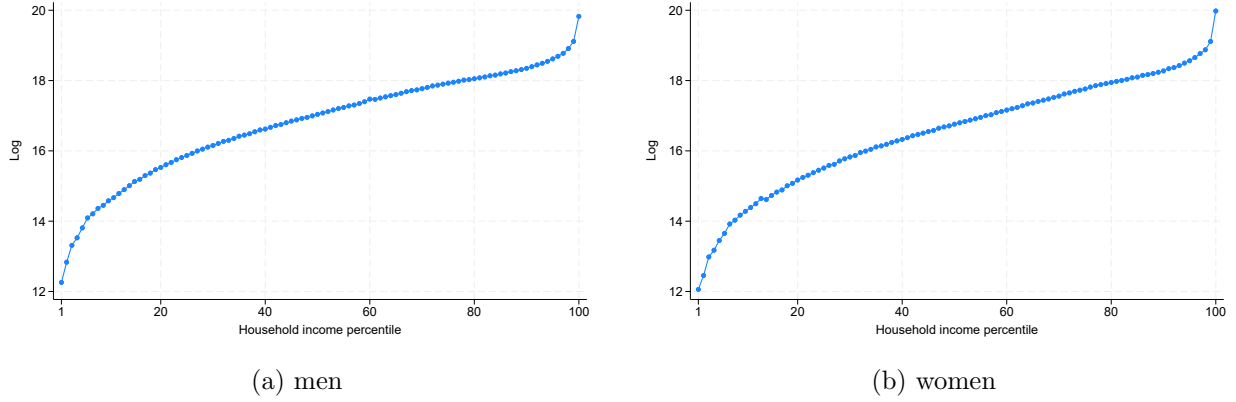
(e) exercise - men



(f) exercise - women

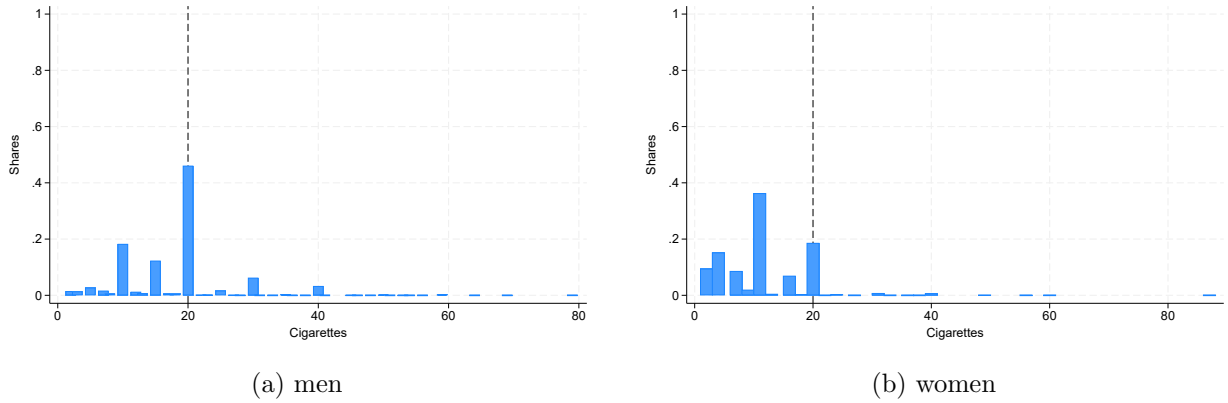
Notes: Sample is limited to the sub-sample of the exam population who are not manual employees, and therefore are supposed to receive biennial health exams (men: 93,078 and women: 134,179). We then show results separately for those who participate in only one health exam between 2009 and 2012 (One-time participants; men: 45,098 and women: 59,052) and those who participate in two health exams over this period (Two-time participants; men: 47,980 and women: 75,127). x-axis shows the household income percentile. y-axis reports the (natural) log of the 5-percentile moving average of behavior rates at a given income percentile minus the log of the average behavior rate across all income percentiles. 23% of men and 26% of women have zero taxable income.

