American Economic Review 2025, 115(3): 772–822
https://doi.org/10.1257/aer.20231133

Do Ordeals Work for Selection Markets? Evidence from Health Insurance Auto-Enrollment†

By Mark Shepard and Myles Wagner*

Are application hassles, or "ordeals," an effective way to limit public program enrollment? We provide new evidence by studying (removal of) an auto-enrollment policy for health insurance, adding an extra step to enroll. This minor ordeal has a major impact, reducing enrollment by 33 percent and differentially excluding young, healthy, and economically disadvantaged people. Using a simple model, we show adverse selection—a classic feature of insurance markets—undermines ordeals' standard rationale of excluding low-value individuals since they are also low-cost and may not be inefficient. Our analysis illustrates why ordeals targeting is unlikely to work well in selection markets. (*JEL* D82, G22, H75, I13, I18)

Should enrolling in public programs be easy or hard? The desirability of enrollment hassles, or "ordeals," for social programs is a classic—and controversial—question in public economics. On the one hand, there is substantial concern about incomplete take-up of programs intended to help the poor (Currie 2006). A growing body of work argues that the bureaucracy, paperwork, and "administrative burden" of enrollment is a major driver of low take-up and source of frustration with and mistrust of government (Herd and Moynihan 2018).

On the other hand, a classic line of thinking in economics argues that ordeals can be useful ways to *target* assistance toward those who need or value it most (Nichols and Zeckhauser 1982; Besley and Coate 1992). The basic idea follows from the

* Shepard: Harvard Kennedy School and NBER (email: mark_shepard@hks.harvard.edu). Wagner: Department of Economics, The Ohio State University (email: wagner.1858@osu.edu). Devin Pope was the coeditor for this article. We thank Amina Abdu, Kendra Singh, Mike Yepes, and Olivia Zhao for excellent research assistance. We thank the editor and three anonymous referees for extremely thoughtful comments. We thank Jason Abaluck, Manasi Deshpande, Keith Ericson, and Ben Handel for thoughtful and constructive discussant comments. For helpful feedback and suggestions, we thank Hunt Allcott, Marcella Alsan, Chris Avery, Peter Blair, Zarek Brot-Goldberg, Sam Burn, Amitabh Chandra, Leemore Dafny, Amy Finkelstein, Peter Ganong, Josh Gottlieb, Jon Gruber, Gordon Hanson, Nathan Hendren, Alex Imas, Tim Layton, Jeff Liebman, Lee Lockwood, Amanda Kowalski, Brigitte Madrian, Sendhil Mullainathan, Matthew Notowidigdo, Carol Propper, Wesley Yin, Richard Zeckhauser, and seminar participants at the AEA meetings, ASHEcon, Boston-Area IO Conference, Brown University, Covered California, Harvard Kennedy School, Harvard-MIT-BU Health Economics, Imperial College London, Massachusetts Health Connector, Queen Mary University, USC Schaeffer, and NBER Health Care, Public Economics, and Economics of Aging meetings. We thank the Massachusetts Health Connector (particularly Michael Norton and Marissa Woltmann) for assistance in providing and interpreting the data. We gratefully acknowledge data funding from Harvard's Lab for Economic Applications and Policy and research support from Harvard Kennedy School's Rappaport Institute for Public Policy, Harvard's Milton Fund, and the National Institute on Aging Grant Number T32-AG000186 (via the NBER). The research protocol was approved by the IRBs of Harvard University(Protocol 23463) and the NBER (Ref 18_281).

† Go to https://doi.org/10.1257/aer.20231133 to visit the article page for additional materials and author disclosure statement(s).

logic of revealed preference. Ordeals work like a nonfinancial "price" of enrolling, and as in standard markets, prices screen out people with low value (demand) for a program. By excluding low-value types, the government saves money and can redirect aid toward those who need it most. This influential "self-targeting" idea has spawned an active empirical debate, with some research finding that it holds in practice (Alatas et al. 2016; Dupas et al. 2016), while other work argues that behavioral frictions may undermine its validity (Bhargava and Manoli 2015; Finkelstein and Notowidigdo 2019; Deshpande and Li 2019). Importantly, the debate has been framed almost entirely around the self-targeting question: Do ordeals effectively screen out *low-value* or *low-need* types in a given setting?

In this paper, we ask whether this is the right way to think about targeting in programs where people vary not just in value or need but also in their *costs*. We observe that many programs—and especially insurance programs—share a key feature of "selection markets" that have been widely studied in the economics literature (Einav, Finkelstein, and Mahoney 2021). In these settings, enrollee costs vary substantially and tend to be *correlated* with value, often because both are driven by the same underlying factor, like risk. For instance, in our health insurance data, the highest-risk (sickest) 10 percent of enrollees incur 15 *times* higher medical costs than the healthiest 10 percent (about \$1,400 versus \$90 per month). Moreover, the healthy are likely to value insurance less, precisely because they have fewer medical needs and use less care. This example illustrates the key correlation in settings with adverse selection: low-value types also tend to be low-cost.

Our paper's central conceptual point is that adverse selection tends to weaken, and when strong enough undermine, the classic self-targeting case for ordeals. When low-value enrollees are also low-cost, excluding them may yield minimal, or even negative, targeting gains. The key question in selection markets is not whether ordeals screen on value, but whether they screen *more strongly* on social value than on costs. This question is theoretically ambiguous and does not follow from the standard revealed preference logic for ordeals.

We formalize this argument with a mix of theory and evidence from a public health insurance program. We use a natural experiment to study descriptively *how much* ordeals matter for take-up and which types of people they screen out. We find that even minor hassles lead to major reductions in take-up among an otherwise uninsured low-income population. Consistent with adverse selection, the excluded group is differentially younger, healthier, and poorer, suggesting ordeals screen out people with low private value (demand) but also low cost of insurance.[1] Using an empirical model estimated with our data, we find that ordeals worsen targeting efficiency, despite successfully screening out low-value types. More generally, we show that adverse selection works alongside behavioral frictions to weaken the (revealed preference) link between demand and efficiency that is key to self-targeting. This makes ordeals relatively poorly suited tools for adverse selection markets.

We begin the paper (in Section I) with a general framework to formalize these ideas about ordeals targeting in selection markets. Ordeals improve welfare if they yield "gains from targeting"—the ability to include efficient (*social value* $>$ *cost*)

[1] This also aligns with the groups most likely to be among the 28 million uninsured in the United States today (Tolbert et al. 2024).

and exclude inefficient (*social value* $<$ *cost*) types—sufficient to outweigh any direct losses from their hassle or administrative costs. We show that targeting gains can be visualized in simple supply/demand-like graphs of marginal value/cost versus quantity enrolled as ordeals vary, analogous to the approach of Einav, Finkelstein, and Cullen (2010) for visualizing welfare in selection markets. As in their graphs, adverse selection implies that the "marginal cost" curve is not flat (as in a nonselection market) but *slopes downward* alongside marginal value, reflecting the positive value-cost correlation driven by enrollee risk. This shrinks the gains from targeting, reflected in a smaller area between marginal value and cost curves above and below their intersection.

We formalize this reduction in what we call the "adverse selection tax," which equals the coefficient in a regression of enrollee (net) cost on social value, or $\hat{\beta} = \text{cov}\left[C_i^{Net}, V_i^{Soc}\right]/\text{var}\left[V_i^{Soc}\right] = \rho \cdot \sigma_C/\sigma_V$.[2] When adverse selection is sufficiently strong (roughly, when $\hat{\beta} > 1$), the marginal cost curve becomes steeper than marginal value, and ordeals induce "*backward sorting*" into insurance even when they correctly sort on value. This idea—analogous to the insights of Marone and Sabety (2022) for menu design and sorting with prices—shows the limits of choice and self-targeting mechanisms in adverse selection markets where demand and efficiency are often misaligned.[3]

In addition, we show a second reason adverse selection tends to undermine ordeals: it makes it more likely that the optimal outcome is *universal*—enrolling or excluding everyone—rather than targeted. We call this second idea "*optimal universality*." Graphically, it occurs when the marginal value (MV) curve lies entirely above or below marginal costs (MC), so the two do not intersect. This is more likely when both MV and MC have a similar downward slope because value and cost are strongly correlated. For instance, consider a case where social value and net enrollee cost align perfectly: $V_i^{Soc} = \delta \cdot C_i^{Net}$. In this case, net welfare $(= V_i^{Soc} - C_i^{Net})$ equals $\left(\delta - 1\right) C_i^{Net}$ for all i, which is uniformly positive or negative depending on $\delta \gtrless 1$. This example illustrates the key idea of optimal universality: a strong value-cost correlation makes it more likely that targeting using ordeals is counterproductive because universal outcomes are superior.

Having developed this framework, we next turn to an empirical analysis of ordeals that lets us both estimate the key model parameters and also learn descriptively about ordeals' impact for health insurance programs. Our empirical setting is the Massachusetts health insurance exchange, a program offering subsidized insurance to low-income people without access to other coverage.[4] The program featured a

[2] Here, $\rho = \text{corr}\left[C_i^{Net}, V_i^{Soc}\right]$, $\sigma_C = \text{std}\left(C_i^{Net}\right)$, and $\sigma_V = \text{std}\left(V_i^{Soc}\right)$, all evaluated across potential enrollees (i). See Section I for the formal definition of social value and net public cost (which is net of fiscal externalities). The adverse selection tax is zero if enrollee costs do not vary ($\sigma_C = 0$) or are uncorrelated with value ($\rho = 0$), and it grows as both of these increase relative to the variation in value.

[3] Conversely, *advantageous* selection—where low-value types have high costs—strengthens the case for ordeals targeting. Because advantageous selection is less common, we do not discuss it in detail. Two settings where it has been found are long-term care insurance (Finkelstein and McGarry 2006) and Medicare supplemental coverage ("medigap") (Fang, Keane, and Silverman 2008).

[4] We study the pre-Obamacare (or ACA) exchange, which operated from 2007 to 2013 and was called Commonwealth Care (or "CommCare"). As a model for the ACA exchanges that followed, CommCare has been a rich source of evidence on demand, competition, and the impact of policies in health insurance markets (see Chandra, Gruber, and McKnight 2011, 2014; Finkelstein, Hendren, and Shepard 2019; Jaffe and Shepard 2020; McIntyre, Shepard, and Wagner 2021; Shepard 2022; Shepard and Forsgren 2023).

unique source of variation in the complexity of enrollment, driven by changing use of an auto-enrollment policy for the program's poorest individuals, who qualified for free insurance. Prior to 2010, the program required only that these individuals *apply* for coverage, submitting paperwork with information to verify eligibility. Approved applicants were then contacted and asked to choose among several plans offered by different insurers (all of which were free). But if they failed to respond—something that occurred surprisingly often—the program *auto-enrolled* them into a plan using a simple algorithm. In essence, this policy used defaults or "choice architecture" (Thaler 2018) to streamline take-up and prevent people from falling through the cracks of the system.

Starting in 2010, the program suspended auto-enrollment. Nonresponsive, or "passive," individuals were no longer enrolled by default; instead, their default became *non-enrollment*. Effectively, this change added an extra step (active plan choice) to the required take-up process. Although not intended to be onerous—people could choose by phone, mail, or online, and all plans remained free—this change is an example of the type of small take-up friction that is common in many US safety net programs.

We use this variation to estimate the causal effect of the ordeal by studying enrollment changes around the 2010 policy shift. We use a difference-in-difference design, comparing changes in new enrollment for the low-income (treatment) group for whom auto-enrollment stops in 2010 versus a slightly higher-income (control) group for whom it was not used throughout. Our rich administrative data let us observe who enrolled actively versus passively prior to 2010, and we can also infer the characteristics of marginal enrollees from compositional changes in enrollment around 2010.

This analysis yields two main findings. First, adding a minor ordeal leads to major reductions in health insurance take-up. Prior to 2010, one-third of low-income new enrollees join the exchange passively via auto-enrollment. When the policy is suspended in 2010, the flow of new enrollment falls by a nearly identical 33 percent. The decline is immediate and persistent, with parallel pre-trends and no concurrent changes for the control group.[5] We also see no evidence of an uptick in active enrollment in 2010, suggesting that passive individuals are unlikely to be deliberately choosing nonresponse (e.g., because they know they will be auto-enrolled). Rather, when subjected to a small hassle, about one-third of eligible individuals simply fail to take up health insurance.

This effect is quite large. For instance, it is similar to the impact of a \$470 (or 57 percent) annual premium increase based on prior evidence (Finkelstein, Hendren, and Shepard 2019) and 1.25–2 times larger than the impact of Massachusetts's uninsurance penalty (Chandra, Gruber, and McKnight 2011). It is an order of magnitude larger than the 1–4 percentage point effects observed from lower-touch "nudges" (like outreach and assistance) in recent work on health insurance (Goldin, Lurie, and McCubbin 2021; Domurat, Menashe, and Yin 2021; Ericson et al. 2023). The

[5] Further evidence comes from a temporary reinstatement of the auto-enrollment policy in late 2010. Consistent with the policy having a causal effect, we find that new enrollment spikes back up to its pre-2010 level, then falls back down when auto-enrollment is again suspended in early 2011.

findings suggest that *fully automatic* enrollment—not just incremental incentives and nudges—may be a key step to further reduce uninsurance in the United States.

Our second descriptive finding is that ordeals differentially screen out low-risk individuals, consistent with adverse selection. Relative to active enrollees, passive enrollees are younger and healthier (e.g., 33 percent less likely to be chronically ill) and especially likely to be young men age 19–34. They incur 44 percent lower medical spending per month—most of which (a 36 percent gap) is predictable by their age and diagnosis risk factors. Because of their lower costs, excluding passive enrollees results in a 15 percent higher average-cost risk pool of enrollees.

We also examine the distributional equity implications of ordeals. We find that passive enrollees are more likely to be very low income, to live in disadvantaged neighborhoods, and to live near safety net hospitals and clinics. This is consistent with ordeals differentially impacting the poor (Bertrand, Mullainathan, and Shafir 2004; Mullainathan and Shafir 2013). But it is also consistent with evidence that the poor have lower *demand* for health insurance, potentially because of access to charity care when uninsured (Finkelstein, Hendren, and Luttmer 2019).

