The Socio-Economic Distribution of Choice Quality: Evidence from Health Insurance in the Netherlands*

Benjamin Handel, UC Berkeley
Jonathan Kolstad, UC Berkeley
Thomas Minten, London School of Economics
Johannes Spinnewijn, London School of Economics

December 20, 2023

Abstract

We study how choice quality relates to socio-economic factors using population-wide data on health insurance choices and utilization in the Netherlands. We document a striking choice quality gradient with respect to socio-economic status, finding that those with higher income, higher education, and training in quantitative fields make meaningfully better choices. Household income for the top 5% of decision makers, in terms of surplus captured, is almost three times higher than for those in the bottom 5%. When jointly including all factors, education degree and field of study are more important predictors of choice quality than income and other financial variables.

*We thank the Central Bureau of Statistics of the Netherlands and especially Annemieke Redeman for help with the data. Chloé de Meulenaer, Miguel Fajardo Steinhäuser and William Parker provided excellent research assistance. We thank Jason Abaluck, Anna Aizer, Saurabh Bhargava, Zarek Brot-Goldberg, John Campbell, Raj Chetty, Rebecca Diamond, Keith Ericson, Amy Finkelstein, Sebastian Fleitas, Jihye Jeon, Tim Layton, George Loewenstein, Neale Mahoney, Sarah Miller, Martin Salm, Josh Schwartzstein, Mark Shepard, Jon Skinner, Justin Sydnor and Richard Thaler for their discussions and comments. We thank seminar participants at Amazon, Delaware, Edinburgh University, IO^2 Stanford seminar series, Erasmus University Rotterdam, KULeuven, LSE, Monash University, UPF Barcelona, University of Virginia, University of Wisconsin, BU-Harvard-MIT health seminar, Statistics Netherlands, the 2021 AEA, 2021 ASHE, Essen Health Conference, KVS, the NBER Health Care, the NBER Public Economics, the Paris-London Public Economics, the CEPR Public Economics, the Sloan Economics of Inattention and the UPENN Behavioral Health meetings for excellent comments. We gratefully acknowledge funding by ERC (grant #716485), ESRC and STICERD. Some of the material in this paper was previously part of the manuscript titled "The Social Determinants of Choice Quality: Evidence from Health Insurance in the Netherlands" (Handel et al. (2020)).

Policymakers increasingly rely on market solutions that provide choice in the provision of products viewed as public goods, such as retirement investments (see, e.g., Hastings et al. (2016) and Chetty et al. (2014)), schooling (see, e.g., Neilsen (2017)), electricity (see, e.g., Ito (2015)), and health insurance (see, e.g., Enthoven, Garber and Singer (2001)). One important argument for facilitating choice in such markets — rather than a uniform product, whether offered directly by the government or a regulated private firm — is the opportunity to match heterogeneous consumers with products that provide them with greater surplus. Thus, consumer choice is a central aspect of market function and an important rationale for these kinds of solutions.

In practice, if consumers make choice errors, as much prior work documents, the welfare gains from greater choice and competition can be diminished, or even eliminated. Furthermore, beyond the average consumer-product match, the distribution of choice quality and surplus determine the degree to which choice-based policy impacts inequality, for example when consumers with lower socio-economic status are less able to make complex decisions or have less opportunity to engage with those decisions (e.g., Mullainathan and Shafir (2013), Campbell (2016)).

In this paper, we investigate consumer choices and their socio-economic determinants, with an emphasis on how inequality in choice quality impacts outcomes. We study this in the context of health insurance provision in the Netherlands. The dimension we focus on is the choice of deductible — the amount in each year a consumer must pay out-of-pocket before insurance payments kick in. The Dutch setting is particularly well suited because we focus solely on the financial aspects of insurance contracts that are orthogonal to other plan differences, making it more straightforward to assess choice quality. Moreover, we can leverage rich administrative data on the universe of the population of the Netherlands (approximately 17 million people) linked to individual insurance choices.

We assess choice quality using a simple choice model together with precise health risk predictions generated with tools from machine learning (Einav et al. (2018)). The distribution of predicted health risk combined with the premium reductions associated with an increased deductible — for every 100 EUR increase in deductible premiums fall by 50 EUR — implies that a large share of consumers would gain from choosing a higher deductible option. Overall, we find that more than 50% of consumers would be better off choosing a higher deductible due to their low predicted health risk, but less than 10% actually do so. We show that the large gap between the model's predicted choices and observed choices cannot be rationalized by reasonable risk preference estimates, neither is it explained by low financial liquidity or moral hazard in our data.

We then study the socio-economic determinants of deductible choice and how it relates to predicted health risk. We document a strong and positive socio-economic gradient in choice quality overall and find that education is particularly important. The take-up rate of a high deductible is 3 times (28 percentage points) higher for a predictably healthy individual with substantial education (graduate degree) compared to a predictably healthy individual with less than high school education. This difference is 13 percentage points for those with a college degree and 5 percentage points for those finishing high school. These associations hold demographics and income fixed. We also find a positive association between income and choice quality, but this association is no longer economically meaningful when holding demographics and education fixed.

Leveraging the granularity of the data, we further document a strong positive relationship between being educated or employed in an analytic field and deductible choice quality, all else equal. For example, statistics majors are 21 percentage points more likely to choose a high deductible when predictably healthy, relative to the collection of other fields. Conversely, those with training in security are 6 percentage points less likely to choose the high deductible when predictably healthy. We illustrate this relationship between the analytic nature of education fields and profession comprehensively across 90 education fields and 68 professions documented in

our data. In contrast, all else equal, we find small associations between individuals' choice quality and their household finances including liquid savings, indebtedness and net worth.

We weave together our findings on heterogeneous choice quality in a welfare framework that classifies decision-making quality as a function of all these socio-economic characteristics jointly, conditional on health. We find that the 5% best decision-makers not only are much more educated and predominantly trained in analytic fields, they also have an average gross income of 105,000 EUR, and net worth of about 250,000 EUR. Conversely, the 5% worst decision makers have average income of 40,000 EUR and net worth of 5,000 EUR. Our results highlight that distributional considerations are important to consider when evaluating policies that embed consumer choice, something that we discuss in more detail in our conclusion.