Figure A15: Mean household income by income percentile



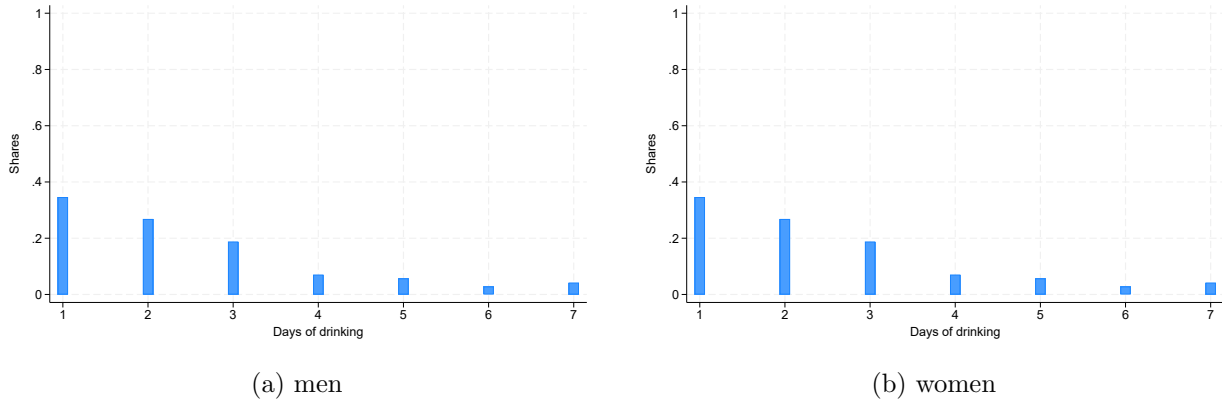
Notes: Sample consists of target population participants (men: 249,325 and women: 243,964). Household income is the log of mean household income over 3 years.

Figure A16: Distribution of answers to smoking question



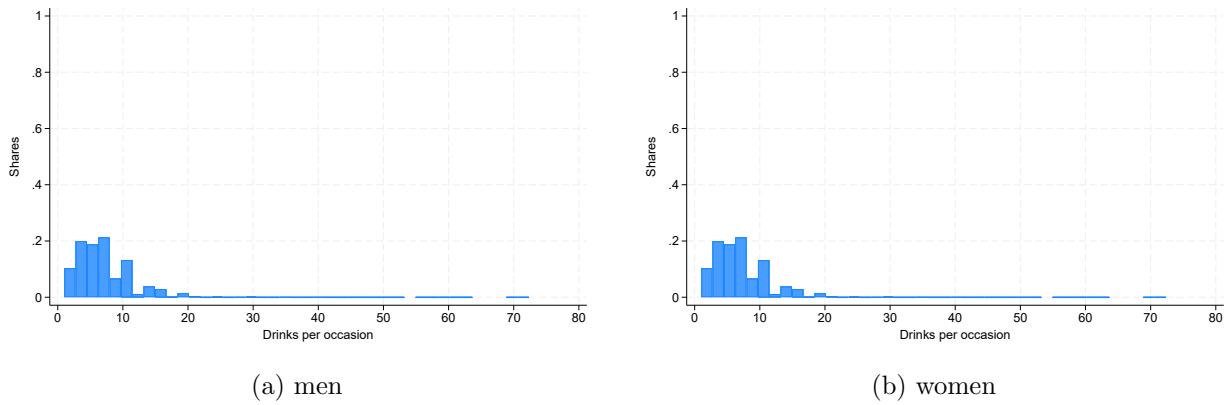
Notes: Figure shows the distribution of answers to question “How many cigarettes do you smoke a day?”. The mean values for men and women are 18.05 (sd: 7.92) and 11.20 (sd: 6.72) respectively. Dashed line indicates a cutoff for being a heavy smoker.