Why does a seemingly small hassle matter so much for enrollment? This fact is striking because the benefits of forgone health insurance are likely meaningful.[6] Our evidence is most consistent with behavioral frictions like inattention, forgetting to act, or simply "going with the flow" in insurance choices.[7] We examine but find little evidence of other explanations, including stigma or unawareness of the program (since everyone in our sample has already applied for coverage), "choice overload" that leads to passivity (Iyengar and Kamenica 2010), or passive enrollees already having another form of duplicate insurance.[8]

The final portion of our paper applies the ordeals welfare framework to our setting using the auto-enrollment natural experiment. We specify a rich model allowing for the key features of insurance problem, including heterogeneity in enrollee value (demand), insurer cost (based on medical claims data), and externalities of insurance via savings on uncompensated care. The key empirical challenge—common to most analyses of ordeals—is to infer enrollee value of insurance, given the nonprice nature of the take-up barrier. We address this challenge by estimating demand among a higher-income segment of exchange enrollees who face positive prices, drawing on RD-style premium variation used in prior work (Finkelstein, Hendren, and Shepard 2019). We then project these demand estimates onto the lower-income population at the level of key observables (cells of age, sex, and medical risk scores). We consider various assumptions for the role of unobserved preferences, as well as alternate methods of estimating value directly from observed medical use in our claims data.

[6] Passive enrollees (while healthier than average) do use significant medical care and experience medical shocks. Based on our model estimates and prior work on the value of health insurance (Finkelstein, Hendren, and Luttmer 2019), coverage should be worth about \$550 to \$1,300 for an average passive enrollee over a typical year-long spell. This is comparable to forgone benefits from failure to take up the EITC or SNAP (Bhargava and Manoli 2015; Finkelstein and Notowidigdo 2019).

[7] Consistent with these ideas, we find that passive nonresponse is more common among immigrants (who may face language barriers), people with signs of address instability, and people transitioning into the exchange from Medicaid (which may involve greater confusion because Medicaid's process is different).

[8] We test this using the state's All Payer Claims Database, where we can see the near universe of health insurance coverage. We see very low rates ($<$ 4 percent) of duplicate enrollment in the exchange plus other coverage and no meaningful change in duplication rates around the end of auto-enrollment in 2010.

This exercise yields three main results. First, ordeals do screen out lower-value enrollees. In our baseline estimate, passive enrollees have a private (social) value of coverage that is 28 percent (34 percent) lower than active types. This finding, which is consistent with the classic ordeals rationale of self-targeting, is robust across a wide range of specifications we consider.

Second, adverse selection substantially reduces, or even reverses, the ordeal's targeting gains. Our estimates suggest substantial cost variation and a strong value-cost correlation that implies an "adverse selection tax" that is large and often exceeds 100 percent. Correspondingly, the value-cost *ratio* of passive enrollees is similar to or (in our main specification) higher than active enrollees, suggesting that ordeals induce counterproductive "backward sorting" into insurance. We also examine the robustness of this conclusion to varying distributional equity goals, by applying a social welfare weight $\mu > 1$ to enrollee welfare. We find that with even modest equity concerns ($\mu > 1.3$), it becomes optimal to enroll *both* active and passive individuals. The ordeal is still nonoptimal, but not because sorting is backward, rather because the optimal outcome is universal.

Finally, we use the model to compare auto-enrollment versus subsidies as ways of expanding take-up. We find that the two have similar targeting properties—both enroll a similar young, healthy, and low-cost population—but that auto-enrollment is much more cost-effective because it does not require new spending on inframarginal enrollees. We find that each extra $1 million in public spending covers 55–66 percent more people if used for auto-enrollment rather than subsidies.

Related Literature.—Our paper contributes to three main strands of literature. The first studies the nature of ordeals targeting for social programs. Starting from the classic analysis of Nichols and Zeckhauser (1982), the debate has centered around whether ordeals screen out people who value or benefit less from assistance (e.g., Alatas et al. 2016; Dupas et al. 2016; Finkelstein and Notowidigdo 2019) or who benefit just as much but have less ability to navigate a complex process (e.g., Bhargava and Manoli 2015; Deshpande and Li 2019; Homonoff and Somerville 2021). This debate is part of a broader literature asking when nonprice targeting is valuable in social programs (e.g., Kleven and Kopczuk 2011; Lieber and Lockwood 2019). We provide evidence in a new and important setting (health insurance) and highlight that the classic debate misses the key role of cost heterogeneity and adverse selection for this question.

Second, our paper contributes to work evaluating "nudges" to increase take-up of social programs, including health insurance (Goldin, Lurie, and McCubbin 2021; Domurat, Menashe, and Yin 2021; Banerjee et al. 2021; Ericson et al. 2023). Our results suggest a much larger impact of fully *removing* hassles by changing the default to auto-enrollment. This complements prior work on the large impact of auto-enrollment in other settings (e.g., Madrian and Shea 2001; Chetty et al. 2014),[9] as well as evidence that defaults create inertia in choosing *among* insurance plans

[9] Recent work on 401(k) pensions by Choukhmane (2021) finds that while auto-enrollment has a large *initial* impact on enrollment and savings, people who are not auto-enrolled largely catch up by saving more in the future. Unlike pensions, health insurance is a domain where failure to enroll can have immediate repercussions if an individual gets sick and incurs medical bills. This suggests auto-enrollment is likely to be a consequential policy for health insurance.

(Handel 2013; Ericson 2014; Polyakova 2016; Brot-Goldberg, Layton et al. 2023). Default effects are a key example of a broader set of "choice frictions" that have been shown to be prevalent in health insurance markets (Abaluck and Gruber 2011, 2023; Bhargava, Loewenstein, and Sydnor 2017). Our paper shows that defaults are also important policies for insurance take-up.

Finally, our paper contributes to the literature asking why uninsurance is so persistent in the United States. A large prior literature has analyzed the impact of financial prices and subsidies for incomplete take-up (Gruber 2008; Dague 2014; Frean, Gruber, and Sommers 2017; Finkelstein, Hendren, and Shepard 2019). We show that ordeals and hassles are also likely to be a key barrier, given the United States' fragmented and nonautomatic health insurance system. There is growing interest in the role of complexity, transaction costs, and "administrative burden" in shaping enrollment, with emerging evidence that this matters for Medicaid take-up (Aizer 2007; Arbogast, Chorniy, and Currie 2022; Wu and Meyer 2023) and for ACA health insurance marketplaces (Drake et al. 2023; McIntyre, Shepard, and Layton 2024). We show, likewise, that imposing even modest hassles leads to non-enrollment by a large share of people, especially the young, healthy, and poor, who are disproportionately uninsured today. Our results suggest that as long as take-up is voluntary, getting to universal coverage will likely require some form of auto-enrollment. They also illustrate the surprising power of a feasible form of auto-enrollment that has recently been considered or implemented in several states' ACA exchanges.[10]

Outline of Paper.—Section I presents a conceptual framework for ordeals targeting with adverse selection. Section II discusses the setting, the auto-enrollment policy, and our data. Section III shows our main results on enrollment impacts, and Section IV presents targeting results. Section V implements our empirical model using the auto-enrollment variation. Finally, Section VI concludes.

I. Conceptual Model: Adverse Selection and Ordeals Targeting

In this section, we present a simple framework for the economics of ordeals in programs characterized by adverse selection, that is, where enrollee value and costs are positively correlated. Adverse selection is a classic feature of insurance, where individual risk (e.g., health status) is the primary driver of the value-cost correlation. But it is also relevant more generally for transfer programs with varying benefit amounts (e.g., by income or family size) since people who receive smaller benefits also cost less to the government. Our central point is that adverse selection reduces—and may even reverse—the efficiency of the standard ordeals rationale of screening out *low-value* types since low-value enrollees may not be *inefficient* enrollees.

This section formalizes this argument using a simple model based on the classic insights of Nichols and Zeckhauser (1982), as well as the more recent ordeals framework of Finkelstein and Notowidigdo (2019). Our key innovation is to connect ordeals to the economics of selection markets, visualized using the graphical

[10] This includes Massachusetts, which reinstated a similar form of auto-enrollment in April 2022, partly based on discussions with them about this research.

framework of Einav, Finkelstein, and Cullen (2010). Our analysis also connects to recent insights about "backward sorting" in selection markets (Marone and Sabety 2022), in which prices also lead to inefficient sorting between insurance options.

A. *Model Setup*

Consider a population of individuals who qualify for a public program—in our setting, free health insurance—but have not yet enrolled. For each individual i, the program generates social value of

$$V_i^{Soc} = \mu_i W_i + E_i, \tag{1}$$

where W_i is the program's private welfare to enrollee i (willingness to pay, or WTP), μ_i is the marginal social welfare weight on individual i (capturing distributional equity concerns), and E_i is the social value of any externalities from i's participation in the program. A Kaldor-Hicks efficiency welfare criterion would involve $\mu_i = 1$ for all i, but it may be natural to think of $\mu_i > 1$ for safety net programs where beneficiaries are lower income. For our empirical work, we simplify by treating μ_i as a constant μ for everyone who qualifies for the program, but in principle, μ_i could vary across eligible groups to capture distributional goals.

For individual i, the program involves net government cost $C_i^{Net} = C_i - FE_i$, which equals direct costs (C_i) minus any offsetting fiscal externalities (FE_i).[11] We assume $C_i^{Net} > 0$ so that there is a real fiscal trade-off of expanding enrollment. Both social value and cost may vary across individuals, potentially creating a rationale for targeting.

The government seeks to target enrollment to maximize total social benefits net of costs. Mathematically, if $A_i \in \{0,1\}$ indicates whether i is enrolled, the government seeks to maximize net social welfare, or $SW = \sum_i \left(V_i^{Soc} - C_i^{Net}\right) \cdot A_i$. We define γ_i as the net contribution to social welfare of enrolling individual i:

$$\text{(Net Welfare)} \qquad \gamma_i = V_i^{Soc} - C_i^{Net} = \left(\mu_i W_i + E_i\right) - C_i^{Net}. \tag{2}$$

If the government had full information, it would optimally enroll everyone for whom $\gamma_i \geq 0$ and exclude those with $\gamma_i < 0$. Equivalently, if we define $R_i \equiv V_i^{Soc}/C_i^{Net}$ as the enrollee's "social value-cost ratio," the government optimally enrolls everyone with $R_i \geq 1$ and excludes those with $R_i < 1$.[12] The metric γ_i is a useful targeting index that shows how a government would optimally prioritize enrollment with

[11] In our empirical setting we think of these variables as follows. $W_i > 0$ is the benefits of insurance to the individual; $C_i > 0$ is the government's direct subsidy cost for insuring them; and $E_i, FE_i \geq 0$ are savings on (uninsured) uncompensated care borne by private hospitals (E_i) and the government (FE_i). The nature of C_i depends on how insurance is provided. We assume either direct public provision (relevant in programs like Medicaid) or zero-profit contracting with private insurers (which we find to be roughly true in the Massachusetts exchange), which implies that C_i equals i's expected insured medical costs.

[12] The social value-cost ratio is closely related to the marginal value of public funds (MVPF) metric (Hendren 2016), which is also a (policy-level) benefit-cost ratio.

full information. In practice, however, the government has limited information, so it must use blunt policies like ordeals, which we turn to next.

Ordeals and Take-Up.—The government has access to a screening mechanism—in our setting, an ordeal—that it uses to limit take-up. Ordeals work by imposing a "friction," $\eta_i \geq 0$, that individuals must overcome to enroll. The friction may vary across individuals and could involve both real costs (e.g., the time and effort of completing paperwork) and behavioral frictions that limit take-up (e.g., inattention). We assume the government can adjust the "intensity" of the ordeal through its policy choices (e.g., how much paperwork to impose). A simple specification that captures this idea is $\eta_i = \sigma \cdot h_i$, where $\sigma \geq 0$ is the ordeal's intensity (a policy choice) and $h_i \geq 0$ captures a person's experienced hassle cost per unit ordeal. The policy of no ordeal is equivalent to setting $\sigma = 0$.

In addition to the ordeal, people may have behavioral biases that affect demand, e.g., biased beliefs about their risk type (Spinnewijn 2017). We denote the bias by ε_i, and the utility governing take-up as $U_i \equiv W_i - \varepsilon_i$, where $\varepsilon_i > 0$ captures undervaluation and $\varepsilon_i < 0$ overvaluation. With the ordeal in place, people take up the program if

$$\text{(3)} \qquad \text{(Take-Up)} \quad U_i = \underbrace{W_i}_{\text{True WTP}} - \underbrace{\varepsilon_i}_{\text{Bias}} \geq \underbrace{\sigma \cdot h_i}_{\text{Ordeal friction}}.$$

A comparison of the conditions for who should optimally enroll ($\gamma_i \geq 0 \Leftrightarrow \mu_i W_i + E_i - C_i^{Net} \geq 0$) versus actual take-up ($W_i - \varepsilon_i - \sigma h_i \geq 0$) shows that there may be both under- and overenrollment among differing groups. All else equal, underenrollment is more likely for disadvantaged groups (with high welfare weights, $\mu_i > 1$), for people with positive externalities ($E_i > 0$) or undervaluation bias ($\varepsilon_i > 0$), and for people with low cost (C_i^{Net}) relative to WTP. Overenrollment is more likely for the opposite cases. Imposing an ordeal improves targeting if it reduces overenrollment more than it exacerbates underenrollment, in a sense that we formalize below.[13]

We denote the share of people who enroll given an ordeal of intensity σ as $D(\sigma) = \Pr(W_i - \varepsilon_i \geq \sigma h_i)$. The share excluded is $1 - D(\sigma)$. The ordeal splits potential enrollees into two groups. For any variable X_i (e.g., value or cost), we denote averages for screened-in enrollees as $\bar{X}_1(\sigma) \equiv E[X_i | W_i - \varepsilon_i \geq \sigma h_i]$, and for excluded individuals as $\bar{X}_0(\sigma) \equiv E[X_i | W_i - \varepsilon_i < \sigma h_i]$.