This paper relates to several distinct literatures, but is closest to prior work on insurance choice including papers without choice frictions (see Einav, Finkelstein and Mahoney (2021) for an overview) and many with choice frictions (see Handel and Schwartzstein (2019) for an overview). Relative to this prior work, the choice we study is simpler and the data we have are much deeper and more comprehensive in terms of socio-economic factors, allowing us to contribute in several key ways. We are able to study choice heterogeneity on numerous dimensions simultaneously for the same population. Prior work on Medicare Part D choices (e.g., Abaluck and Gruber (2011), Abaluck and Gruber (2023), Ketcham et al. (2012), Ericson (2014), Polyakova (2016), Ho, Hogan and Scott Morton (2017), Brown and Jeon (2023)) typically have the largest / most representative samples, but those are also the studies that have more limited measures of socio-economic heterogeneity. On the other hand, studies with richer heterogeneity (see, e.g., Bhargava, Loewenstein and Sydnor (2017)), Fang, Keane and Silverman (2008)) occur either in specific contexts such as a large employer, or have limited sample size due to the nature of data used.[1] We are not aware of other prior studies in this space that have the depth of data we use for underlying socio-economic factors, especially at the scale of an entire country. Our analysis also relates to papers that study choice quality and the incidence of consumer frictions in other domains (e.g., Allcott, Lockwood and Taubinsky (2019), Dubois, Griffith and O'Connell (2020) and Brown et al. (2023)). Most notably, a number of papers leverage registry data to study choice quality and default effects at scale in household finance (e.g., Chetty et al. (2014), Andersen et al. (2020)).

I Institutional Context and Data

All individuals in the Netherlands are obligated to directly buy health insurance from a private health insurance market.[2] The Health Insurance Act of 2006 introduced a managed competition model in which the government strictly regulates the contents of the basic health insurance package (see Kroneman et al. (2016) for a comprehensive overview of the Netherlands health system). The regulation also (i) prohibits price discrimination, (ii) prohibits the rejection of individuals from purchasing insurance and (iii) mandates that all individuals purchase basic coverage. Insurers compete for consumers on premiums, provider networks, and supplementary insurance offerings, which covers dental care and extra physical therapy. In 2015, there were 25 health insurers that together offered 53 separate insurance contracts. Yearly premiums for the mandatory health insurance with the smallest possible deductible have a mean of 1195 EUR and a fairly compact distribution around this mean (see Online

[1] Other notable papers on insurance choice quality in other contexts include, e.g., Abaluck and Gruber (2023), Brot-Goldberg et al. (2023), Handel and Kolstad (2015*b*), and Domurat, Menashe and Yin (2021). Our paper also relates to two papers on the voluntary deductible in the Netherlands Van Winssen, Van Kleef and Van de Ven (2015) and Van Winssen, Van Kleef and Van de Ven (2016).

[2] We focus on individual choices. This corresponds to the regulatory structure of the market in the Netherlands where all adults must select their own plan annually. We note that families may make choices together (e.g. for children). This is one reason we limit our analysis to those over 18.

Appendix Figure A.2). Consumers enroll between mid-November and the end of December for the following year. During that period, health insurers advertise their insurance packages through various media. If no action is taken by the consumer, she automatically extends her current contract. Relatively few consumers switch insurers each year (6.8% of individuals in 2015).

Regulation of deductible options for the basic coverage has been a central topic of the policy debate in the Dutch Parliment. Each individual faces a compulsory deductible (375 EUR in 2015), but can opt for an extra voluntary deductible of 100, 200, 300, 400 or 500 EUR on top of this compulsory deductible (maximum total deductible of 875 EUR in 2015). The compulsory deductible, introduced in 2008, has gradually increased from 150 EUR in 2008 to 385 EUR in 2017, while the options for the extra voluntary deductible have remained the same. By opting for a higher deductible, consumers receive a premium reduction. The default option, if the consumer does not choose a higher deductible, is the baseline compulsory deductible (375 EUR in 2015). The right part of Figure A.2 in the Online Appendix shows the (unweighted) histogram of premium reductions consumers can get by electing the additional 500 EUR deductible across health plans in 2015. The distribution has a mean of 233 EUR and most of the mass lies between 200 and 300 EUR, making the deductible election a quite standardized decision across all insurance contracts.

I.A Data and Sample

We use data on health insurance choices and health expenditures for all individuals in the Netherlands. The data is linked at Statistics Netherlands to other administrative registers, which provide detailed information on individuals' income, wealth, education, employment and other demographic variables.

Sample We restrict attention to all individuals who are at least 18 years old in January of the year in which they decide on their health insurance contract and deductible. Each adult must make an individual decision annually. We exclude from the sample adults who have incomplete health data records in the two previous years. The remaining sample consists of about 13.25 million adults in each year. As explained in Section II, we use a random sample of 1.25 million of these individuals to estimate and calibrate a cost prediction model, leaving approximately 12 million adults each year for the analyses, which we call our baseline sample. The Online Appendix provides sample summary statistics and distributions of health care expenditures for the year 2015.

Health Insurance Deductible Choices We obtain data on health insurance contract choices from 2013 to 2017 from Vektis, an organization that is responsible for the collection of data from all health insurers. Our data include only information on an insurer and deductible choice. We do not observe whether individuals purchase supplementary insurance, but these choice dimensions are orthogonal to the deductible choice except for minor price differences. Table A.3 in the Online Appendix shows the take-up of different deductible amounts in 2015. The take-up rate of an extra deductible in our sample is 9.06% in 2015. More than 2 out of 3 individuals opting for an extra deductible take the maximum extra deductible of 500 EUR.

Health Care Costs Data on health care costs contain annual health care expenditures by category. The categories included in spending covered by the deductible are medicines, hospital care, geriatric care, paramedical care and physiotherapy, mental health care, aids and tools for health, health care in foreign countries, health care transport, multidisciplinary care, sensory handicap care, and other care. The aggregate distribution is skewed with about 19 percent of individuals making zero expenditures and more than 10 percent of individuals spending

more than 5000 EUR (see Online Appendix Figure A.1). Note that we also have data on preventive, maternal and GP care, but these are covered at zero cost by all insurers by law.

Education, Financial, and Demographic Data We obtain information on other variables from a number of administrative registers and link these to the health and insurance data. Our data includes standard demographics like age, gender and household status. We use third-party reported information from tax registers on household income and household wealth. The former includes pre-tax income from labor, self-employment and capital and government transfers. The latter includes information on net worth, liquid and other financial assets, mortgage and other debt. We also observe data on the highest formal education level attained for more than half of the sample. These data also include information on the specific field of study for individuals who proceed past high school as well as each individual's employment sector. We provide more detail about the different registers and variables in Online Appendix A.2.

II Deductible Choice and Health Risk

Our setting allows us to explore deductible choice quality using a very simple calculation in which we can abstract away from many aspects of insurance demand such as risk preferences, liquidity and moral hazard.