Figure A17: Distribution of answers to drinking frequency question



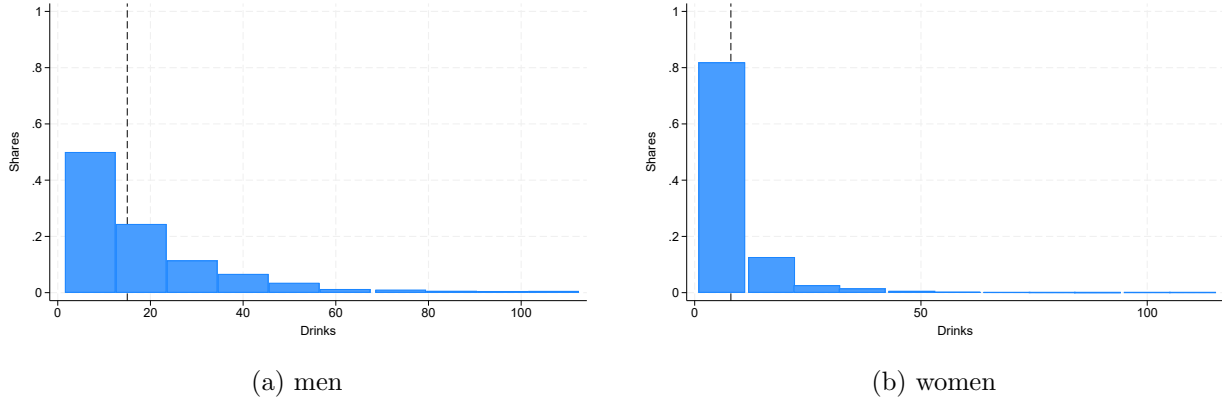
Notes: Figure shows the distribution of answers to question "How many days do you drink a week?". The mean values for men and women are 2.48 (sd: 1.61) and 1.78 (sd: 1.32) respectively.

Figure A18: Distribution of answers to drinking amount question



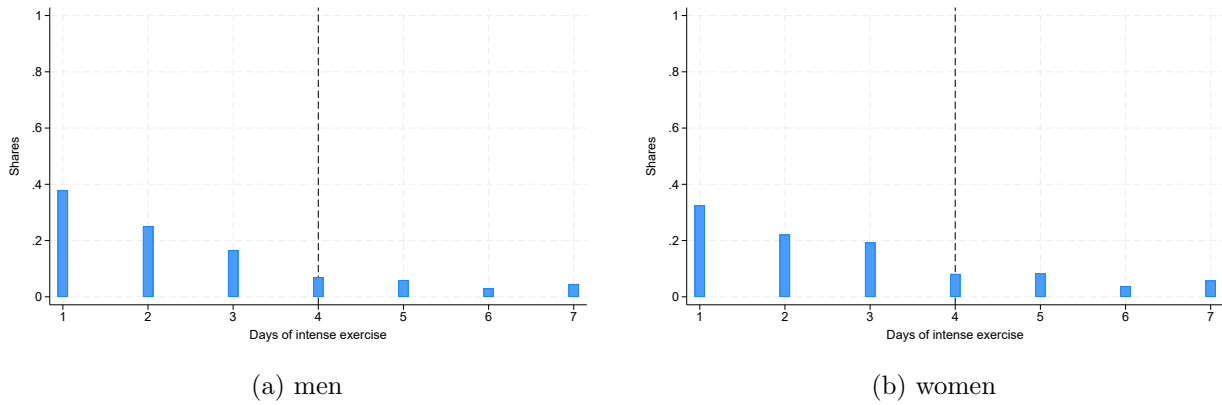
Notes: Figure shows the distribution of answers to question "How many glasses do you drink at one time?". The mean values for men and women are 6.80 (sd: 4.28) and 3.66 (sd: 2.85) respectively.

Figure A19: Distribution of weekly alcohol consumption



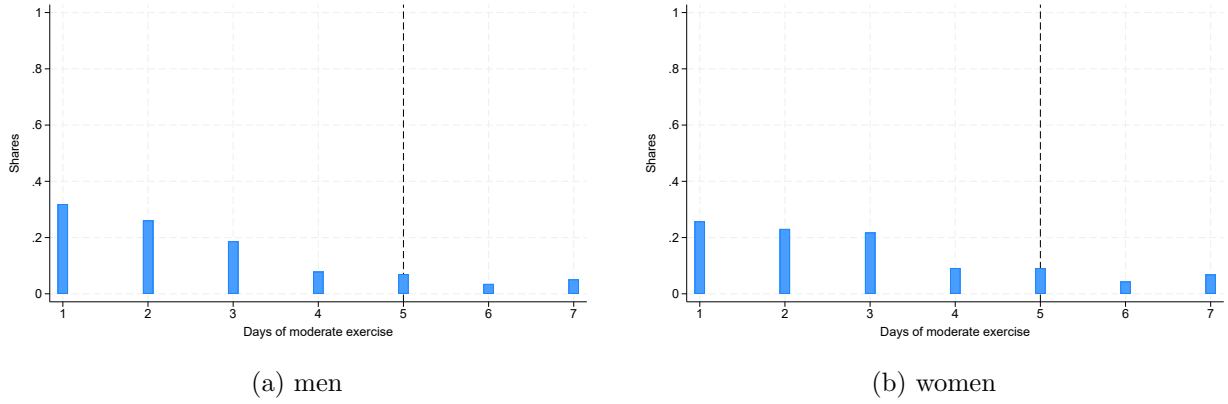
Notes: Figure shows the distribution of weekly alcohol consumption. The mean values for men and women are 18.27 (sd: 19.49) and 7.41 (sd: 11.17) respectively. Dashed line indicates a cutoff for being a heavy drinker.