[13] One way to understand misallocation is to define the "wedge" between optimal enrollment versus take-up utility (absent the ordeal) as

$$\text{(4)} \qquad \Delta_i \equiv \gamma_i - U_i = [(\mu_i - 1)W_i + E_i + \varepsilon_i] - C_i^{Net}.$$

In an ideal world, this take-up wedge would be zero, ensuring that people enrolled if and only if $\gamma_i \geq 0$. Imposing an ordeal works like a reduction in take-up utility, so it shifts the wedge from Δ_i to $(\Delta_i + \sigma h_i)$. This will tend to improve welfare if the distribution of $(\Delta_i + \sigma h_i)$ is closer to zero than the distribution of Δ_i. This point is related to the result of Allcott et al. (2022) that "nudges" tend to improve welfare if they reduce the *variance* of net wedges between socially optimal and actual consumption of a good.

In addition to their impact on take-up, ordeals may impose "direct" or "excess" costs, including both hassle/psychological costs to enrollees and administrative costs to the government. The nature of these costs depends on the specifics of the ordeal and the model of behavior and welfare (Ericson 2020).[14] Rather than specify it in detail, we write the ordeal's total direct/excess cost as a general function, $L(\sigma) \geq 0$, which we assume is weakly positive. As we show below, direct costs are separable from the effect of ordeals on social welfare via *targeting* (who is enrolled versus excluded), which is our focus in this paper.

B. *When Are Ordeals Optimal?*

We now lay out the general conditions under which an ordeal is desirable, which we relate to adverse selection in the next subsection. Consider an ordeal of strength σ that generates enrollment $D(\sigma)$. Net social welfare under this policy is

$$SW_{Ordeal}(\sigma) = D(\sigma) \cdot \underbrace{\left[\bar{V}_1^{Soc}(\sigma) - \bar{C}_1^{Net}(\sigma)\right]}_{=\bar{\gamma}_1(\sigma)} - L(\sigma), \tag{5}$$

where $L(\sigma) \geq 0$ is the total direct cost of the ordeal via hassles and administrative costs. To be welfare improving, an ordeal must at least be superior to two trivial alternate policies:

- **Shutting down the program** (**no enrollment**), which results in $SW_0 = 0$, and
- **Enrolling everyone** (**full enrollment**), which results in $SW_1 = E[\gamma_i] \equiv \bar{\gamma}$.

Relative to these alternatives, the ordeal's extra social welfare is $\Delta SW_{Ordeal}(\sigma) = SW_{Ordeal}(\sigma) - \max\{0, \bar{\gamma}\}$, or:[15]

$$\Delta SW_{Ordeal}(\sigma) = \underbrace{\min\left\{D(\sigma)\,\bar{\gamma}_1, \left[1 - D(\sigma)\right] \cdot (-\bar{\gamma}_0)\right\}}_{\text{Gains from Targeting, } GT(\sigma)} - \underbrace{L(\sigma)}_{\text{Direct cost}}, \tag{6}$$

where we now suppress the dependence of $\bar{\gamma}_{0/1}(.)$ on σ for conciseness. The first term in expression (6) is the ordeal's "gains from targeting," or $GT(\sigma)$. This captures how effectively the ordeal screens or "targets" enrollment to positive net-welfare individuals ($\gamma_i > 0$), relative to the alternatives of full exclusion and inclusion. We show below that $GT(\sigma)$ corresponds exactly to areas between (appropriately defined) marginal value and cost curves of an ordeal, allowing us to display these

[14] In the classic model, ordeals impose a "real" hassle cost on enrollee i of σh_i, which is identical to their impact on take-up behavior, but no costs on non-enrollees (who need not incur the hassle) or administrative costs for the government. Thus, in the classic setup, $L(\sigma) = D(\sigma) \cdot \sigma \bar{h}_1(\sigma)$. However, Ericson (2020) notes that policies like defaults may impact take-up through behavioral frictions like inattention that do not involve real welfare costs for (already-attentive) enrollees. Additionally, some barriers like stigma may impose psychological costs even on non-enrollees. The general $L(\sigma)$ allows our model to capture any of these cases.

[15] To derive this, we use the fact that $\bar{\gamma}$ is the welfare of the average enrollee in the full population, so for any σ, $\bar{\gamma} = D(\sigma) \cdot \bar{\gamma}_1(\sigma) + \left[1 - D(\sigma)\right] \cdot \bar{\gamma}_0(\sigma)$. Note that our analysis implicitly normalizes the size of the full population (enrollees plus non-enrollees) to be 1.0.

gains graphically. The second term, $L(\sigma)$, is the ordeal's total direct costs, which need not be incurred if the government simply excludes or includes everyone.

The key takeaway of this expression is that an ordeal is desirable only if it achieves positive gains from targeting large enough to exceed the ordeal's direct costs. Positive gains from targeting, in turn, requires that included groups be favorable (positive net welfare) and excluded groups be unfavorable (negative net welfare):

$$\text{(Positive Gains from Targeting)} \quad \bar{\gamma}_1(\sigma) > 0 > \bar{\gamma}_0(\sigma). \tag{7}$$

A necessary condition for (7) is that the ordeal induces "effective targeting" between included and excluded groups, or $\Delta\gamma \equiv \bar{\gamma}_1 - \bar{\gamma}_0 > 0$. We call the term $\Delta\gamma$ the "*targeting efficacy*." It is straightforward to show that $GT(\sigma) > 0$ only if $\Delta\gamma > 0$ and that $GT(\sigma)$ is an increasing function $\Delta\gamma$.[16]

There are two reasons the gains from targeting condition in (7) may fail, both of which, we will argue, become more likely with adverse selection. The two reasons are

- **Backward Sorting:** $\bar{\gamma}_1(\sigma) < 0 < \bar{\gamma}_0(\sigma)$. The ordeal sorts "backward" by including inefficient and excluding efficient enrollees. Note that this implies ineffective targeting, or $\Delta\gamma < 0$.
- **Optimal Universality:** Either $\bar{\gamma}_1, \bar{\gamma}_0 > 0$ or $\bar{\gamma}_1, \bar{\gamma}_0 < 0$. It is better to simply include or enroll everyone, rather than screening with the ordeal. Note that this may be true even if targeting is "effective" ($\Delta\gamma > 0$).

In our empirical work, we analyze these conditions for a *particular* ordeal (at a given intensity σ) since this is what we observe. Conceptually, with more variation, these conditions could be assessed *globally* across all $\sigma > 0$ for a given ordeal, which is what we depict in our graphs below.

The Classic Ordeals Debate.—How do these conditions for ordeal desirability relate to the classic ordeals debate? The classic rationale for ordeals going back to Nichols and Zeckhauser (1982) is that they result in "self-screening" or "self-targeting," in which people who highly value the program enroll, while low-value types drop out. Intuitively, hassle costs screen consumers just like prices in standard markets, with high-value consumers willing and low-value consumers unwilling to buy a good. In its classic formulation, self-screening is a statement about screening on private welfare, W_i. Under self-screening,

$$\text{(Self-screening)} \quad \Delta W \equiv \bar{W}_1 - \bar{W}_0 > 0. \tag{8}$$

[16] The gains from targeting from (6) yields

$$GT(\sigma) = D(\sigma)\left[1 - D(\sigma)\right] \cdot \Delta\gamma - K(\bar{\gamma}),$$

where $K(\bar{\gamma}) \equiv \max\left\{\left[1 - D(\sigma)\right] \cdot \bar{\gamma}, -D(\sigma) \cdot \bar{\gamma}\right\} \geq 0$ is a (nonnegative) correction that captures the fact that targeting is less desirable when a program's overall average welfare ($\bar{\gamma}$) is either very positive or very negative. Because the second term subtracts a nonnegative value, $GT(\sigma) > 0$ only if $\Delta\gamma > 0$.

In a model without behavioral biases ($\varepsilon_i = 0$) and homogeneous hassle costs ($h_i = \bar{h}\ \forall i$), self-screening must hold as a consequence of rational choice. The classic critiques of self-screening, therefore, focus on ways that biases or hassles may be larger for high-value types—in our notation, $\mathrm{cov}[W_i, \varepsilon_i] > 0$ and/or $\mathrm{cov}[W_i, h_i] > 0$. For instance, work on the "psychology of scarcity" argues that the poor, for whom social programs are especially valuable, may also experience the largest biases and hassle costs of overcoming ordeals (Bertrand, Mullainathan, and Shafir 2004; Mullainathan and Shafir 2013).[17]

Notice, however, that self-screening on *private* welfare (W_i) is not equivalent to favorable screening on *social value*, $V_i^{Soc} = \mu_i W_i + E_i$. This distinction is often missed in ordeals analyses that do not clearly delineate private versus social value. We say that an ordeal achieves favorable *social value sorting* if

$$\text{(9)} \qquad \text{(Social value sorting)} \qquad \Delta V^{Soc} \equiv \bar{V}_1^{Soc} - \bar{V}_0^{Soc} > 0.$$

In addition to the ways self-screening can fail, social value sorting can fail if ordeals differentially exclude people with high-welfare weights (μ_i) or with large positive externalities (E_i). This is likewise consistent with the "psychology of scarcity" ideas if ordeals differentially screen out poorer individuals (for whom μ_i is larger in standard welfare functions).

However, we emphasize that the right metric of targeting is not private welfare or even social value but net social welfare, $\gamma_i = V_i^{Soc} - C_i^{Net}$, or what we have called favorable *targeting efficacy*:

$$\text{(10)} \quad \text{(Targeting efficacy)} \quad \Delta\gamma \equiv \bar{\gamma}_1 - \bar{\gamma}_0 = \underbrace{\left(\bar{V}_1^{Soc} - \bar{V}_0^{Soc}\right)}_{\text{Social Value sorting}} - \underbrace{\left(\bar{C}_1^{Net} - \bar{C}_0^{Net}\right)}_{\text{Cost sorting}} > 0.$$

It is straightforward to see that targeting efficacy and value sorting coincide only in the special case where there is no offsetting sorting on costs. This is reasonable for programs with *constant costs* or more generally where costs are *uncorrelated* with value. For example, this might be reasonable for slots in a public childcare program or for a welfare program that gives everyone the same benefit amount. But it is unlikely to apply to insurance programs and other settings characterized by cost heterogeneity and adverse selection, which we turn to next.

C. *Ordeals Targeting and Adverse Selection*

How do the conditions for ordeals being optimal relate to adverse selection? In this subsection, we use our model to analyze the social welfare impact of ordeals. We show that the targeting impacts of ordeals can be visualized in a simple graphical framework, following the approach of Einav and Finkelstein (2011) for selection markets. This lets us visualize the role of adverse selection for the gains from targeting and therefore the desirability of ordeals.

[17] In a related vein, Spinnewijn (2015, 2017) argue that behavioral biases tend to reduce the slope of the social value curve relative to demand, making revealed preference sorting less efficient.

While the classic ordeals debate has tended to focus on the wedge between individual choice and enrollees' true private welfare (W_i) or true social value (V_i^{Soc}), we use our framework to illustrate how the economics of adverse selection can create an analogous wedge between V_i^{Soc} and *net* social welfare, $\gamma_i = V_i^{Soc} - C_i^{Net}$. Thus, even when ordeals successfully induce self-screening and favorable value sorting, adverse selection can erode or even reverse the gains from targeting.

Adverse Selection and Targeting.—Adverse selection is a feature typically associated with insurance and other "selection markets," where it is known to unravel trade and distort market outcomes. However, the underlying features driving adverse selection may also be relevant for thinking about targeting in social programs. These two key features are

1. **Cost Heterogeneity:** C_i^{Net} varies across enrollees (with variance $\sigma_C^2 > 0$).
2. **Value-Cost Correlation:** C_i^{Net} correlates positively with V_i^{Soc}, or $\rho = \operatorname{corr}\left[V_i^{Soc}, C_i^{Net}\right] > 0$.[18]

These two features characterize many insurance programs where an individual's value (demand) and cost are both heavily driven by their risk. For instance, in health insurance, sicker individuals tend to have both higher value for insurance and higher expected costs. Adverse selection tends to result in $\bar{C}_1^{Net} - \bar{C}_0^{Net}$ having the same sign as $\bar{V}_1^{Soc} - \bar{V}_0^{Soc}$. Under adverse selection, positive value sorting $(\bar{V}_1^{Soc} - \bar{V}_0^{Soc} > 0)$ is not enough for an ordeal to be desirable; it is possible to have small or even negative targeting efficacy ($\Delta\gamma \approx 0$ or $\Delta\gamma < 0$) if sorting on costs is sufficiently large.

While we focus on adverse selection, *advantageous* selection may be relevant in some settings, like long-term care insurance. Under advantageous selection, costs vary $(\sigma_C^2 > 0)$, but the value-cost correlation is negative $(\rho < 0)$. As a result, ordeals will generally target more effectively than without selection since low-value types (who self-screen out) will also have high costs.

Graphical Analysis.—We show that the gains from targeting under adverse selection can be illustrated using the familiar graphical framework of Einav, Finkelstein, and Cullen (2010) for welfare in selection markets. The intuition is that different levels of the intensity of an ordeal, given by σ in our framework, trace out marginal value and marginal cost curves in much the same way as different prices generate demand and marginal cost curves in the original Einav, Finkelstein, and Cullen (2010) analysis. For a given ordeal of strength σ, we define the marginal social value curve $MV(\sigma) = E\left[V_i^{Soc} | W_i - \varepsilon_i = \sigma h_i\right]$ as the expected social value of those for whom a marginally stronger ordeal would cause not to enroll. Likewise, we define the marginal cost curve as $MC(\sigma) = E\left[C_i^{Net} | W_i - \varepsilon_i = \sigma h_i\right]$. It is straightforward to see that the conditional means in equation (10) ($\bar{V}_1^{Soc}$, $\bar{V}_0^{Soc}$, $\bar{C}_1^{Net}$ and $\bar{C}_0^{Net}$) are the average values of $MV(\sigma)$ and $MC(\sigma)$ to the left and right of $D(\sigma)$.

[18] In many settings, this condition is presented as a positive correlation between direct costs C_i and private welfare W_i. For the purpose of this discussion, we assume that W_i and V_i^{Soc} are highly correlated, as are C_i and C_i^{Net}, so these conditions are aligned.