II.A Optimal Deductible Choice

We represent the individual's decision problem as the binary choice between the baseline deductible of 375 EUR and adopting a high deductible of 875 EUR with the latter corresponding to a premium reduction of 250 EUR. For those expecting to spend less than 375 EUR — a large group as we show — the savings on premiums make the high deductible a preferred choice. More formally, we can approximate an individual's expected utility by:

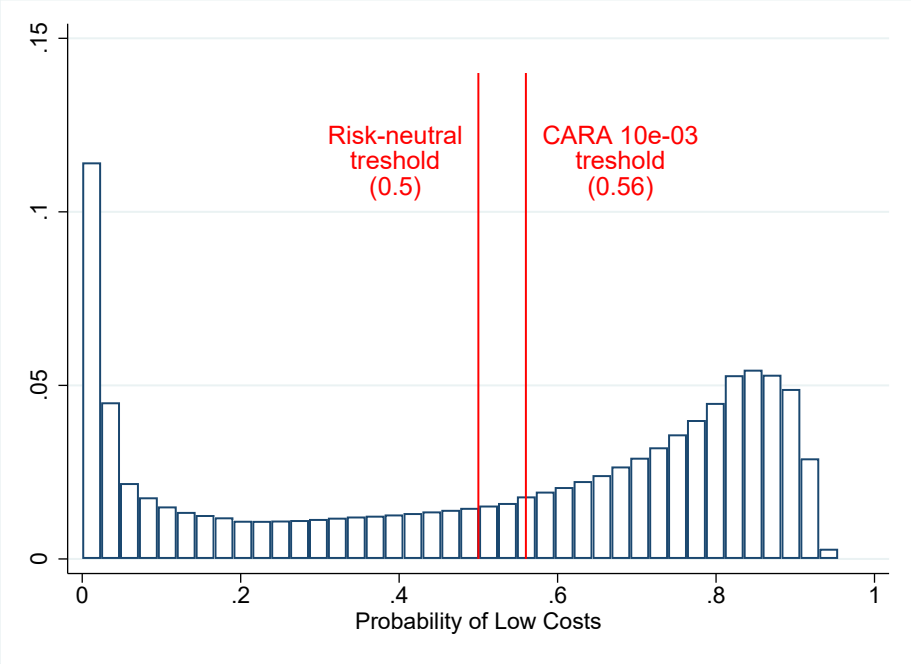
$$U_{i,d} \approx \pi_i u_i(W_i - p_d) + (1 - \pi_i)\, u_i(W_i - p_d - d), \tag{1}$$

where π_i denotes the chance that expenditures stay below 375 EUR, W_i denotes individual i's wealth, p_d is the annual premium for each deductible level, denoted by d. Our main analysis relies on the simple, risk neutral solution to the model. Solving for the expected payoff terms, $\bar{\pi} = 0.5$ is the (approximate) threshold leaving individuals indifferent between the two deductible options. Thus, to assess choices we simply need a prediction of an individual's probability of spending more than 375 EUR (π_i) and their deductible choices.[3]

We develop a prediction model to estimate the likelihood an individual spends more than 375 EUR as a binary classification algorithm. The yearly predictions of π_i are made using an ensemble learning model consisting of a random forest model, a boosted regression trees model and a LASSO model (see e.g. Einav et al. (2018)). We only include predictors that are known at the time of choice including gender, age, income ($t-2$, $t-1$), work status, education level, education field, and past health spending per category ($t-2, t-1$). We use a training sample of 1.25 million individuals, while all the results shown for the remainder of this paper use a hold-out sample of approximately 12 million observations each year. Online Appendix Figure A.5 describes the precision and fit of this model, which are both very strong, along with the model's outcomes.

[3] In theory, the optimal decision depends on the probability distribution of expenditures between 375 EUR and 875 EUR too, but the share of expenditures that fall in this range is small, as shown in Online Appendix Figure A.9. Our main analyses ignore interior choices between the two levels, which are less frequent and not easily rationalized under standard preferences. Our empirical results are robust to alternative modeling choices, as shown in Online Appendix Table A.5.

Figure 1: Distribution of Cost Probability Predictions



Notes: This figure shows the distribution of the predicted probabilities of having health costs below 375 EUR. These probabilities are obtained when predicting the binary variable (having insurable health costs below 375) with the ensemble machine learner (see Online Appendix A.3 for further details). The figure presents the risk-neutral threshold for someone to choose the 500 EUR incremental deductible if the incremental premium reduction is the modal incremental premium reduction of 250 EUR. It then presents the same threshold for extreme risk-aversion (CARA coefficient $1 * 10^{-3}$).

Combining our simple model with the cost predictions clearly demonstrates why many individuals benefit from switching to the higher deductible. Figure 1 shows the distribution of predicted cost relative to the threshold of $\bar{\pi} = 0.5$. A substantial share of individuals have predicted costs that fall well below this threshold and thus can expect to benefit from choosing the high deductible. In fact, the specifics of the setting offer a novel opportunity to abstract away from a richer model of preferences and behavior. The case of classical risk aversion makes it clear why. For standard but lower value of absolute risk aversion of 10^{-5}, this threshold increases very slightly to 0.5006. Even for an extreme level of absolute risk aversion of 10^{-3}, this threshold is still only 0.56.[4] We include this alternative threshold in Figure 1, which demonstrates visually why results are unchanged: there are very few individuals whose expected spending falls within a range affected by those assumptions. In addition, we can leverage the richness of the data to rule out other natural determinants of insurance choice. In particular, we will show that the lack of liquid savings has limited impact on deductible choice, something that prior conceptual research by Ericson and Sydnor (2018) illustrates can have meaningful and multi-faceted effects on optimal choices.

We also find a limited role for moral hazard, a potentially important factor in optimal plan choices. Moral hazard could cause consumers to reduce care consumption in response to greater cost sharing (e.g., Newhouse (1993), Einav, Finkelstein and Schrimpf (2010), Brot-Goldberg et al. (2017)). Under a classical model of moral hazard, our framework under-predicts value from the high deductible plan since it rules out reductions in care that are lower in value than the associated cost savings. Since our empirical results focus on significant under-adoption of higher deductibles, having the lower bound interpretation does not impact the main import of our

[4]See discussion in Barseghyan et al. (2018) for typical risk preference estimates in different contexts.

results. Moreover, we show in the Online Appendix A.5 that deductible choice has a small impact on realized spending, holding all else equal, suggesting that the combined impact of selection on private information and moral hazard is small relative to the value embedded in the deductible choice.

Overall, while one can thus build a much richer model of standard consumer preferences, behavior and constraints, the qualitative findings from our simpler model should remain unchanged moving to these more complex frameworks.

II.B Observed Deductible Choice

We next turn to studying how deductible choices relate to predicted health risk. Panel A in Figure 2 plots the empirical relationship between predicted health risk and deductible choice and shows the optimal choice in the simplified, risk-neutral model for comparison. Two key facts emerge. First, as expected, people who are healthier are more likely to elect the higher incremental 500 EUR deductible. Second, the relationship between risk and deductible choice is substantially weaker than one would expect if consumers were making optimal choices, not subject to any frictions. For example, the share of consumers in the healthiest predicted health bin electing the high deductible is only 17%. These individuals face a 90% chance of having costs below the lowest deductible, exposing themselves to an expected cost of only about 50 EUR when taking the highest deductible. Still, more than 80% of them forego on the 250 EUR savings in premium. The same two key facts are confirmed when using only within-individual variation in predicted health risk (Online Appendix Table A.4).