Figure A20: Distribution of answers to vigorous exercise question



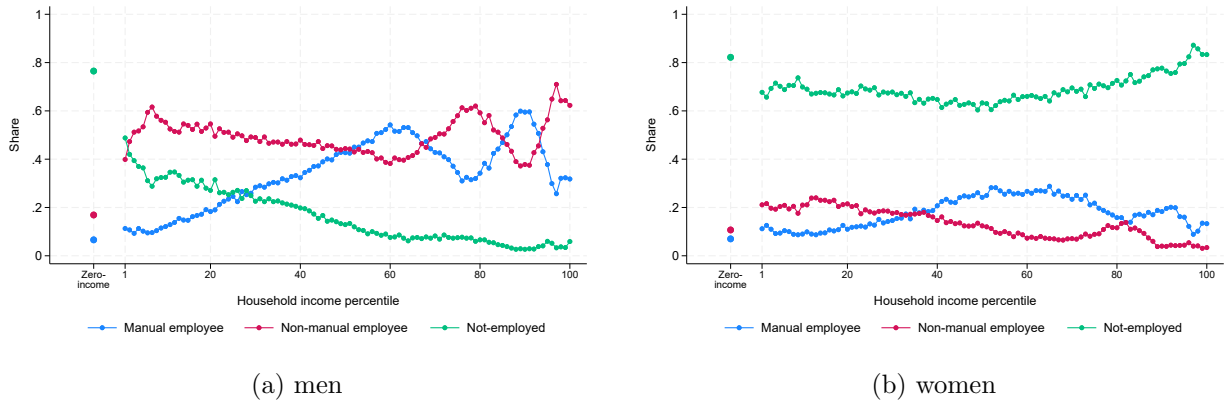
Notes: Figure shows the distribution of answers to question "How many days do you do a vigorous workout more than 20 minutes a week?". The mean values for men and women are 2.45 (sd: 1.66) and 2.72 (sd: 1.77) respectively. Dashed line indicates a cutoff for sufficient exercise.

Figure A21: Distribution of answers to moderate exercise question



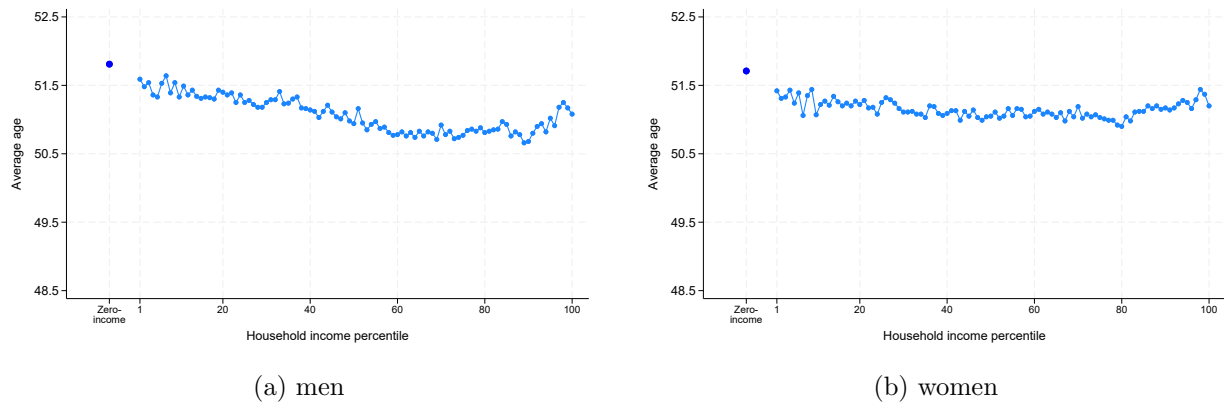
Notes: Figure shows the distribution of answers to question “How many days do you do a moderate workout more than 30 minutes a week?”. The mean values for men and women are 2.63 (sd: 1.70) and 2.93 (sd: 1.79) respectively. Dashed line indicates a cutoff for sufficient exercise.