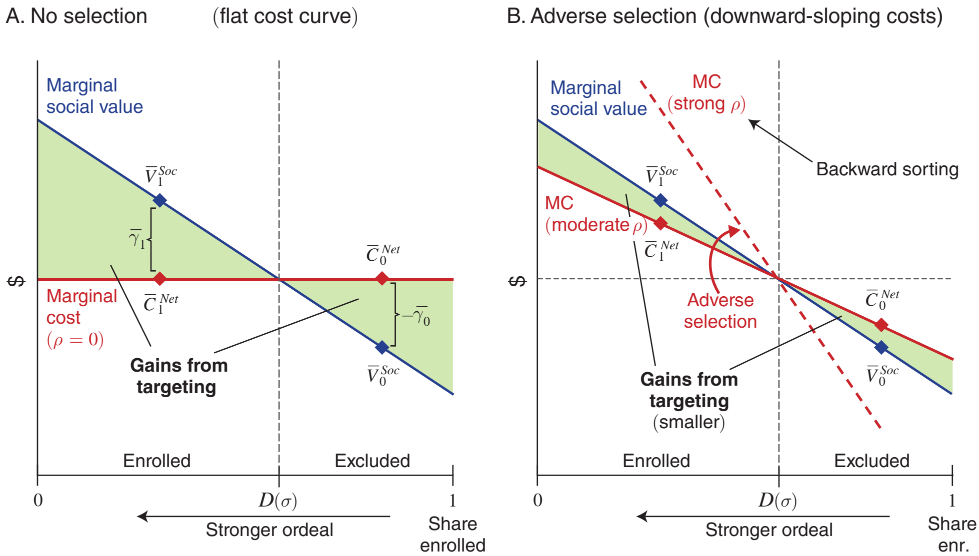

Figure 1. Gains from Ordeals Targeting with No Selection versus Adverse Selection

Notes: The figure shows the gains from targeting from ordeals in two cases: (i) the "standard" ordeals case without selection (a flat marginal cost curve, panel A) and (ii) with adverse selection (downward-sloping cost curve, panel B). Both panels depict enrollee value and cost curves for marginal enrollees as the ordeal strengthens and enrollment drops (moving right to left), using a setup similar to Einav, Finkelstein, and Cullen (2010). The green shaded areas are the "gains from targeting," which shrink or become negative under adverse selection.

The key impact of adverse selection in this framework is to make the marginal cost curve *downward sloping* since low-value types also have low costs. This, we argue, reduces or reverses an ordeal's gains from targeting, potentially leading to backward sorting. Further, it makes it more likely that $MV(\sigma)$ lies entirely above or below $MC(\sigma)$, the condition for optimal universality.

Figure 1 illustrates this adverse selection logic graphically, showing how adverse selection reduces or reverses the gains from targeting. The curves in each panel depict the marginal social value (blue) and cost (red) curves as the ordeal gets stronger (moving right to left), an ordeals version of standard demand and marginal cost curves from Einav, Finkelstein, and Cullen (2010). The diamonds are average value and cost for included/excluded enrollees under an ordeal, optimally set to maximize targeting gains. Both panels show the same downward-sloping marginal value curve, reflecting the case in which the ordeal favorably sorts on social value, $\bar{V}_1^{Soc} - \bar{V}_0^{Soc} > 0$. The areas between the value and cost curves, shaded in green, correspond to the gains from targeting, $GT(\sigma)$,[19] and are increasing in $\Delta\gamma = \bar{\gamma}_1 - \bar{\gamma}_0$, as shown in the graph.

Panel A illustrates the classic ordeals case with *no selection* (i.e., where costs are constant or uncorrelated with value), represented by a flat marginal cost curve that intersects marginal value at an interior point. As a result, targeting efficacy $(\bar{\gamma}_1 - \bar{\gamma}_0)$

[19] Technically, gains from targeting equals the smaller of the two shaded triangles.

is equivalent to social value sorting $(\bar{V}_1 - \bar{V}_0)$ because there is zero sorting on cost. An ordeal, therefore, achieves positive gains from targeting as long as the value curve is downward sloping, that is, $\Delta V^{Soc} > 0$. This is the key idea underlying the classic "self-screening" and "social value sorting" rationales for ordeals described above.

Panel B shows how this changes with *adverse selection*. The marginal value curve remains downward sloping, but now the marginal cost curve is also downward sloping, capturing the positive value-cost correlation. We show a case where the $MC(\sigma)$ curve rotates around its intersection point with $MV(\sigma)$, so the two curves continue to intersect. Because of this rotation, the gains from targeting (as shown in the green shaded area) are substantially reduced (when ρ is modest) and may be negative (when ρ is large). The key question for targeting efficacy is no longer whether the marginal value curve is downward sloping but whether it is *steeper* than marginal costs. In the case illustrated by the dashed red curve—where $MC(\sigma)$ is steeper than $MV(\sigma)$—the ordeal leads to "*backward sorting*." In this case, the ordeal targets inversely from what is desirable: those who are enrolled have negative surplus, while those who are excluded have positive surplus. This type of backward sorting is closely related to the idea that price-based sorting may also be inefficient in insurance markets (Marone and Sabety 2022).[20]

Figure 2 shows a second way adverse selection may undermine the optimality of ordeals: by leading to "optimal universality." We show both the no-selection and "modest" adverse selection $MC(\sigma)$ curves from the prior figure but now consider what happens if the $MV(\sigma)$ is higher, e.g., because society places a higher welfare weight (μ) on program enrollees. With no selection, a more modest but still positive ordeal is optimal because the marginal value and cost curves continue to intersect. But with adverse selection, the *MV* curve lies *entirely above MC*, implying that full enrollment (zero ordeal) is optimal. The same idea applies in reverse if the marginal value curve is lower (via a lower μ), with adverse selection making it more likely that no enrollment is optimal (see Supplemental Appendix Figure A.1). Intuitively, adverse selection makes these "universal" optima more likely because the similar downward slope of *MV* and *MC* makes them less likely to intersect within a given range.

Mathematical Analysis.—We now formalize these arguments. We start with the claim that adverse selection reduces or reverses the gains from targeting—the sorting argument shown in Figure 1, panel B. Note that given estimates of V_i^{Soc} and C_i^{Net}, we can quantify the value-cost relationship by considering the linear projection of enrollee costs onto value: $C_i^{Net} = \bar{C} + \hat{\beta} \times V_i^{Soc} + \omega_i$, where $\bar{C}$ is the mean of net costs and ω_i is a residual capturing cost heterogeneity orthogonal to value. This projection can always be performed and results in the standard regression coefficient $\hat{\beta} = \rho \cdot \sigma_C / \sigma_V$, where σ_C and σ_V are the standard deviations of cost and value, and

[20] Sorting may be improved if ordeals (or prices) can be targeted only at high-cost enrollees (Bundorf, Levin, and Mahoney 2012), but this is typically not done because it would be inequitable to the sick. In a different context, the fact that "prior authorization" hassles are targeted at high-cost prescription drugs may explain why these yield savings in excess of their costs (Brot-Goldberg, Burn et al. 2023).

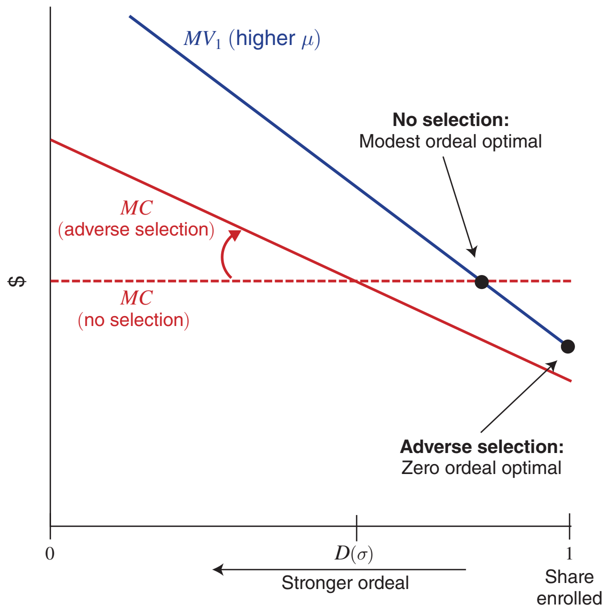

FIGURE 2. OPTIMAL UNIVERSALITY WITH ADVERSE SELECTION

Notes: The figure shows how adverse selection increases the likelihood of "optimal universality" when the social marginal value (MV) curve is shifted upward (relative to Figure 1) due to a higher social welfare weight, μ. With no selection, the new marginal value curve (MV_1) still intersects marginal cost (MC), implying that a (more modest) ordeal is still optimal. With adverse selection, MV_1 lies entirely above MC, implying full enrollment (zero ordeal) is now optimal.

$\rho \in [-1, 1]$ is the value-cost correlation. Applying this projection to the terms for targeting efficacy in (10) yields[21]

$$(11)\qquad \underbrace{\bar{\gamma}_1 - \bar{\gamma}_0}_{\text{Targeting Efficacy}} = \underbrace{\left(\bar{V}_1^{Soc} - \bar{V}_0^{Soc}\right)}_{\text{Social Value sorting}} \times \underbrace{\left[1 - \overbrace{\left(\rho \cdot \frac{\sigma_C}{\sigma_V}\right)}^{\text{Adverse Selection Tax }(\hat{\beta})} - \widetilde{\Delta\omega}\right]}_{\text{Correction for value-cost correlation}},$$

where $\widetilde{\Delta\omega} \equiv (\bar{\omega}_1 - \bar{\omega}_0)/\left(\bar{V}_1^{Soc} - \bar{V}_0^{Soc}\right)$ captures the ordeal's sorting on idiosyncratic costs. We call $\hat{\beta}$ the "adverse selection tax" since it captures the degree to which adverse selection (a large covariance between value and costs) "taxes away" the welfare gains from favorable sorting on value.

Equation (11) formalizes the relationship between social value sorting $\left(\bar{V}_1^{Soc} - \bar{V}_0^{Soc}\right)$ and the true targeting efficacy, $\bar{\gamma}_1 - \bar{\gamma}_0$. If program costs are either constant across enrollees ($\sigma_C = 0$) or uncorrelated with enrollee value ($\rho = 0$), social welfare gains are approximately equal to value sorting. However, as cost heterogeneity (σ_C) and the value-cost correlation (ρ) grow more positive—precisely the two key features of adverse selection laid out above—the adverse selection tax

[21] We get this from applying the projection to get $\bar{C}_1^{Net} - \bar{C}_0^{Net} = \hat{\beta} \times \left(\bar{V}_1^{Soc} - \bar{V}_0^{Soc}\right) + (\bar{\omega}_1 - \bar{\omega}_0)$, which can be rearranged to yield the expression in (11).

grows, and gains from targeting are diminished. Further, if $\hat{\beta}$ grows large enough that

$$\hat{\beta} = \rho \cdot \frac{\sigma_C}{\sigma_V} > 1 - \widetilde{\Delta\omega}, \tag{12}$$

the correction term becomes negative, and the ordeal leads to backward sorting (on social welfare) despite favorable sorting on value. This corresponds to a "steeper" marginal cost than marginal value curve in Figure 1, panel B. If $\widetilde{\Delta\omega} \geq 0$—which occurs if an ordeal does not screen, or screens unfavorably, on idiosyncratic costs (the case we usually find in our empirical work)—a sufficient condition for backward sorting is $\hat{\beta} > 1$, or $\rho > \sigma_V/\sigma_C$.

This analysis provides insight into why ordeals will generally work poorly in settings with strong adverse selection, where $\hat{\beta} > 1$. In these settings, *any* ordeal that sorts favorably on value will sort *backward* on efficiency, unless it happens to screen in people with low *idiosyncratic* costs ($\widetilde{\Delta\omega} < 0$), something that while possible, is not implied by economic theory. More generally, even modest adverse selection ($\hat{\beta} \in (0,1]$, or $\rho \in (0, \sigma_V/\sigma_C]$) "taxes" away the gains from value sorting in proportion to $\hat{\beta}$, making the real welfare gains much smaller.[22]

We now formalize the claim that adverse selection makes optimal universality more likely, as depicted in Figure 2. As in the figure, we consider how shifts in marginal social value driven by a higher/lower social welfare weight (μ) affect the optimality of a given ordeal with strength σ.[23] For the ordeal to yield targeting gains per condition (7), it must be the case that $\bar{\gamma}_1(\sigma) > 0 > \bar{\gamma}_0(\sigma)$, or $\bar{V}_1^{Soc}(\sigma;\mu) - \bar{C}_1^{Net}(\sigma) > 0 > \bar{V}_0^{Soc}(\sigma;\mu) - \bar{C}_0^{Net}(\sigma)$, where we highlight that $\bar{V}_1^{Soc}$ and $\bar{V}_0^{Soc}$ are both (increasing) functions of μ. These inequalities, therefore, implicitly define a range of μ over which the ordeal is desirable: $\mu \in [\mu^*_{min}, \mu^*_{max}] \equiv [(\bar{C}_1^{Net} - \bar{E}_1)/\bar{W}_1, (\bar{C}_0^{Net} - \bar{E}_0)/\bar{W}_0]$ as long as $\mu^*_{min} \leq \mu^*_{max}$. Relative to no selection ($\bar{C}_1^{Net} = \bar{C}_0^{Net}$), adverse selection rotates the cost curve, making $\bar{C}_1^{Net} > \bar{C}_0^{Net}$, which pushes upward μ^*_{min} and downward μ^*_{max}. Thus, adverse selection *narrows the range* of social preferences $[\mu^*_{min}, \mu^*_{max}]$ over which ordeals are preferred to universal policies. (See Supplemental Appendix Figure A.1 for a visualization of this argument.) Further, for sufficiently strong adverse selection, this range becomes null, implying that there is no μ at which the ordeal is optimal.