This section has highlighted that the gap between the 'frictionless', rational choice model and observed behavior is large and cannot be credibly explained with standard consumer preferences or constraints (e.g., risk aversion, liquidity constraints, moral hazard). However, there are a plethora of models with choice barriers one could write down that could help rationalize the data (e.g., inertia, limited attention, misperceptions). Online Appendix A.6 simulates the choices for a set of alternative models of decision-making that are proposed in the literature. For example, a model with imperfect information and switching costs comes close to replicating the choice patterns we observe. Regardless of the nature of the choice barriers, the evidence shows that the choice barriers need to be large to explain the gap between the observed behavior and the behavior predicted by the model. The next section will show that the gap and, thus, the barriers required to explain the gap, differ substantially across individuals with different demographic, educational, and financial characteristics, indicating an important socio-economic gradient in choice quality.

III Socio-Economic Determinants of Deductible Choice

This section examines how different individual socio-economic factors change deductible choice with respect to health risk and then ranks individuals based on choice quality, conditional on their health, to compare choice quality across socio-economic characteristics.

III.A Socio-Economic Factors

We first study how individuals' deductible choices relate to different socio-economic factors. We do so by (i) presenting non-parametric graphical evidence examining specific characteristics and (ii) with a regression framework that examines the impact of those characteristics conditional on many other characteristics. We rely on a simple OLS regression in a linear probability model:[5]

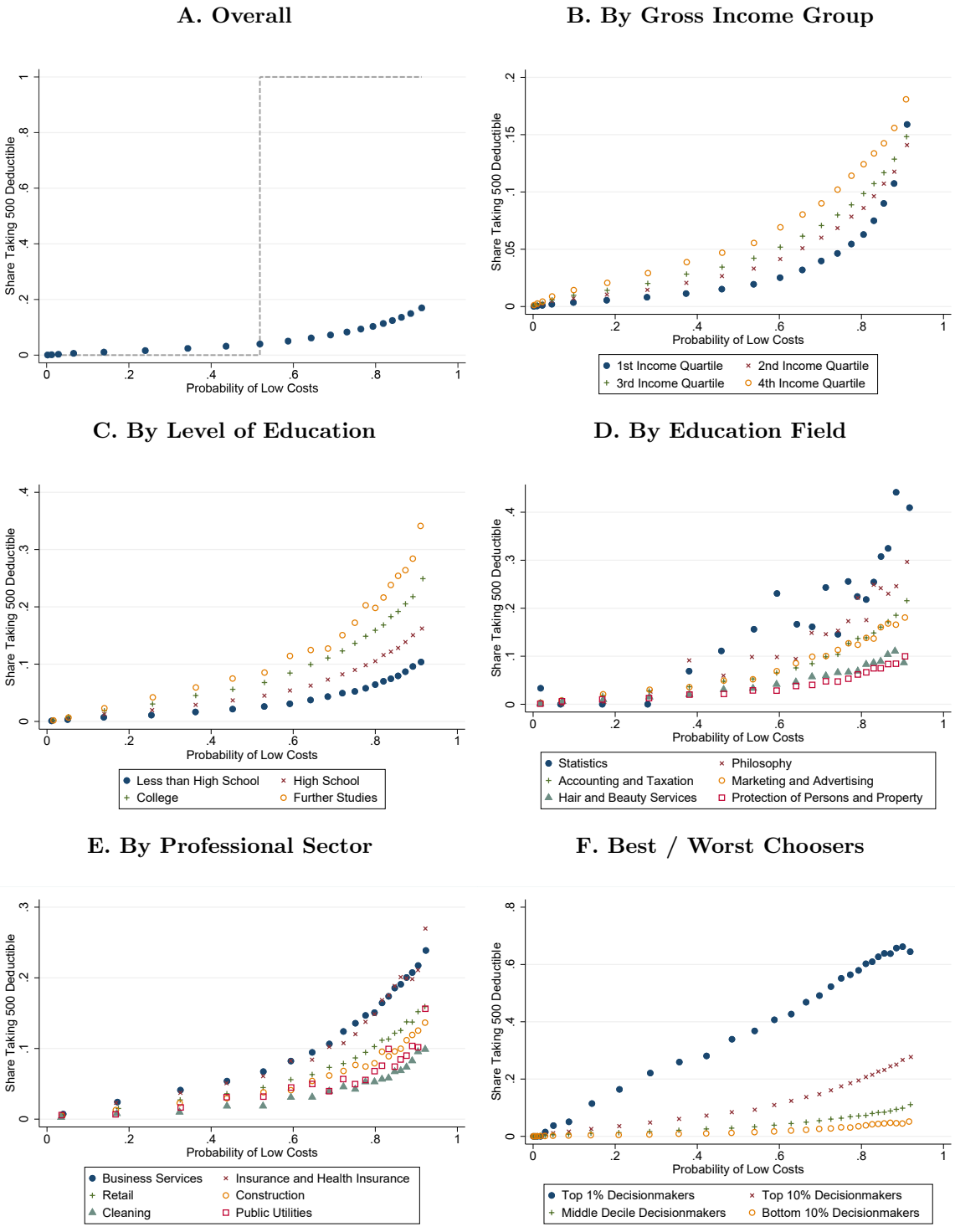
[5] We present alternative specifications, e.g., a probit model, in Online Appendix Table A.5, with little difference in findings.

$$Y = \alpha + \gamma X + [\beta + \nu X]\pi + \epsilon \tag{2}$$

where Y is an indicator variable taking the value of 1 when an individual takes the extra 500 EUR deductible and 0 otherwise, π is the predicted probability of having costs lower than 375 EUR, and X includes all other variables of interest. The primary coefficients of interest are γ and ν. The former captures how different observables affect the intercept, i.e., the average take-up of the 500EUR deductible by individuals who are the sickest (with $\pi_i = 0$). The latter measures how different factors affect the relationship between risk and deductible choice. $\gamma+\nu$ captures the impact on average take-up by individuals who are the healthiest (with $\pi_i = 1$). Each regression also includes year and insurer fixed effects. The insurer fixed effects control for potential differences in insurer marketing / steering, provider network, and/or differences in insurer incremental deductible premium, though as we showed earlier there is limited dispersion in the latter.

Figure 2 plots the relationship between health and deductible take up by income and education in Panels B and C respectively. Both panels show an important gradient in the deductible take-up when people are predicted to be healthy. For example, those in the healthiest predicted risk vigintile with a college degree (i.e., bachelor or master) elect the higher deductible about 25% of the time and those with an advanced degree choose the highest deductible 35% of the time. In contrast, those with less than high school education in the healthiest predicted decile elect the higher deductible only 10% of the time and those with high school education only approximately 15% of the time. For all of these education levels, when people are predicted to be sick they almost never elect the higher deductible. The relations are qualitatively similar when comparing groups with different gross household income (including capital income and government transfers). Higher levels of income are associated with higher take-up of high deductible among the healthiest, though the differences are less pronounced. For example, the average take-up rate of the highest income quartile remains below 20%.