Figure A22: Worker type by income percentile



Notes: Sample consists of target population (men: 249,325 and women: 243,964). 23% of men and 26% of women have zero taxable income.

Figure A23: Mean age by income percentile



Notes: Sample is limited to exam population, and age is defined as the age at which we observe them taking their first health exam in 2009-2012. 23% of men and 26% of women have zero taxable income.

Appendix Tables

Table A1: Average health risks: Korea vs. United States

	Korea		U.S.	
	Men	Women	Men	Women
High blood pressure	0.305	0.186	0.200	0.137
High blood glucose	0.113	0.054	0.155	0.068
High BMI	0.026	0.046	0.379	0.391
High cholesterol	0.101	0.130	0.169	0.220
Observations	2,246	3,112	1,062	1,093

Notes: Table compares 2009-2012 data from the U.S. National Health and Nutrition Examination Survey or NHANES ([National Center for Health Statistics 2009-2012](#)) with data from the Korean National Health and Nutrition Examination Survey, or KNHANES ([Korea Centers for Disease Control and Prevention 2009-2012](#)). We select participants aged 47 and 57 (average age in NHANES: 51.97 (men) and 51.71 (women); KNHANES: 52.03 (men) and 51.90 (women)). Outcomes are defined using the same cutoffs as in the main analyses (see footnote 7.) All means are weighted using the appropriate survey weights.

Table A2: Annual mortality rates - Korea vs. United States

	Korea			US		
	Overall	Men	Women	Overall	Men	Women
All-cause mortality						
Age 35-44	0.001	0.002	0.001	0.002	0.002	0.001
Age 45-54	0.003	0.005	0.002	0.004	0.005	0.003
Age 55-64	0.007	0.011	0.004	0.009	0.011	0.007
Age 65+	0.037	0.045	0.032	0.047	0.050	0.045
Share external mortality						
Age 35-44	0.349	0.369	0.300	0.345	0.397	0.258
Age 45-54	0.220	0.233	0.184	0.168	0.192	0.128
Age 55-64	0.137	0.147	0.111	0.064	0.074	0.048
Age 65+	0.065	0.076	0.056	0.025	0.030	0.021
Share CV mortality						
Age 35-44	0.143	0.154	0.116	0.218	0.225	0.205
Age 45-54	0.174	0.172	0.179	0.299	0.322	0.261
Age 55-64	0.225	0.217	0.247	0.357	0.383	0.318
Age 65+	0.344	0.328	0.358	0.467	0.490	0.448

Note: Based on the official mortality data in 2006 from [Statistics Korea \(2023\)](#) and [Centers for Disease Control and Prevention \(2010\)](#). Share external mortality is the share of all-cause mortality from external causes and Share CV mortality is the share of all-cause mortality from cardiovascular disease.

Table A3: Bivariate relationship between 8-year mortality and health risk factors

	All-cause		Non-external		Cardiovascular	
	Men	Women	Men	Women	Men	Women
Mean mortality	0.027	0.010	0.018	0.007	0.004	0.001
Biomarkers						
High blood pressure	0.251 (0.040)	0.365 (0.079)	0.260 (0.049)	0.430 (0.096)	0.629 (0.109)	1.129 (0.263)
High blood glucose	0.726 (0.058)	1.110 (0.154)	0.767 (0.072)	1.224 (0.189)	0.385 (0.133)	1.921 (0.503)
High BMI	0.082 (0.090)	0.334 (0.142)	0.205 (0.117)	0.404 (0.173)	0.763 (0.283)	-0.089 (0.330)
High cholesterol	-0.103 (0.042)	-0.125 (0.060)	-0.138 (0.051)	-0.201 (0.070)	0.283 (0.119)	-0.321 (0.155)
Health behaviors						
Smoking	0.347 (0.044)	1.009 (0.190)	0.370 (0.055)	0.671 (0.209)	0.423 (0.112)	1.466 (0.590)
Drinking	-0.337 (0.029)	-0.153 (0.062)	-0.378 (0.036)	-0.195 (0.072)	-0.326 (0.073)	-0.426 (0.152)
Exercise	-0.260 (0.032)	-0.078 (0.049)	-0.210 (0.039)	-0.120 (0.058)	-0.344 (0.080)	-0.572 (0.142)

Note: Sample is limited to exam population (men: 167,233 and women: 181,903). Each cell reports the coefficient from a separate bivariate regression of the 8-year mortality measure on biomarkers and health behaviors. Robust standard errors are in parentheses. Each column's coefficients and standard errors are divided by that column's mean mortality, so entries read as relative to the average mortality rate.