Broader Implications for Transfer Programs.—While our emphasis has been on insurance programs, our framework also sheds light on many *transfer* programs where recipient value and public costs are naturally correlated via the (varying) *benefit amounts*, which are both a benefit to enrollees and a cost to the government. For instance, in many means-tested programs, benefit amounts vary with enrollee income or family status. This suggests that the logic of correlated value and costs may apply,

[22] One reason $\hat{\beta}$ is likely to be large in low-income populations is that σ_V (at least for private WTP) tends to be small because marginal utility of consumption is high, while σ_C is much larger, reflecting variation in health needs.

[23] We make this argument for a particular σ, but an analogous argument applies across a *full range* of values of σ to show that adverse selection makes it more likely that the $MV(\sigma)$ and $MC(\sigma)$ curves do not intersect over this range.

and self-targeting may not translate into significant welfare gains. Instead, the desirability of ordeals may depend on whether low-benefit-amount enrollees also tend to be those the government wishes to screen out for other reasons (e.g., because they are less poor, so have a lower social welfare weight).

Our analysis can help interpret the findings in past work. For instance, both Finkelstein and Notowidigdo (2019) (studying SNAP) and Bhargava and Manoli (2015) (studying the EITC) find that hassles on average screen out people who receive smaller benefit amounts from these programs. But the normative implications are different. In SNAP, low-benefit types are generally *higher-income* individuals, for whom economic need is less. But in the EITC, low-benefit types were generally *lower-income* individuals *without kids*, for whom need may be high. By contrast, ordeals screening works well in programs that distribute supplies with *equal costs* for all participants, as in free chlorine solution for water treatment (Dupas et al. 2016).

Connection to Economics of Nudges.—Our analysis of ordeals relates to the broader economics of "nudges" (Thaler and Sunstein 2008) and similar nonprice interventions. Although the vast majority of this literature focuses on empirical impacts and positive economics, recent work by Allcott et al. (2022) unpacks the welfare implications of nudges. Their work emphasizes that simple *average treatment effects* on demand or adoption of ostensibly beneficial goods or behaviors may be a misleading guide to welfare. Instead, the key welfare question is whether a nudge reduces *choice distortions*, by inducing people to consume or behave more in line with what is socially optimal.[24] A nudge improves social welfare only if it reduces (more than it exacerbates) baseline under- and overconsumption of a good relative to the social optimum.

This aligns closely with our analysis of take-up and targeting with ordeals for social programs. An ordeal improves welfare only if it corrects (more than it exacerbates) errors of overenrollment (enrolling $\gamma_i < 0$ types) and of underenrollment (excluding $\gamma_i > 0$ types) that occur with alternate policies like full inclusion and exclusion. This is exactly what is captured by our targeting efficacy statistic, $\Delta\gamma = \bar{\gamma}_1 - \bar{\gamma}_0$, and by our expression for "gains from targeting" in (6). Indeed, there is a close parallel between our model and the setup of Allcott et al. (2022),[25] suggesting a deep connection between the welfare economics of nudges and ordeals. This also suggests that thinking about nudges through the lens of *optimal targeting* may be a fruitful way to understand their welfare impacts.

[24] Allcott et al. (2022) show that this occurs when a nudge reduces the *variance* of "net distortions," or the (individual-specific) wedge between choice utility and social welfare arising from behavioral biases, externalities, and other factors like markups and taxes. These wedges may be either positive or negative, so a smaller variance implies behavior more in line with social welfare.

[25] Importantly, we allow C_i^{Net} to vary (whereas marginal cost is fixed in their model) because we are studying a selection market. Finally, their model is more complex because it allows prices to endogenously adjust to nudges (via their impact on supply/demand), which necessitates an analysis of price pass-through impacts that we can ignore.

II. Setting, Auto-Enrollment Policy, and Data

A. *Massachusetts Exchange Setting*

CommCare Exchange.—We study Commonwealth Care ("CommCare"), a subsidized insurance exchange in Massachusetts that operated from 2006 to 2013 before shifting form in 2014 at the ACA's implementation. CommCare covered low-income adults with family income below 300 percent of the federal poverty level (FPL, or "poverty") and without access to insurance from another source, including an employer or public program (i.e., Medicare or Medicaid). We focus on the population with income below 100 percent of FPL for whom the auto-enrollment policy applied. Given eligibility rules for other programs, this group is almost entirely childless adults age 19–64.[26]

CommCare offered generous insurance at heavily subsidized premiums. The program specified a detailed benefit structure (i.e., cost sharing rules and covered medical services) that private insurers were required to follows. Each insurer offered a single plan with the standardized benefits but could differ in its network of hospitals and doctors. For the below-poverty group we focus on, benefits were equivalent to Medicaid—that is, broad covered services with essentially no patient cost sharing (the actuarial value is 99.5 percent)—and all plans were fully subsidized ($0 premium). This setup is similar to Medicaid managed care programs. As in Medicaid, there is no financial cost to insurance, and the only barriers are enrollment hassles. An important difference from Medicaid, however, is that CommCare does *not* have retroactive coverage; coverage starts the first day of the month *after* completing enrollment.[27] Therefore, enrollment delays have a meaningful impact, including the risk of getting acutely ill and incurring medical debts before enrollment takes effect.

Application and Enrollment Process.—It is well-known that there is substantial "churn" into and out of eligibility for different forms of health insurance, e.g., due to job changes, income fluctuation, or family status changes. Therefore, many people newly need health insurance and apply for public coverage. For CommCare, the enrollment process involves two steps, as shown in Figure 3. Step 1 is to apply for eligibility. This requires completing a six-page application that asks about income, demographics, family status, and access to other health insurance (see Supplemental Appendix H for snapshots of the form). The state used this information to determine eligibility for Medicaid or CommCare (dual eligibility should not occur) and to sort people into income-based subsidy groups in CommCare. Although the application form is a meaningful hassle, many individuals get help from a social worker or medical staffer in completing it, often just after having visited a medical provider while uninsured.

[26] Medicare covers seniors age 65+, and Massachusetts Medicaid covers children up to 300 percent of FPL, parents with dependent children up to 133 percent of FPL, and pregnant women up to 200 percent of FPL. In addition to the nonelderly, CommCare covered a small number of immigrants age 65+ not eligible for Medicare. As we discuss below, we drop immigrant enrollees from our sample.

[27] By contrast, Medicaid covers medical bills incurred prior to enrollment, typically with a 90-day retroactive period. As a result, Medicaid eligibles have a form of "conditional coverage" that is not available from CommCare.

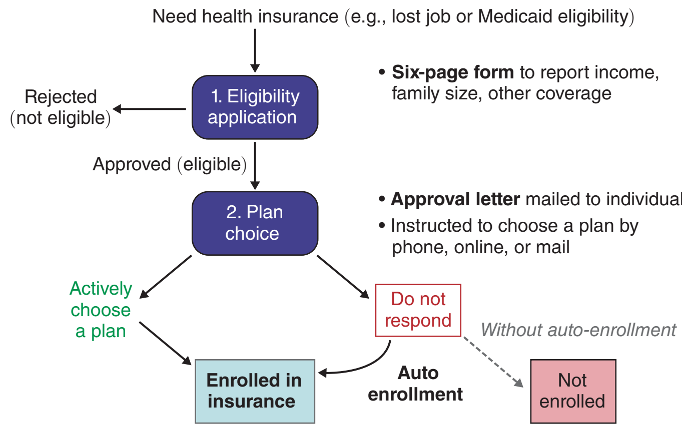

Figure 3. Enrollment Process and Auto-Enrollment Policy

Notes: The figure diagrams the enrollment process for the Massachusetts health insurance exchange we study (CommCare). Prospective enrollees who need health insurance must follow a two-step process. First, they apply for eligibility, completing a six-page form with information on income, family status, and other coverage. Second, if approved, they are mailed an approval letter and asked to choose a (free) health plan by phone, online, or mail. The auto-enrollment policy applies to approved individuals who do not respond to this approval letter within 14 days ("passive" individuals). With auto-enrollment (the policy from 2007 to 2009), they are auto-enrolled into a state-selected plan; without auto-enrollment (post-2010 policy), they are not enrolled unless and until they actively respond.

The second enrollment step is to choose a plan. After determining eligibility, the state notified an individual (by mail and/or email) and provided information on available plans and associated premiums. Supplemental Appendix H shows this two-page approval letter. To complete enrollment, individuals were asked to choose a plan by calling, going online, or circling a plan choice and returning it by mail. Relative to the initial application, this step was quite simple. However, without auto-enrollment, individuals still had to take action to enroll. Moreover, the action needed to be taken *independently* in response to the approval letter, which could be lost, misunderstood, or forgotten.

B. *Auto-Enrollment Policy and Timeline*

Auto-Enrollment Policy.—CommCare's auto-enrollment policy set the default outcome for people determined eligible (step 1 of the process) but who did not respond when asked to choose a plan (step 2; see Figure 3). The policy applied only to below-poverty enrollees, for whom all plans were free.[28] This allowed regulators to borrow a policy widely used in Medicaid managed care that "auto-assigns" passive new enrollees into a state-selected plan. Aggregate statistics suggest that

[28] Auto-enrollment was generally not used for above-poverty enrollees because premiums varied across plans and were typically nonzero, raising concerns about auto-enrolling people into plans that generated a financial debt for them. There were two limited exceptions of auto-enrollment for 100–150 percent of poverty enrollees, both of which are excluded from our main sample (see discussion below): (i) for reenrollees prior to 2010 who reenrolled with a gap of less than 12 months and (ii) for new enrollees during the single month of December 2007 (fiscal year 2008m6).

auto-assignment in Medicaid is very common: the median state auto-assigns 45 percent of new enrollees (Smith et al. 2015). However, we are not aware of any *causal* evidence on this policy's impact on take-up, likely because of a lack of variation in its use.

Auto-enrollment applied when individuals entered the market, but with different rules for two groups: (i) "new enrollees" joining for the first time and (ii) "reenrollees" joining after a gap in coverage. We focus our main analysis on new enrollees. New individuals were mailed a coverage approval letter and given 14 days to actively choose a plan before being auto-enrolled if they failed to respond. This lets us observe mode of enrollment (active versus passive) directly in our administrative data.[29]

There was one notable exception to the process for new enrollees near CommCare's inception in 2007 when the state "auto-converted" a large population from its pre-RomneyCare uncompensated care pool (UCP). These individuals did not complete a new eligibility application but were determined eligible based on information from their original UCP application, often completed months beforehand. Consistent with the long lag, many of these UCP individuals failed to respond and were auto-enrolled, creating a large spike in auto-enrollment in early 2007. Because of these distinct circumstances, we focus our main analysis on the "steady-state" auto-enrollment period (fiscal years 2008–2009), with the initial period (2007) analyzed for comparison and robustness.[30]

Policy Timeline.—We examine auto-enrollment policy changes during fiscal year (FY) 2010 (which ran from July 2009 to June 2010). Facing a Great Recession–related budget shortfall, CommCare needed to cut spending. The program had raised enrollee premiums and copays the prior year, and it was eager to avoid doing so again. Suspending auto-enrollment provided an alternative to reduce enrollment and therefore subsidy spending. The exchange did so as of the start of FY 2010, with (because of a lagged impact) a final group of passive enrollees joining in 2010m1 (July 2009). These cuts proved quite effective, and CommCare unexpectedly came in under budget during 2010. As a result, the program temporarily reinstated auto-enrollment in the final three months of FY 2010. After this, facing continued budget pressures, it was permanently canceled in 2011.

These changes give us variation to estimate the causal impact of auto-enrollment. To be valid, it is important that there not be other concurrent shocks or policy changes that affect enrollment around the same time. Based on background research and discussions with the exchange administrator, this appears to be true, with one exception: an eligibility cut for noncitizen enrollees in 2010m4 (October 2009), two months after the auto-enrollment suspension. To avoid biasing our results, we

[29] By contrast, most reenrollees were *immediately* auto-enrolled in their former plan (without a 14-day window to actively choose), and auto-reenrollment was also used for some above-poverty enrollees (our control group). For these reasons, we exclude reenrollees from our main sample, reporting effects on them in robustness analysis (see Supplemental Appendix B.2).

[30] Supplemental Appendix C.5 compares our main targeting analysis for the 2008–2009 sample (see Section IVA) to the results for 2007. Interestingly, while auto-enrollment is much more common in early 2007, we find very similar targeting (active versus passive enrollee characteristics) in both periods.

exclude noncitizen enrollees from our sample in all periods.[31] Aside from this, other enrollment-relevant policies did not change.[32] Nonetheless, to address any unobserved demand shocks, we also use a control group of higher-income enrollees not subject to auto-enrollment.

Other Policy Details.—Although our analysis focuses on enrollment impacts, other policy details are of interest, including rules for plan auto-assignment. The plan assignment rule had two parts. Passive enrollees with prior enrollment with an insurer in the past 12 months (either in CommCare or Medicaid) were auto-assigned to that insurer. Other new enrollees were randomly assigned to plans, with probability shares following a schedule giving more weight to plans with lower (state-paid) premiums. After enrollment, all new/reenrollees (both active and passive) could freely switch plans within 60 days of starting coverage. In practice, the vast majority (96 percent of passive and 98 percent of active enrollees) stick with their initial plan, consistent with other work finding that default health plan assignment is very sticky (Brot-Goldberg, Layton et al. 2023).

These policies raise two interesting issues that we have not explored in this paper. First, random assignment could allow for inferring causal plan effects, as in recent work on Medicaid (Geruso, Layton, and Wallace 2020). In practice, we find evidence of slight demographic imbalance across plans, suggesting the presence of hard-to-observe exceptions to random assignment. We therefore have not pursued this topic further. Second, giving higher probability weights to lower-price insurers should affect competitive incentives. This topic is interesting but would require a different research design to study; we therefore leave it for future work.

C. *Data and Descriptive Statistics*

Exchange Admin Data and Sample Definition.—Our primary data come from deidentified CommCare administrative records for fiscal years 2007–2014, spanning November 2006 to December 2013 (Massachusetts Health Connector 2014). For all enrollees, we observe a panel of individual-level demographics and monthly plan enrollment, linked to insurance claims and risk scores. Observed demographics include age, gender, zip code of residence, and family income as a percentage of the poverty line. Insurance claims let us measure individuals' medical conditions and health care use and costs while enrolled. Importantly, the data include a flag for whether each new enrollee is auto-enrolled or actively chooses a plan. This lets us

[31] The eligibility change was for legal immigrant residents (typically green card holders) who had not yet cleared their "five-year bar" requirement to receive federal Medicaid matching funds—a group the state calls "aliens with special status" (AWSS). Starting in October 2009, the AWSS group was not eligible to newly enroll in CommCare, and existing AWSS enrollees were shifted into a parallel program. We observe a flag for AWSS status and enrollment in this parallel program, which lets us exclude these individuals from the sample in all periods.