FIGURE 2: DEDUCTIBLE TAKE-UP BY HEALTH RISK AND SOCIO-ECONOMIC DETERMINANTS



Notes: This figure show binned scatter plots of the relationship between the predicted probability of having costs below 375 EUR (staying under the minimum deductible range) and the take-up of the extra 500 EUR deductible. Panel A presents this relationship for the entire population, alongside the risk-neutral threshold at $\pi = .5$ for someone to choose the 500 EUR incremental deductible. The following four panels show deductible choices as a function of predicted health (i) household gross income quartile in Panel B (ii) education level in Panel C (iii) 6 fields of study in Panel D and (iv) 6 professional sectors in Panel E. Panel B excludes individuals with gross income below minimum social assistance, which mostly consists of students, self-employed and households with negative capital income. For Panels D and E, refer to Tables A.6 and A.7 for an overview of the deductible take-up in all fields and sectors, respectively. Panel F presents the relationship for the best and worst cohorts of decisionmakers, conditional on predicted health risk, as estimated in our regressions and defined in Section III.C.

Table 1 presents results from the regression model in equation 2, including baseline demographics, but focusing on income and education. The estimated intercept and slope coefficients for the different characteristics correspond to γ and ν respectively in equation (2).

There is significant and economically meaningful variation in slopes, as expected based on the graphical evidence. The effects, however, are predominantly driven by differences in education. The interaction with the predicted health risk is indeed substantially larger for those with higher education reflecting the fact that individuals are more responsive to their health status in selecting the higher deductible with higher education levels. An individual in good health — *ex ante* very high probability of being low cost — who has completed graduate studies beyond college is 18 percentage points more likely to take up the high deductible than an equivalent person with less than a high school education.

Controlling for other factors, the interaction of income and the gradient of take-up is small in magnitude. The highest income quartile is only about 4 percentage points more likely to take up the high deductible if they are in good health compared to the lowest income quartile, all else equal. The three bottom income quartiles have basically the same take-up rates. Thus, though both income and education have similar and substantive relationships with choice quality independently, when considered together the results show a much stronger impact of education than of income.[6]

In comparison to the variation in slopes, there is generally little variation in the intercepts. There are statistically significant differences in responsiveness to underlying health risks, though the magnitude of the effects are relatively small. As can be expected from the graphical evidence, some of these differences change when relaxing the linearity assumption on the relation between take-up and risk, but they are consistently small. The regressions in Table 1 also include controls for age, gender and household composition on deductible choice, controlling for health risk, income and education level. Despite the relative simplicity of the models we estimate, the overall patterns are very robust to alternative specifications. For brevity, we present those results in Online Appendix Table A.5.

III.B Human vs. Financial Capital

Table 1 demonstrates that the strongest relationship between deductible take-up and observable characteristics is for education level. This is indicative of the potential role of expertise, cognitive ability or information frictions in insurance choices. To shed more light on the role these effects may play we perform the same analysis as above but use richer data on the specific field of education and professional sector of employment.

We first explore this graphically, plotting the relationship between deductible choice and predicted health risk by education field and professional sector in the panels D and E in Figure 2. Since there are many education fields and professional sectors, in the figures we present only 6 specific fields and sectors that are indicative of the broader patterns. Statistics majors are the most responsive to predicted health risk: they choose the additional deductible approximately 43% of the time when they are in the healthiest predicted health bin and choose the additional deductible almost never when they are in the sickest predicted bin. The effect stands in stark contrast to those with training in "Protection of Persons and Property" or "Hair and Beauty Services." Even for the healthiest group in those fields, take-up of the higher deductible is only approximately 10%. Similarly, for professions that are more analytical in nature, deductible choice is also higher for those with low risk. Those in the "insurance and health insurance" industry choose the high deductible option approximate 28% of the time in the healthiest group, compared to only 10% of the time for those in "cleaning."

[6] Online Appendix Figure A.8 provides a graphical illustration of the flattening of the income gradient in deductible take-up when controlling for other factors including education.

Table 1: Deductible Take-Up: Baseline Regression Estimates

	(1) Baseline		(2) Education Field		(3) Professional Sector		(4) Liquidity and Financials	
	intercept	*slope*	*intercept*	*slope*	*intercept*	*slope*	*intercept*	*slope*
2nd Income Quartile	0.004***	-0.007***	0.005***	-0.016***	0.009***	-0.032***	0.005***	-0.022***
3rd Income Quartile	0.004***	0.007***	0.005***	-0.003***	0.010***	-0.018***	0.006***	-0.021***
4th Income Quartile	0.002***	0.039***	0.005***	0.025***	0.011***	0.012***	0.007***	-0.000
High School	-0.011***	0.057***	-0.012***	0.059***	-0.014***	0.055***	-0.010***	0.048***
College Degree	-0.034***	0.165***	-0.033***	0.165***	-0.039***	0.165***	-0.031***	0.152***
Further Studies	-0.047***	0.226***	-0.046***	0.227***	-0.054***	0.236***	-0.045***	0.217***
Statistics			-0.042**	0.247***				
Philosophy			-0.003	0.046***				
Accounting and Taxation			-0.003***	0.024***				
Marketing and Advertising			-0.000	-0.004				
Hair and Beauty			0.007***	-0.035***				
Protection of Persons			0.008***	-0.068***				
Business Services					-0.012***	0.045***		
Insurance					-0.025***	0.078***		
Retail					-0.002***	-0.002*		
Construction					-0.001	-0.018***		
Cleaning					0.003***	-0.033***		
Public Utilities					0.006***	-0.008*		
2nd Net Worth Quartile							0.003***	-0.004***
3rd Net Worth Quartile							0.000*	0.021***
4th Net Worth Quartile							-0.002***	0.061***
Has Savings > 2000EUR							-0.006***	0.028***
Has Mortgage Debt							-0.000	0.005***
Has Other Debt							0.005***	-0.023***
Constant	-0.041***		-0.043***		-0.050***		-0.042***	
Prob. Low Costs		0.098***		0.101***		0.117***		0.094***
Baseline Controls	YES		YES		YES		YES	
Year and Insurer FE	YES		YES		YES		YES	
Observations	57,100,388		30,799,129		32,299,835		57,013,765	

Notes: This table plots coefficients from our regressions studying deductible choice, as explained in Section III. Each variable is interacted with the probability of having low health expenses; the impact on the intercept is reported in the first column, and the impact on the slope in the second column. The dependent variable in all specifications is a dummy that takes value of 1 when the individual takes up the extra 500 EUR deductible. The prob. costs < 375 EUR variable is obtained from our prediction algorithm. All regressions include baseline demographic controls, income quartiles and education dummies. The reference groups are the 1st income quartile and those with education lower than high school respectively. Columns (2)-(4) include additional controls: in Column (2), dummies for six selected educational fields of study, as well as their interactions with health risk. The reference category for field of study is all other fields of study; in Column (3) dummies for six selected professional sectors, as well as their interactions with health risk. The reference category is all other sectors; in Column (4), a dummy for liquidity (household savings>2000EUR), a dummy for having household mortgage debt and other household debt, household net worth quartiles, as well as their interactions with predicted health risk. *** p<0.01, ** p<0.05, * p<0.1 with robust standard errors.