Table A4: Predicted mortality rates

	All-cause		Non-external		Cardiovascular	
	Men	Women	Men	Women	Men	Women
8-year Observed mortality rate						
Mean	0.027	0.010	0.018	0.007	0.004	0.001
8-year Predicted mortality rate						
<i>Four biomarkers</i>						
Mean	0.027	0.010	0.018	0.007	0.004	0.001
SD	0.011	0.004	0.008	0.003	0.002	0.001
<i>Four biomarkers and three health behaviors</i>						
Mean	0.027	0.010	0.018	0.007	0.004	0.001
SD	0.024	0.013	0.019	0.010	0.010	0.004

Note: Sample is limited to the exam population (men: 167,233 and women: 181,903). Predicted mortality is generated using a linear probability model. The “four biomarkers” prediction is generated by regressing (separately for men and women) a mortality measure on a fully saturated set of 81 ($=3^4$) indicators for the three possible health states (high, elevated or normal) for each of the four biomarkers. The “four biomarkers and three health behaviors” is generated analogously, but now we include a fully saturated set of 2,187 ($=3^7$) indicators for the three possible values of each of the seven variables; for the health behaviors, the three possible values are heavy, any, or none for smoking and drinking, and none, any, or sufficient for exercise.

Table A5: 8-year mortality rates

	Target population		Exam population		Non-exam population	
	Men	Women	Men	Women	Men	Women
8-year Mortality rate						
All-cause	0.045	0.016	0.026	0.009	0.085	0.037
Non-external	0.029	0.012	0.017	0.006	0.055	0.027
Cardiovascular	0.008	0.002	0.004	0.002	0.015	0.005
Observations	249,325	243,964	167,233	181,903	82,092	62,061

Note: The exam population and non-exam population are mutually exclusive and exhaustive subsets of the target population. For all three group, 8-year mortality is measured from the first year we could have observed them taking a 2009-2012 health exam after they turned 50.

Table A6: 8-year mortality rates by participation frequency

	One-time participants		Two-time participants	
	Men	Women	Men	Women
8-year Mortality rate				
All-cause	0.046	0.016	0.018	0.007
Non-external	0.031	0.012	0.011	0.005
Cardiovascular	0.008	0.002	0.003	0.001
Observations	45,098	59,052	47,980	75,127

Note: 8-year mortality is measured for the eight years after the first health exam.

Table A7: Correlation of individual household income percentiles over time

Correlation	Men	Women
2006 and 2007 income	0.922	0.900
2007 and 2008 income	0.924	0.898
2008 and 2006 income	0.881	0.846

Note: Analysis is limited to the exam population (men: 167,233 and women: 181,903). Household income percentiles is defined by yearly household income relative to the household income of all others in the individual's birth cohort and sex.

Table A8: Summary statistics

	Target population		Exam population		Non-exam population	
	Men	Women	Men	Women	Men	Women
Type of worker (free health exam regime)						
Share manual employee (annual, mandatory)	0.271	0.149	0.332	0.168	0.147	0.092
Share non-manual employee (biennial, mandatory)	0.420	0.129	0.456	0.135	0.347	0.113
Share not employed (biennial, non mandatory)	0.309	0.722	0.212	0.697	0.506	0.795
Birth year						
1955	0.109	0.114	0.111	0.117	0.103	0.106
1956	0.107	0.107	0.113	0.113	0.096	0.091
1957	0.119	0.118	0.119	0.117	0.121	0.119
1958	0.120	0.118	0.124	0.122	0.113	0.108
1959	0.130	0.131	0.126	0.128	0.138	0.138
1960	0.140	0.141	0.14	0.143	0.139	0.137
1961	0.139	0.140	0.133	0.133	0.151	0.161
1962	0.136	0.131	0.135	0.128	0.139	0.141
Average age in 2009	50.29	50.32	50.33	50.37	50.20	50.16
Share with zero income	0.228	0.261	0.149	0.223	0.389	0.373
Observations	249,325	243,964	167,233	181,903	82,092	62,061

Note: The exam population and non-exam population are mutually exclusive and exhaustive subsets of the target population.