[32] The start of 2010 did see the entry of a new insurer (CeltiCare). But for the below-poverty group, this expanded the choice set of available free plans, which should (if anything) increase enrollment, pushing in the opposite direction of our findings. In practice, CeltiCare had a narrow network and was not popular, with only 1.5 percent of below-poverty active choosers selecting it during 2010–2011. We therefore view the new availability of CeltiCare as having a negligible impact.

construct the key variables for our main analysis: monthly counts, characteristics, and outcomes for passive and active enrollees.[33]

We are interested in the policy's impact on enrollment totals and composition. For enrollment impacts, the main outcome of interest is counts of new enrollees joining CommCare per month (a flow measure). We use our panel data and a simple model to translate this into an effect on steady-state enrollment (a stock measure). For composition, we use variables on demographics, diagnoses, and medical spending during an individual's enrollment spell.

We make several limitations to our main CommCare analysis sample. First, we limit attention to new enrollees who (when they joined the market) were in one of two income groups: (i) the 0–100 percent of poverty "treatment" group and (ii) a 100–200 percent of poverty "control" group not subject to auto-enrollment. Second, we exclude from our sample noncitizen enrollees who (as described above) faced an eligibility cutback in October 2009, shortly after the auto-enrollment change (in August 2009). Finally, we limit our main sample period to FY 2008–2011 for analyses of the treatment group and to 2009–2011 for difference-in-differences (DD) regressions comparing treatment and control groups. We exclude 2007 because of the different nature of auto-enrollment during that year (see discussion above). For DD regressions, we further exclude 2008 because of other policy changes that affected the control group in mid-to-late 2008.[34] We end our analysis in 2011 because of a change in plan choice rules for the treatment group at the start of 2012 (see Shepard 2022).

Other Datasets.—We draw on two additional datasets for specific pieces of our analysis:

- **American Community Survey (ACS):** For context on uninsurance in Massachusetts, we use the ACS (Ruggles et al. 2015) to estimate the CommCare-eligible uninsured population by income group, following a method used by Finkelstein, Hendren, and Shepard 2019. Details are in Supplemental Appendix A.1.
- **Massachusetts All-Payer Claims Database (APCD):** We use the state's APCD (version 3.0, with data for 2009–2013) (Massachusetts CHIA 2014) to examine whether CommCare enrollees are enrolled in duplicate private insurance, as a possible reason for failing to actively enroll. The APCD is well suited for this purpose because it lets us observe a near-universe of Massachusetts health insurance plans and measure simultaneous coverage. Supplemental Appendix D describes the data construction method and shows that the APCD's enrollment counts for CommCare closely match our administrative data.

[33] We observe this flag for the FY 2007–2009 period when auto-enrollment is in effect, but due to a technical issue, it is missing during the policy's temporary reinstatement in April–June 2010. For this latter period, we report only aggregate data for all enrollees.

[34] Specifically, for individuals above 150 percent of poverty, the state's insurance mandate penalty took effect in December 2007 (FY 2008m6), leading to a spike in new enrollment. Also in December 2007, there was a large auto-enrollment for the 100–150 percent poverty group. For the whole 100–200 percent poverty control group, there was a change in plan premiums and subsidies at the start of FY 2009 (July 2008). Importantly, none of these changes applied to the treatment group, and policy for the control group was stable throughout the 2009–2011 period used in our DD analysis.

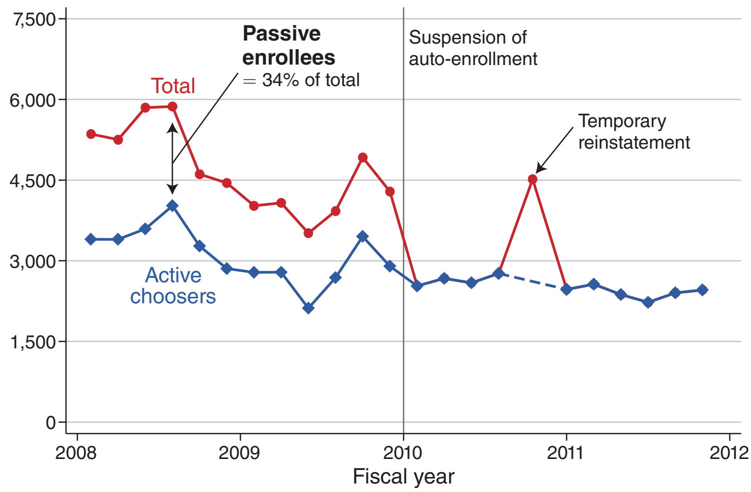

Figure 4. Active versus Passive New Enrollment into the Massachusetts Exchange

Notes: The graph shows counts of new enrollees per month for the below-poverty group subject to auto-enrollment. The red series is total new enrollment, the blue is active choosers, and the gap between these is passive auto-enrollment. The vertical line indicates the timing of auto-enrollment's suspension at the start of fiscal year 2010. After this, total enrollment equals active choosers, except for the period of auto-enrollment's temporary reinstatement (during which we lack the flag to separate active versus passive enrollment). Data are bimonthly averages to smooth over fluctuations.

Descriptive Statistics.—Figure 4 shows data on new enrollment per month in the treatment group (0–100 percent of poverty) over the main 2008–2011 period.[35] The figure plots both total new enrollment (in red) and the count of active choosers (in blue), with the gap between these being passive enrollees. Passive enrollees represent a sizable 34 percent share of new enrollment during 2008–2009, and new enrollment falls sharply when auto-enrollment was suspended at the start of 2010. The decline is almost identical to the number of passive enrollees during 2008–2009. Moreover, when the policy is briefly reinstated at the end of 2010, enrollment spikes up to a similar level as at the end of 2009. Together, these facts are consistent with auto-enrollment having a causal effect roughly equal to the full number of passive enrollees in the pre-period.

Supplemental Appendix Table A.1 further summarizes enrollment statistics, including enrollment counts for the 100–200 percent of poverty group and on total market enrollment and new versus reenrollment. Supplemental Appendix Table A.2 reports average consumer attributes; we defer a discussion of these to Section IV, where we compare active versus passive enrollees.

[35] The points are bimonthly averages to smooth over noise; see Supplemental Appendix Figure A.2 for the raw monthly data over the full 2007–2011 period. As that figure shows, auto-enrollment spiked during early 2007 because of the autoconversion of the state's uncompensated care pool.

III. Causal Impact of Auto-Enrollment Policy

This section presents our estimates of the impact on take-up of suspending auto-enrollment in 2010. After presenting results in Section IIIA, we provide context on the magnitude in Section IIIB.

A. *Impact on Health Insurance Enrollment*

We use the 2010 policy change to estimate the causal impact of auto-enrollment. To do so, we run difference-in-difference regressions on counts of monthly new enrollment, comparing the 0–100 percent of poverty "treatment" group (for whom auto-enrollment is in place through 2009 and suspended in 2010) to the 100–200 percent of poverty "control" group (for whom auto-enrollment was not in place throughout). The DD regression is

$$(13) \qquad NewEnr_{g,t} = \alpha_g + \beta_t + \gamma \cdot \mathbf{1}\{g = Treat,\ t \geq 2010\} + \varepsilon_{g,t},$$

where $NewEnr_{g,t}$ is (scaled) new enrollment for income group g (treatment or control) at time t, α_g is a group fixed effect (for the treatment and control groups), β_t is a time fixed effect, and $\varepsilon_{g,t}$ is an error. We run (13) on data from 2009 to 2011, excluding the period of temporary reinstatement of auto-enrollment at the end of 2010.[36] The dependent variable is "scaled" new enrollment, equal to a group's raw monthly counts divided by its average new enrollment in the pre-2010 period. This ensuring $NewEnr_{g,t}$ has a mean of 1.0 for each g in the pre-period and lets us interpret estimates as proportional effects. The coefficient of interest is γ, which is the DD estimate of the impact of turning off auto-enrollment (i.e., adding the active choice ordeal).

Figure 5 plots the data for the regression in (13) and reports the main DD estimate. Panel A shows results for *total* new enrollment (active plus passive). Trends for both groups are parallel in the pre-period, and treatment group enrollment drops sharply and persistently at the policy change. The DD estimate of $\gamma = -0.326$ implies that suspending auto-enrollment reduced new enrollment by 32.6 percent of the pre-period mean. In the reverse direction, new enrollment was 48 percent $(= 0.326/(1 - 0.326))$ higher when auto-enrollment was in place.

Figure 5, panel B shows the impact on the number of *actively choosing* new enrollees. In principle, auto-enrollment might induce some attentive individuals to be "purposely passive" because they know the stakes are low, e.g., if they view CommCare plans as roughly equivalent and are happy to let the regulator select for them.[37] If this were true, we would expect these purposely passive individuals to actively enroll when auto-enrollment stops in 2010, resulting in an uptick in *active*

[36] The time unit (t) is bimonthly periods, averaging over new enrollment in pairs of months, which smooths over a few single months when auto-enrollment appears not to have occurred followed by a surge in auto-enrollment the next month. We calculate standard errors using the normal linear model given the small samples sizes but verify that robust standard errors are essentially the same.

[37] Enrollees were informed about the auto-enrollment policy in the coverage approval letter, which stated, *"If you do not choose a health plan by [date], the Connector will choose one for you."* After early 2010, this language was removed, and enrollees were sent periodic reminder letters if they had qualified but not enrolled in coverage.

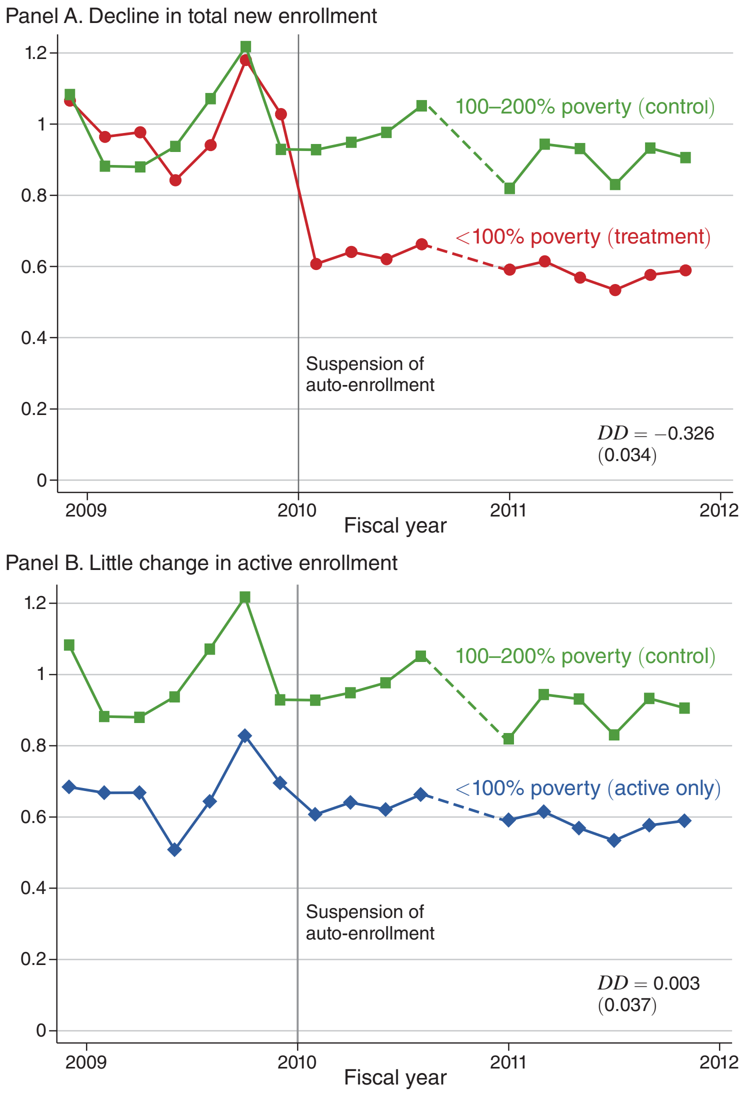

Figure 5. Enrollment Impact of Auto-Enrollment's Suspension

Note: The figure shows scaled new enrollment per month into CommCare and estimates of the DD specification (13) for estimating the causal effect of auto-enrollment's suspension. Each panel compares trends for below-poverty enrollees (the treatment group) versus 100–200 percent of poverty enrollees (the control group, not auto-enrolled). Each income group's series is rescaled by dividing by the group's pre-period mean new enrollment, which makes DD estimates interpretable as a proportional change. The temporary reinstatement period is excluded (as indicated with dashed lines). Panel A shows that total new enrollment falls sharply (by 32.6 percent) for the treatment group at the start of 2010, consistent with a causal effect of the policy. Panel B shows that the number of active new enrollees is flat through the policy change.

enrollment. Instead, Figure 5, panel B shows that there was no change in active new enrollment around the policy change, with a DD estimate of almost exactly zero ($\gamma = 0.003$) and no sign of an uptick in the two years following the policy change. As a further test, Supplemental Appendix Figure A.3 shows that we see no evidence of compositional changes in the characteristics of active enrollees, which we would expect if some people shifted to active choice.

This evidence suggests two facts about the ordeal of requiring active plan choice to get insurance. First, failure to actively enroll is unlikely to have been a strategic or purposeful decision; instead, passivity is more likely due to inattention or misunderstanding of enrollment rules. Second, active choice is unlikely to involve significant costs to inframarginal enrollees. If it did, we would expect some to substitute toward passivity when auto-enrollment is an option.