Columns (2) and (3) in Table 1 present the corresponding regression analysis, including baseline controls for predicted health risk, income, education level, age, gender and household structure. Even controlling for these

Table 2: Deductible Take-up and Field of Study

Education Field	(1) Take-up of 500 Deductible	(2) Probability Low Costs	(3) Take-up of 500 Ded. | Being Predictably Healthy
1 **Statistics**	29%	87%	34%
2 Mathematics	21%	85%	27%
3 Physics	21%	91%	26%
4 Architecture and town planning	18%	88%	21%
5 Physical science	18%	82%	22%
6 Earth science	18%	88%	21%
7 **Philosophy and ethics**	17%	82%	21%
8 Medicine	17%	83%	20%
16 Sociology and cultural studies	14%	82%	18%
17 Mining and extraction	14%	91%	17%
18 Economics	14%	84%	17%
19 Humanities and Arts	14%	84%	18%
41 **Accounting and taxation**	11%	78%	14%
42 Agriculture, forestry and fishery	10%	81%	13%
43 **Marketing and advertising**	10%	80%	13%
83 Secretarial and office work	5%	65%	7%
84 **Protection of persons and property**	4%	78%	6%
85 Child care and youth services	4%	66%	6%
86 Computer use	4%	65%	6%
87 **Hair and beauty services**	4%	65%	5%
90 Literacy and numeracy	2%	62%	4%

Notes: For a selection of fields of study, this table shows: in Column (1), the fraction of individuals who take-up the 500 EUR extra deductible, in Column (2), the fraction of individuals with a probability of low costs < 375 EUR, and in Column (3), the fraction of individuals who take-up the 500 EUR extra deductible, conditional on having predicted health costs < 375 EUR. The full list of fields is provided in Online Appendix Table A.6.

other factors, more quantitative / analytic fields of study or profession are more responsive to predicted health when making deductible choices. For example, among the predictably healthy, someone with statistics training is 28.2% more likely to choose a higher deductible than someone with hair and beauty training, controlling for age, income, gender, and education level.

Table 2 presents the relationship between the specific field of study and deductible choice for a broad selection of fields. Columns 1 and 2 show the share taking up the high deductible and the predicted low-cost probability respectively. The primary results of interest are presented in column 3, which shows the rate of take-up of the high deductible among those with a high probability of having low cost — the group for which we expect high adoption under the standard model. The table shows that quantitative fields are grouped at the top of the table, exhibiting greater responsiveness to predicted health risk when making deductible choices, while those in less quantitative fields are grouped at the bottom of the table, exhibiting lower responsiveness.[7] Online Appendix Table A.7 shows a very similar gradient by professional sector.

[7] An exhaustive list of education fields is presented in Online Appendix Table A.6.

Moving from human to financial capital, we can leverage the availability of a range of financial variables in addition to income to confirm the limited importance of household finances. Column (4) in Table 1 shows that - while controlling for demographics, education and income - household liquid savings are positively correlated with deductible take up: having liquid savings of greater than 2000 EUR is associated with a 2.2 percentage point increase in deductible take up for those who are predictably healthy. Liquidity and debt constraints could either increase the demand for insurance (to avoid large expenditures) or reduce the demand for insurance (to avoid paying the premium) (see Ericson and Sydnor (2018)). The sign of the effect we find is consistent with the former explanation. In line with this, we also find that households who are in debt (excluding mortgage debt) are also less likely to take-up the deductible. The effects, however, are small in both cases and only hold for those in good health. Finally, we find that take-up rate for wealthier individuals is higher on average. The differences become meaningful (about 6% percentage points) for the highest wealth quartile. Note that these effects are again fully driven by individuals with better health. That is, wealthier individuals are more responsive to taking the incremental deductible as they become healthier. Hence, rather than capturing wealth effects on insurance choices, this results could be simply indicative of choice barriers for people with fewer financial resources.

III.C Heterogeneity in Choice Quality

We now rank types of individuals based on choice quality and examine the socio-economic factors that predict the best and worst choosers. To evaluate choice quality, we use expected cost savings as our measure of consumer welfare.[8] Using this welfare measure, we find that approximately 52% of consumers would have been ex ante better off with the extra 500 EUR deductible, but less than 7% of consumers chose it. Overall, only 54.4% of individuals chose the cost-minimizing deductible given their predicted health risk. The average amount of money left on the table per individual is 66.2 EUR. While small in absolute value, these savings are roughly half of the total surplus at stake in the decision, which defined as $|250 - (1 - \pi_i)500|$ comes down to 145 EUR on average.

We use this welfare metric together with our regression estimates to rank individuals based on choice quality, conditional on health. That is, we use the estimates of regression (2) with health risk interacted with all socio-economic determinants to predict deductible choice probabilities $d(X_{it}, \pi_{it})$, which we then translate into cost savings conditional on health. We then average the cost savings over the different health risks using the population distribution of predicted health risks. We finally rank individuals from worst to best decision makers based on how much value they are estimated to leave on the table on average across a representative distribution of population health. We provide more detail on this procedure in Online Appendix A.7.2.

Figure 2 (panel F) illustrates the overall heterogeneity in choice quality in the population using this procedure, plotting the responsiveness of deductible choices to health risk for different quantiles of choice quality. The performance of the very best decision makers is striking relative to the others. The take-up rate of the top 1% of decision makers is much steeper, coming close to the 45-degree line, with high and appropriate take-up of the extra deductible when healthy. The median quality decision-maker, on the other hand, comes close to sticking to the compulsory deductible regardless of the underlying health risk, bearing significant expected losses due to over-insurance and doing only slightly better than the bottom 10% of decision-makers.

Table 3 compares the observable characteristics for the best and worst decision makers and paints a telling picture of who is making the best choices in our context. Not surprisingly, we find substantial differences in education, both in terms of the overall level and educational field. For example, those with college education are

[8] Under risk-neutrality, the money-metric welfare loss between a utility-maximizing choice and an observed choice equals the expected cost savings from choosing the deductible that minimize one's expected out-of-pocket expenditures. As we discussed in Section II, our model predictions hardly are robust to alternative models of preferences, so that risk neutral decision making is an excellent approximation of the choice process in our environment.

Table 3: Best and Worst Decision Makers

	Top 5% decisionmakers	*Bottom 5% decisionmakers*		*Top 5% decisionmakers*	*Bottom 5% decisionmakers*
	Mean			Over/underrepresentation	
Demographics			**Education level**		
Gender (male)	62%	28%	Less than high school	0.30	2.99
Age	36	63	High school	0.82	0.33
Has children	59%	34%	College	3.48	0.00
Has a partner	46%	90%	Further Studies	15.57	0.00
			Unknown	0.08	1.05
Financials					
Gross income	105,801	39,347	**Education field**		
Net worth	250,632	4,969	Statistics	19.66	0.00
Has Mortgage Debt	64%	19%	Philosophy	13.14	0.00
Has Other Debt	27%	53%	Economics	6.95	0.01
Has Savings >2000EUR	91%	38%	Tax and administration	3.30	0.01
			Marketing and advertising	1.91	0.06
	Over/underrepresentation		Hair and beauty services	0.64	1.79
Work Status			Protection of persons	0.38	2.24
Student	2.80	0.16			
Retired	0.07	2.47	**Professional sector**		
Self-employed	2.07	0.05	Business services	2.77	0.09
Employee	1.16	0.31	Insurance	2.13	0.07
On Benefits	0.32	1.94	Retail	1.10	0.34
			Construction	0.75	0.24
			Cleaning	0.26	1.40
			Public utilities	1.51	0.11
Observations					11,369,800

Notes: This table presents observable characteristics for the groups that our model considers to be the top 5% and the bottom 5% decision makers. The entries give either the average value of the variable in each group or the ratio of the proportion of consumers with that characteristic in each group relative to the proportion of consumers with that characteristic in the population overall. For example, the group of best decision makers has 6.95 time more economics majors, proportionally, than the population overall.

3.48 times more likely to be in the best decision-making group and with further education are 15.57 times more likely. Individuals with quantitative degrees or occupations are similarly over-represented in this top group. For example, statisticians are 19.66 times and economists 6.95 times more likely to be present in the group of top decision-makers, while those in cleaning are 0.26 times as likely to be in this group.

While we have found that demographic and financial variables provide relatively limited explanatory power in addition to education, the differences between the best and worst decision-makers are striking. The best decision-makers have an average gross income of 105K EUR and net worth of about 250K EUR. The worst decision makers, though, only have an average income of 40k EUR and net worth of 5K EUR. Better decision makers are much more likely to have liquid savings, a mortgage, and much less like to be indebted otherwise. We also find that better decision-makers are significantly younger, more likely to be male and more likely to have children.

IV Conclusion

Using granular data from the Netherlands, we characterized nationwide quality in deductible choices and found that (i) these choices were poor on average, in line with prior work on default options, and (ii) higher SES consumers make better choices than lower SES consumers, with a meaningful impact on realized surplus. Most notably, highly educated individuals who have more quantitative training make better choices than their counterparts, holding constant other key factors like income, net worth, and health risk. A variety of other socio-economic factors have more limited impacts on choice quality, including household income, net worth and liquidity.

Our results on the overall choice quality, and on the choice quality - SES gradient in particular, have important policy implications, as they provide a new perspective on the potential distributional consequences of policies allowing for choice or relying on markets in the provision of public services. At a micro level, it will also be valuable to assess how different policy and technology solutions can improve choices in different markets and regulatory environments, both overall and for lower SES consumers specifically. For example, one could design the choice menu to combat the regressive nature of choice quality by matching the default option closer to the typical low SES consumer than to the typical high SES consumer. Targeted defaults as a function of key consumer characteristics, as discussed in Handel and Kolstad (2015*a*) and Abaluck and Gruber (2023), are another interesting path forward from a policy design standpoint.

While it has not been the focus of this paper to test different decision-making models, we found that default effects in combination with imperfect information about health risks can explain the observed empirical patterns well. However, it is unclear whether the better choices we document for higher-SES consumers are due to increased attention, relative to lower-SES consumers, or due to better active decisions once paying attention (e.g., Brot-Goldberg et al. (2023)). If higher SES consumers are more attentive but not much more sophisticated otherwise, this has important implications for the welfare impacts of policies and our understanding of the potential for insurance markets to deliver value. A field experiment at scale (e.g., Banerjee et al. (2021)) distinguishing between distinct behavioral foundations and/or distinct behaviorally-motivated policies (e.g., Bhargava, Loewenstein and Sydnor (2017)) could provide valuable additional insights, especially if linked to socio-economic data similar to what we use in this study. Finally, in the spirit of understanding the mechanisms underlying the choice quality gradient, in ongoing work (Handel et al. (2023)) we are assessing differential exposure to peers, at work and at home, and the consequent impact on choices.

References

Abaluck, Jason, and Jonathan Gruber. 2011. "Choice Inconsistencies Among the Elderly: Evidence from Plan Choice in the Medicare Part D Program." *American Economic Review*, 101(4): 1180–1210. 2

Abaluck, Jason, and Jonathan Gruber. 2023. "When Less is More: Improving Choices in Health Insurance Markets." *Review of Economic Studies*, 90(3): 1011–1040. 2, 14

Allcott, Hunt, Benjamin B Lockwood, and Dmitry Taubinsky. 2019. "Regressive sin taxes, with an application to the optimal soda tax." *The Quarterly Journal of Economics*, 134(3): 1557–1626. 2

Andersen, Steffen, John Y. Campbell, Kasper Meisner Nielsen, and Tarun Ramadorai. 2020. "Sources of Inaction in Household Finance: Evidence from the Danish Mortgage Market." *American Economic Review*, 110(October): 3184–3230. 2

Baicker, Katherine, Sendhil Mullainathan, and Joshua Schwartzstein. 2015. "Behavioral hazard in health insurance." *The Quarterly Journal of Economics*, 130(4): 1623–1667. 17

Banerjee, Abhijit, Amy Finkelstein, Rema Hanna, Benjamin A Olken, Arianna Ornaghi, and Sudarno Sumarto. 2021. "The Challenges of Universal Health Insurance in Developing Countries: Experimental Evidence from Indonesia's National Health Insurance." *American Economic Review*, 111(9): 3035–63. 14

Barseghyan, Levon, Francesca Molinari, Ted O'Donoghue, and Joshua C. Teitelbaum. 2018. "Estimating Risk Preferences in the Field." *Journal of Economic Literature*, 56(2): 501–64. 5

Bhargava, Saurabh, George Loewenstein, and Justin Sydnor. 2017. "Choose to Lose: Health Plan Choices from a Menu with Dominated Option." *The Quarterly Journal of Economics*, 132(3): 1319–1372. 2, 14

Brot-Goldberg, Zarek C, Amitabh Chandra, Benjamin R Handel, and Jonathan T Kolstad. 2017. "What Does a Deductible Do? The Impact of Cost-Sharing on Health Care Prices, Quantities, and Spending Dynamics." *The Quarterly Journal of Economics*, 132(3): 1261–1318. 5, 16

Brot-Goldberg, Zarek, Timothy Layton, Boris Vabson, and Adelina Yanyue Wang. 2023. "The behavioral foundations of default effects: theory and evidence from Medicare Part D." *American Economic Review*, 113(10): 2718–2758. 2, 14, 18

Brown, Zach Y, Christopher Hansman, Jordan Keener, and Andre F Veiga. 2023. "Information and Disparities in Health Care Quality: Evidence from GP Choice in England." National Bureau of Economic Research Working Paper 31033. 2