Effect on Steady-State Enrollment.—The results so far are on the *flow* of new enrollees, which falls immediately when auto-enrollment ends. The *stock* of total enrollment, however, changes more gradually, as existing enrollees exit, while fewer new enrollees enter each month. To estimate the impact on steady-state enrollment, Supplemental Appendix B.3 uses the data to calibrate a simple stock-flow model. We find that suspending auto-enrollment reduces steady-state enrollment by 24 percent; or in the reverse direction, enrollment is 32 percent higher with auto-enrollment in place. (This estimate is slightly smaller than the impact on new enrollment because passive enrollees have shorter durations.) The estimates from the stock-flow model are highly consistent with the raw data on the stock of below-poverty enrollment, which falls by 23 percent from late 2009 to the end of 2011 (Supplemental Appendix Figure A.7).

Robustness: Alternate Specifications and Effects on Reenrollment.—These estimates are quite robust to alternate specifications and control groups. Supplemental Appendix Table A.3 shows that the estimated 33 percent fall in new enrollment is little changed when we (i) use alternate income groups as controls (e.g., 100–150 percent FPL only, or 100–300 percent FPL), (ii) use no control group (a simple pre/post difference), and (iii) include the "temporary reinstatement" period in the regressions. Additionally, while the analysis so far has been limited to new enrollees, Supplemental Appendix B.2 shows that there are similar impacts on the number of reenrollees joining the exchange after a break in coverage. We find that reenrollment falls 35–39 percent at the start of 2010, very similar to the 32.6 percent fall for new enrollment. We therefore conclude that our main estimates on new enrollees are representative of the policy's overall impact.

B. *Magnitude: Comparison to Other Take-Up Policies*

How should we interpret the magnitude of the impact of auto-enrollment—a 48 percent increase in new enrollment and 32 percent increase in steady state? Several benchmarks provide context for this estimate. First, relative to other "nudge" interventions to increase health insurance take-up, these are very large impacts. Several recent randomized experiments have tested nudges like reminder mailings/phone calls, simplified plan information, and a simpler take-up process (Domurat, Menashe, and Yin 2021; Myerson et al. 2021; Ericson et al. 2023). These studies find take-up impacts of 1–4 percentage points among a similar passive population (people who have qualified for coverage but not chosen a plan).[38] Similarly, evidence from

[38] Goldin, Lurie, and McCubbin (2021) study a similar mail outreach intervention on uninsured individuals identified in tax filings. They likewise find a modest take-up impact of +1.1 percentage points, though even this small impact led to a meaningful decline in mortality among the marginally insured.

Aizawa and Kim (2020) suggests that a threefold increase in government advertising of ACA Marketplaces would increase market-level enrollment by 1.3 percentage points (or 7.6 percent). By contrast, our auto-enrollment policy leads to an *order of magnitude larger* impact: nearly complete take-up among the passive group and a 30–50 percent increase in the total enrolled population. These results suggest that while information and simplification matter, *making enrollment the default* may be critical to substantially boost take-up.

A second benchmark is the impact of financial incentives. Our estimated steady-state impact of auto-enrollment is nearly identical to the 33 percent effect of subsidies that reduce enrollees' premiums by \$39–\$40 per month, or \$468–\$480 per year (a 57 percent average reduction), in prior evidence from the Massachusetts exchange (Finkelstein, Hendren, and Shepard 2019). It is somewhat larger than the 20–26 percent impact of introducing Massachusetts's uninsurance penalty (Chandra, Gruber, and McKnight, 2011).[39] Therefore, auto-enrollment has an impact comparable to sizable changes in financial incentives.

Despite its large impact, the targeted nature of the auto-enrollment policy—applying only to people who had already qualified for coverage—meant that its impact on overall uninsurance was more modest. Using ACS data, we estimate that Massachusetts had about 300,000 uninsured people in 2009, of whom about 62,000 had incomes below poverty and were likely CommCare eligible. Relative to this denominator, auto-enrollment's 14,900-person impact (see Supplemental Appendix B.3) represents a 24 percent decline in the eligible uninsured population.

IV. Targeting Implications of Auto-Enrollment

In this section, we study the targeting implications of auto-enrollment. Who are the marginal enrollees, and how do they compare to inframarginal (active) enrollees? How does auto-enrollment affect the market risk pool? What mechanisms may explain passive individuals' failure to actively enroll? These questions matter both for the policy's positive economic implications and for its welfare interpretation. Section IVA provides descriptive evidence on targeting implications, comparing marginal (passive) versus inframarginal (active) enrollees on characteristics related to the value and cost of insurance. Section IVB shows evidence that auto-enrollment is unlikely to be (invalidly) enrolling individuals with duplicate private health insurance. Section IVC assesses mechanisms, both rational and behavioral, for why a small hassle deters so many people from taking up free coverage.

A. *Targeting Implications and Impact on Market Risk Pool*

To study the targeting implications of auto-enrollment—that is, inferring its marginal versus inframarginal enrollees—we employ two methods. The first is motivated by our finding in Section IIIA that the number and composition of active

[39] Evidence from the ACA—which involves a somewhat higher-income population than in CommCare—suggests smaller impacts of both subsidies and uninsurance penalties (see, e.g., Frean, Gruber, and Sommers 2017; Lurie, Sacks, and Heim 2019). The 32 percent impact of auto-enrollment is even larger relative to subsidies and penalties based on these ACA estimates.

enrollees is unaffected by the end of auto-enrollment in 2010. This suggests that passive behavior is in a sense "exogenous" to the policy environment. If correct, this means that *observed passive* enrollees (prior to 2010) are also *marginal* enrollees who would not have enrolled without the policy in place.[40] Thus, we are in the fortunate position of directly observing who is a marginal versus inframarginal enrollee (something that is rarely true in the targeting literature). A simple comparison of passive versus active enrollees, therefore, should faithfully characterize marginal versus inframarginal individuals. We use this method for our main analysis, controlling for entry timing using cohort fixed effects.[41]

Our second method uses the *policy change* to infer marginal enrollee characteristics from compositional changes in new enrollment at the start of 2010. This method has the advantage of not requiring the assumption of exogenous passivity. However, it is statistically much less powerful and may suffer problems if enrollee attributes are trending over time. We therefore implement it as a robustness check, using the simple active versus passive comparison for our main estimates.

Characteristics of Passive Enrollees.—Table 1 shows the results from our main method comparing passive versus active enrollees. Overall, the results suggests four main patterns about passive (relative to active) enrollees:

Younger, Healthier, and More Male: Passive enrollees are younger by 3.8 years on average and are 22 percent more likely to fall into the youngest age (19–34) group. They are also more likely to be male, with an especially large share (44 percent higher) of young men age 19–34, a group often called "young invincibles" in insurance discussions. Likewise, passive enrollees are healthier, with 33 percent lower rates of any chronic illness and 49 percent lower rates of severe chronic illness. Overall, passive enrollees have 36 percent lower medical risk scores, a measure of predicted medical costs based on age, sex, and diagnoses.[42] Figure 6 visualizes these patterns in a different way by plotting the passive enrollment rate by age, sex, and risk score groups. Passive rates decline with age and risk, though they exceed 20 percent even for the oldest and sickest groups.

Lower Medical Costs: Consistent with their youth and health, passive enrollees incur 44 percent lower monthly medical costs (\$228 per month versus \$408 for active enrollees) and are more likely to have 0 spending. The slightly larger gap for spending (-44 percent) relative to risk score (-36 percent) suggests passive enrollees may also be unobservably healthy. Because the government pays insurers

[40] More generally, one could think of passive enrollees as falling into two groups: (i) "always passives," who are passive regardless of the policy, and (ii) "conditional passives," who are passive under auto-enrollment but make sure to actively enroll when it is gone. Our evidence in Section IIIA suggests that there are few if any conditional passives in our setting.

[41] This lets us control for any time trends (e.g., medical cost growth) that could affect results if passive rates vary over time. In practice, these fixed effects have little impact on results. The specific method is as follows. Let $Y_{i,c}$ be a characteristic/outcome for new enrollee i who joins CommCare in entry cohort c (i.e., in a given year-month). We regress $Y_{ic} = \alpha_c + \delta \cdot \mathbf{1}\{Passive_i\} + \varepsilon_{i,c}$, which includes a cohort fixed effect (α_c). Table 1 reports the mean for active enrollees ($\bar{Y}_{active}$), the adjusted mean for passive enrollees ($= \bar{Y}_{active} + \delta$), and the difference between the two (δ).

[42] We use the HHS-HCC risk score (silver-CSR version), as used in the ACA Marketplaces, calculated based on diagnoses observed on claims during an enrollee's first 12 months enrolled.

TABLE 1—TARGETING IMPLICATIONS: COMPARING ACTIVE VERSUS PASSIVE ENROLLEES

Variable	Active Enr. (1)	Passive Enr. (2)	Diff (3)	(SE) (4)	% Diff (5)
Panel A. Age and sex					
Average age (years)	35.6	31.8	−3.8	(0.1)	−11
Age 19–34	0.535	0.652	+0.118	(0.003)	+22
Age 35–54	0.339	0.271	−0.068	(0.003)	−20
Age 55+	0.126	0.077	−0.049	(0.002)	−39
Share male	0.538	0.625	+0.087	(0.003)	+16
Male age 19–34	0.286	0.411	+0.125	(0.003)	+44
Panel B. Health status and medical spending					
Any chronic illness	0.641	0.427	−0.215	(0.003)	−33
Severe chronic illness	0.158	0.081	−0.077	(0.002)	−49
Risk score (HCC)	1.011	0.644	−0.367	(0.015)	−36
Average cost ($/month)	$408	$228	−$181	(5.6)	−44
Any spending (> $0)	0.894	0.709	−0.185	(0.003)	−21
Panel C. Income and area disadvantage					
Income/poverty line	0.248	0.200	−0.049	(0.004)	−19
High-disadvantage area	0.320	0.401	+0.082	(0.003)	+25
Share Black (in zip code)	0.082	0.106	+0.024	(0.001)	+29
Share Hispanic (in zip code)	0.137	0.162	+0.025	(0.001)	+18
Near safety net hosp./CHC	0.371	0.458	+0.087	(0.003)	+23
Panel D. Duration enrolled					
Average (month)	16.5	11.9	−4.6	(0.1)	−28
Share 1–3 months	0.154	0.228	+0.075	(0.002)	+48
Share 12+ months	0.559	0.441	−0.119	(0.003)	−21
Share 16+ months	0.297	0.168	−0.129	(0.003)	−43

Notes: The table shows differences in characteristics/outcomes for passive versus active enrollees in our main sample of below-poverty new CommCare enrollees during FY 2008–2009. Estimates control for entry cohort fixed effects and (for all variables except "Duration" in panel D) are weighted averages by months enrolled (capped at 12 months). Health and cost measures are based on claims during the enrollee's first 12 months enrolled. Chronic illnesses follow a classification of ICD-9 diagnosis codes shared with us by David Cutler. Risk score is based on the HHS-HCC model (silver-CSR version) used for risk adjustment in the ACA, renormalized to have mean 1.0 in the CommCare data. Income refers to family income as a share of the federal poverty level. High-disadvantage areas are zip codes (ZCTAs) in the seventy-fifth percentile or higher of the social deprivation index (SDI) produced by the Robert Graham Center (2005-2000) based on ACS data (see https://www.graham-center.org/maps-data-tools/social-deprivation-index.html), which also includes data on zip code–level shares of Black and Hispanic people. "Near safety net hospital or Community Health Center (CHC)" refers to the share of enrollees living in zip codes within two miles of one of these facilities (Google Maps API 2013).

using risk-adjusted capitation, passive enrollees' lower risk scores imply that the government also incurs lower costs to cover them.[43]

More Economically Disadvantaged: Passive enrollees are more disadvantaged across several metrics. Their incomes are slightly lower (20 percent versus 25 percent of poverty). Their differences in neighborhood characteristics (based on zip code) are larger. Passive enrollees are 25 percent more likely to live in a zip code in

[43] Up to 2009, CommCare used a crude risk adjustment system that varied rates by age-sex-region cells. Under this system (which we can observe), the average government payment for passives was 8 percent less than for active enrollees ($344 versus $373 per month). Starting in 2010, the program shifted to a stronger diagnosis-based risk adjustment, similar to the HCC risk scores we report. Although we lack full data until 2011 on CommCare's risk adjuster, the 36 percent lower HCC scores suggest rates would be substantially lower for passives.

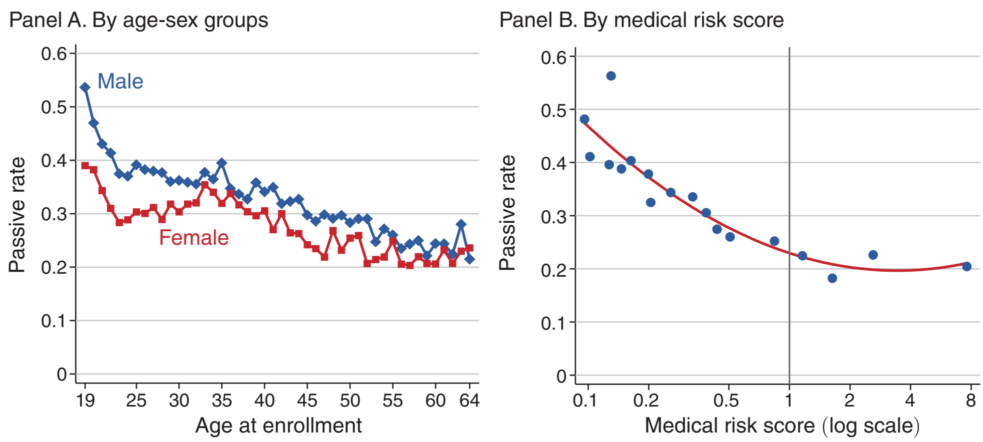

Figure 6. Passive Enrollment Rate by Age, Gender, and Medical Risk

Notes: The figure plots variation in the passive enrollment rate—the share of new enrollees who join passively—by age-sex groups (panel A) and medical risk score bins (panel B). The data are for our main sample: new enrollees in the relevant below-poverty income group during fiscal years 2008–2009. The medical risk score is the HHS-HCC risk score (silver-CSR version) used by the ACA Marketplaces, calculated based on diagnoses observed on claims during the first 12 months of enrollment.

the top quartile of the Social Deprivation Index, a measure based on census data.[44] Their zip codes include a higher share of Black and Hispanic residents.