Brown, Z., and J. Jeon. 2023. "Endogenous Information and Simplifying Insurance Choice." University of Michigan, Working Paper. 2

Campbell, John Y. 2016. "Restoring Rational Choice: The Challenge of Consumer Financial Regulation." *American Economic Review*, 106(5): 1–30. 1

Chetty, Raj, John N. Friedman, Søren Leth-Petersen, Torben Heien Nielsen, and Tore Olsen. 2014. "Active vs. Passive Decisions and Crowd-Out in Retirement Savings Accounts: Evidence from Denmark." *The Quarterly Journal of Economics*, 129(3): 1141–1219. 1, 2

Chiappori, Pierre-Andre, and Bernard Salanie. 2000. "Testing for Asymmetric Information in Insurance Markets." *Journal of Political Economy*, 108: 56–78. 16

Domurat, Richard, Isaac Menashe, and Wesley Yin. 2021. "The Role of Behavioral Frictions in Health Insurance Marketplace Enrollment and Risk: Evidence from a Field Experiment." *American Economic Review*, 111(5): 1549–74. 2

Dubois, Pierre, Rachel Griffith, and Martin O'Connell. 2020. "How Well Targeted Are Soda Taxes?" *American Economic Review*, 110(11): 3661–3704. 2

Einav, Liran, Amy Finkelstein, and Neale Mahoney. 2021. "The IO of Selection Markets." *Handbook of Industrial Organization*, 5(1): 389–426. 2

Einav, Liran, Amy Finkelstein, and Paul Schrimpf. 2010. "Optimal Mandates and the Welfare Cost of Asymmetric Information: Evidence from the U.K. Annuity Market." *Econometrica*, 78(3): 1031–1092. 5

Einav, Liran, Amy Finkelstein, and Paul Schrimpf. 2015. "The Response of Drug Expenditure to Nonlinear Contract Design: Evidence from Medicare Part D." *Quarterly Journal of Economics*, 130(2): 841–899. 16

Einav, Liran, Amy Finkelstein, Sendhil Mullainathan, and Ziad Obermeyer. 2018. "Predictive Modeling of U.S. Health Spending in Late Life." *Science*, 360(6396): 1462–1465. 1, 4, 5, 6

Enthoven, Alain, Alan Garber, and Sara Singer. 2001. "Near-Universal Coverage Through Health Plan Competition: An Insurance Exchange Approach." *Covering America: Real Remedies for the Uninsured, Editors J. Meyer and E. Wicks*, Washington DC: 155–172. 1

Ericson, Keith. 2014. "Market Design When Firms Interact with Inertial Consumers: Evidence from Medicare Part D." *American Economic Journal: Economic Policy*, 6(1): 38–64. 2

Ericson, Keith Marzilli, and Justin R Sydnor. 2018. "Liquidity Constraints and the Value of Insurance." Working Paper. 5, 12

Fang, Hanming, Michael Keane, and Daniel Silverman. 2008. "Sources of Advantageous Selection: Evidence from the Medigap Insurance Market." *Journal of Political Economy*, 116(2): 303–350. 2

Handel, Benjamin. 2013. "Adverse Selection and Inertia in Health Insurance Markets: When Nudging Hurts." *American Economic Review*, 103(7): 2643–2682. 18

Handel, Benjamin, and Jonathan Kolstad. 2015*a*. "Getting the Most From Marketplaces: Smart Policies on Health Insurance Choice." *Brookings Hamilton Project Discussion Paper 2015-08*. 14

Handel, Benjamin, and Jonathan Kolstad. 2015*b*. "Health Insurance For Humans: Information Frictions, Plan Choice, and Consumer Welfare." *American Economic Review*, 105(8): 2449–2500. 2

Handel, Benjamin, and Joshua Schwartzstein. 2019. "Behavioral Economics and Health Care Markets." *Handbook of Behavioral Economics*. 2

Handel, Benjamin, Jonathan Kolstad, Thomas Minten, and Johannes Spinnewijn. 2020. "The Social Determinants of Choice Quality: Evidence from Health Insurance in the Netherlands." Working Paper No. 27785. 1

Handel, Benjamin, Jonathan Kolstad, Thomas Minten, and Johannes Spinnewijn. 2023. "Liquidity Constraints and the Value of Insurance." Working Paper. 14

Hastings, Justine S, Christopher A Neilson, Anely Ramirez, and Seth D Zimmerman. 2016. "(Un) informed college and major choice: Evidence from linked survey and administrative data." *Economics of Education Review*, 51: 136–151. 1

Ho, Kate, Joseph Hogan, and Fiona Scott Morton. 2017. "The Impact of Consumer Inattention on Insurer Pricing in the Medicare Part D Program." *The RAND Journal of Economics*, 48(4): 877–905. 2, 18

Ito, Koichiro. 2015. "Do Consumers Respond to Marginal or Average Price? Evidence from Nonlinear Electricity Pricing." *American Economic Review*, 104(2): 537–563. 1

Ketcham, Jonathan, Claudio Lucarelli, Eugenio Miravete, and Christopher Roebuck. 2012. "Sinking, Swimming, or Learning to Swim in Medicare Part D?" *The American Economic Review*, 102(6): 2639–2673. 2

Kőszegi, Botond, and Matthew Rabin. 2007. "Reference-dependent risk attitudes." *American Economic Review*, 97(4): 1047–1073. 18

Kroneman, Madelon, Wienke Boerma, M van den Berg, P Groenewegen, J de Jong, and E. van Ginneken. 2016. "Netherlands: Health System Review." *Health Systems in Transition*, 18(2): 1–240. 2

Mullainathan, Sendhil, and Eldar Shafir. 2013. *Scarcity: Why having too little means so much.* Macmillan. 1

Neilsen, Christopher. 2017. "Targeted Vouchers, Competition Among Schools, and the Academic Achievements of Poor Students." Princeton Working Paper. 1

Newhouse, Joseph P. 1993. *Free for All? Lessons From the RAND Health Insurance Experiment.* Harvard University Press. 5, 16

Polyakova, Maria. 2016. "Regulation of Insurance with Adverse Selection and Switching Costs: Evidence from Medicare Part D." *American Economic Journal: Applied Economics*, 8(3): 165–195. 2

Remmerswaal, Minke, Jan Boone, and Rudy Douven. 2023. "Minimum generosity levels in a competitive health insurance market." *Journal of Health Economics*, 102782. 17

Sydnor, Justin. 2010. "Over(insuring) Modest Risks." *American Economic Journal: Applied Economics*, 2(4): 177–199. 18

Van Winssen, KPM, RC Van Kleef, and WPMM Van de Ven. 2015. "How profitable is a voluntary deductible in health insurance for the consumer?" *Health Policy*, 119(5): 688–695. 2

Van Winssen, KPM, RC Van Kleef, and WPMM Van de Ven. 2016. "Potential determinants of deductible uptake in health insurance: How to increase uptake in The Netherlands?" *The European Journal of Health Economics*, 17(9): 1059–1072. 2