Shorter Durations: Passive enrollees are enrolled for shorter periods, with average durations 4.6 months (or 28 percent) shorter. Although we do not observe the reason for these shorter spells, an analysis of the time pattern of exits (see Supplemental Appendix C.2) suggests a combination of two factors: (i) a higher rate of brief 1–3 month spells and (ii) a higher exit rate during annual eligibility redetermination (12–14 months into the spell). The latter is consistent with a failure to complete redetermination paperwork, another administrative hassle.

A natural question is whether measured risk differences are driven by passive enrollees' shorter durations (see "Shorter Durations" above), which limits the period over which medical conditions can be observed in claims data. In practice, this does not appear to be a major source of bias. Supplemental Appendix C.1 shows that health differences are robust to using shorter measurement periods (including using just the first month enrolled) and to examining a balanced panel of active and passive enrollees enrolled for the same duration.

In line with their residence in lower-income neighborhoods, passive enrollees are also more likely to live nearby (within two miles) a safety net hospital or community health center. This proximity raises the question of whether they use

[44] We use the Social Deprivation Index (SDI) developed by the Robert Graham Center (see https://www.grahamcenter.org/maps-data-tools/social-deprivation-index.html, accessed January 1, 2025). SDI is an index of area-level deprivation derived from ACS data, based on income, education, housing, employment, and other demographics. We define "high disadvantage" as neighborhoods in the top quartile of the SDI based on the national distribution.

more "uncompensated care"—an important social cost of uninsurance (Finkelstein, Mahoney, and Notowidigdo 2018) that we include in our model in Section I. Supplemental Appendix C.3 presents analysis to test this idea. A limitation is that we cannot directly observe care used by active versus passive individuals when *uninsured*. However, based on care use when insured, passive enrollees obtain a larger share of their care from standard sources of uncompensated care, including emergency rooms and safety net hospitals.

Interpreting the Differences.—Overall, this evidence is consistent with the two main features of our ordeals targeting framework in Section I: *self-targeting* and *adverse selection*. Consistent with self-targeting, passive enrollees (those screened out by ordeals) have attributes consistent with lower demand (value) for health insurance. This includes the young and healthy, who on average need less medical care, and shorter-duration enrollees, who may only have a brief need for public coverage (e.g., between jobs). Demand for health insurance also tends to be low among the poor (Finkelstein, Hendren, and Luttmer 2019; Finkelstein, Hendren, and Shepard 2019; Tebaldi forthcoming).

But consistent with adverse selection, these same low-demand individuals also incur much lower costs. Passive enrollees incur 44 percent lower monthly medical costs, and including their shorter durations, their average per spell costs are 60 percent lower. This is natural in an adverse selection market where both value and costs are driven by an enrollee's medical risk (and by their enrolled duration). As a result, our theory suggests that *self-targeting* may not translate into *socially* beneficial targeting. We evaluate this idea more formally using our empirical model in Section V.

Robustness: Inference Using the Policy Change (*and Risk Pool Impacts*).—As a robustness check, we use the 2010 policy change to infer marginal enrollees. Prior to 2010, new enrollees include both active and passive individuals; afterward, only active choosers enroll. Marginal enrollees' characteristics, therefore, can be inferred from the *compositional* change at the start of 2010. To implement this, we run DD regressions analogous to equation (13) but with a dependent variable of characteristics/outcomes of new enrollees. Regressions are run on individual-level data, clustering standard errors at the income group-by-month level.

Figure 7 shows the raw data and DD estimates for two key risk pool variables: average risk score (panel A) and average cost (panel B) for new enrollees. There is a clear increase in both measures for the treatment group (red) relative to controls (green) after auto-enrollment is suspended.[45] The effects are large, with DD estimates suggesting a 0.146 increase in average enrollee risk (implying 14.6 percent higher costs) and \$57.6 increase in average monthly cost (also about a 15 percent increase). This implies that marginal enrollees screened out are lower risk and lower-cost, just as we found in Table 1. We can further compare the methods quantitatively by calculating what Table 1 predicts for the analogous change in average

[45] Counterintuitively, prior to 2010, the controls have higher risk scores but similar costs to the treatment group, and this pattern flips in 2010+. This occurs because CommCare provided more generous benefits to the treatment group, including dental care and slightly lower copays, which results in higher costs partly through a moral hazard effect (see Chandra, Gruber, and McKnight 2014).

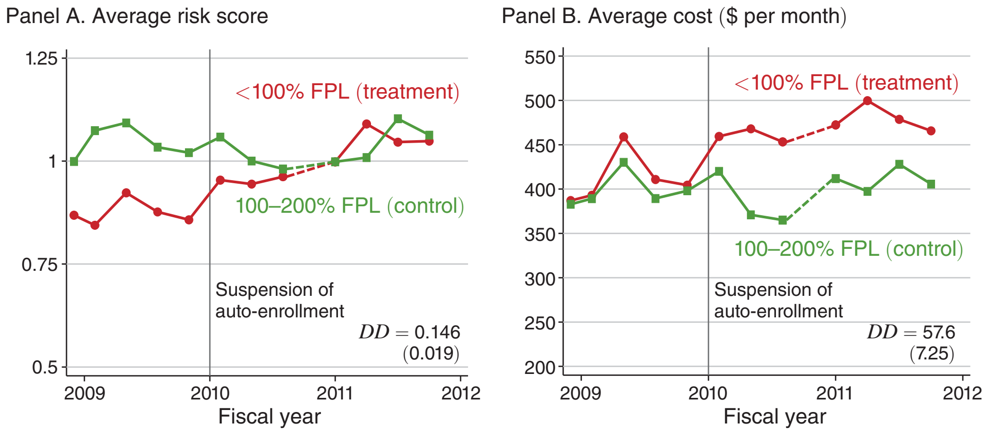

Figure 7. Effect of Auto-Enrollment Suspension on Enrollee Risk Pool

Notes: The figure shows data on average risk score (panel A) and monthly medical costs (panel B) for new enrollees, and estimates of the DD specification (13) using quarterly time periods. Each panel shows trends for below-poverty enrollees (the treatment group) versus 100–200 percent of poverty enrollees (the control group). The temporary reinstatement period is excluded (as indicated with dashed lines). When auto-enrollment is suspended, average risk score rose by 14.6 percent of the market average (which is 1.0), and average medical costs rose by $57.60 per month, also about a 15 percent increase. Both are consistent with the suspension of auto-enrollment resulting in higher-cost risk pools.

risk score and cost, assuming that passive behavior is exogenous.[46] This exercise predicts a 0.119 increase in average risk score and $58.8 increase in average cost, which are very close to (and statistically indistinguishable from) the DD estimates in Figure 7.[47]

B. *Do Passive Enrollees Have Duplicate Private Insurance?*

A relevant question for the targeting implications of auto-enrollment is whether it enrolls people who *already* have private health insurance, making CommCare duplicative. Although duplication is not supposed to occur—CommCare applicants must attest to not having access to any other health insurance (including any *offer* of job-based coverage)—enforcement could be imperfect. If auto-enrollment "overenrolls" individuals who already have other coverage, it would be a failure of "statutory targeting" based on program eligibility rules, something that has been observed for transfer programs in a developing country context (Alatas et al. 2016).

[46] To do so, note that for any variable Y, $\bar{Y}_{Pre2010} = s_P \bar{Y}_P + (1 - s_P)\bar{Y}_A$ and $\bar{Y}_{Post2010} = \bar{Y}_A$, where "$P$" and "$A$" subscripts refer to passive and active enrollees. Therefore, $\Delta\bar{Y} = \bar{Y}_{Post2010} - \bar{Y}_{Pre2010} = s_P \cdot (\bar{Y}_A - \bar{Y}_P)$. We calculate $\Delta\bar{Y}$ using the estimates for $\bar{Y}_A$ and $\bar{Y}_P$ in Table 1 and $s_P = 0.326$ from Figure 5.

[47] Supplemental Appendix C.4 shows a similar robustness analysis for all variables in Table 1; the Supplemental Appendix also describes the methods in greater detail. For all variables, our main method and the DD estimates are directionally similar, always generating estimates of the same sign. Moreover, the methods usually yield quantitatively similar estimates with overlapping confidence intervals.

To test this story, we draw on evidence from the Massachusetts APCD to measure rates of simultaneous duplicate coverage in CommCare and private insurance, a measure of whether "overenrollment" occurred in practice.[48] We define the "duplication rate" as the share of CommCare enrollment months during which the member was simultaneously enrolled in other private insurance.[49] Supplemental Appendix D.1 provides additional details on the data and method.

Overall, we find little evidence of meaningful duplicate coverage in CommCare. The average duplication rate is quite low, just 3.1 percent of enrollee-months, and the rate is even lower at the beginning of enrollment spells when auto-enrollment occurs (see Supplemental Appendix Figure A.13). Moreover, there is little evidence that duplication is higher for passive enrollees. Although we cannot distinguish active versus passive enrollees in the APCD, we can study how duplication rates *change* for new enrollees into CommCare just before versus after auto-enrollment is suspended in 2010. In practice, the duplication rate rises slightly after the policy change, consistent with marginal (passive) enrollees having lower duplication rates. However, duplication rates are low both before and after the change. Our overall conclusion is that duplicate coverage is rare and is unlikely to explain failure to actively take up coverage.

C. *Mechanisms: Why Do People Fail to Take Up Free Insurance?*

Why do so many people fail to enroll in free health insurance when faced with a small hassle? In this subsection, we provide descriptive evidence to assess the mechanisms involved, including both rational and behavioral explanations. We argue that non-enrollment is unlikely to be explained by fully rational and informed stories, in which individuals are passive because they do not need or benefit from (free) public health insurance. Instead, we argue that behavioral "frictions" are likely involved, with the most likely frictions being inattention and limited understanding of program rules.

Evidence against Fully Rational Non-enrollment.—We start by providing evidence against fully rational and informed non-enrollment. We start by noting that several facts about the institutional setup make this a priori less likely. First, everyone in our sample—including passive enrollees—has already *chosen* to apply for public coverage (in step 1 of the process). This suggests that they have some awareness of the program and a desire to enroll. Moreover, the insurance is free and extremely generous, with 0 deductible and close to 0 cost sharing (the actuarial value exceeds 99 percent). Although there are some limits (e.g., on networks), it seems implausible that enrollees would face fewer limits or costs if they were uninsured, the relevant counterfactual.

[48] Ideally, we would want to measure the *counterfactual* of whether CommCare enrollees obtain other insurance if they were (exogenously) kicked out of CommCare. While we cannot measure this counterfactual directly, the observed duplication rate provides suggestive evidence on whether overenrollment is a problem in general.

[49] We do not include duplicate coverage in CommCare plus Medicaid because the two programs use a unified enrollment system, which should automatically prevent duplicate enrollment. Most of the same insurers operate in both programs, and we have some concerns that the insurance type is sometimes mislabeled, which could lead to false positives.

Some simple facts further indicate that passive enrollees are likely to obtain meaningful benefits from health insurance. Although passive enrollees are *relatively* healthy, they are not *uniformly* so. Indeed, over 40 percent have a chronic illness, and 8 percent have a severe chronic illness (Table 1). Their average spending of $228 per month is large relative to their very low incomes (the individual poverty line in 2009 was $903/month). Supplemental Appendix Figure A.11 shows that passive enrollees experience meaningful rates of medical shocks (e.g., high-cost months, emergency hospitalizations) that while less frequent, still occur 60–75 percent as often as for active enrollees. Further, Figure 6 shows that even among the oldest and sickest enrollees, passive rates exceed 20 percent. Thus, while good health is predictive of being passive, it is clearly not the full explanation.

Finally, we argue that access to charity care is unlikely to be a perfect substitute for formal insurance that drives its (true) value down to near zero. First, passive enrollees use a meaningful amount of care in categories that are less available via charity care, including prescription drugs.[50] Second, the prior literature on the value of insurance to the poor suggests that while value is *low*, it is far above *zero*. For instance, a key paper in this literature, Finkelstein, Hendren, and Luttmer (2019), finds that the individual value of insurance is just 20–48 percent of insured medical expenses. Applied to our passive enrollees (who spend $228 per month when insured), this would imply a value of $46 to $109 per month—or $550 to $1,300 over a typical 12-month enrollment spell. This is a sizable amount. For instance, it is comparable to forgone benefits from failing to take up the EITC or SNAP (Bhargava and Manoli 2015; Finkelstein and Notowidigdo 2019) and from losses due to insurance plan choice errors (Abaluck and Gruber 2011; Bhargava, Loewenstein, and Sydnor 2017).

Evidence on Behavioral Frictions.—We test two types of behavioral explanations: (i) those in which the *complexity of plan choice* is the key barrier and (ii) those in which *taking action* is the key barrier, for instance, because of inattention or misunderstanding the steps required to enroll. We find little evidence of (i) but suggestive evidence consistent with (ii).

Choice Overload.—One reason people might be passive when asked to select a health plan is that they become overwhelmed by the choice, as in models of "choice overload" (Iyengar and Kamenica 2010). We note that choice overload is a priori less likely in the CommCare setting, which featured a relatively simple choice set with at most four to five plans available.[51] Further, the passive enrollment rate is unrelated to the choice set size, which varies across areas due to selective insurer entry. Supplemental Appendix Table A.7 shows that the passive rate varies in a narrow range of 33–35 percent across all choice set sizes, including at 34 percent in

[50] We observe that 25 percent of passive enrollees take a regular prescription medication every month they are enrolled, with an average cost of $45 per month. Over a typical 12-month enrollment spell, these prescription costs alone would add up to $540.

[51] There were four plans prior to 2010, and a fifth (CeltiCare) entered during 2010. This is much simpler than other US insurance programs. For instance, Medicare Advantage features an average choice set with 33 options (see https://www.kff.org/medicare/issue-brief/medicare-advantage-2021-spotlight-first-look), and Medicare Part D feature 25–35 plan options (see https://www.kff.org/medicare/fact-sheet/an-overview-of-the-medicare-part-d-prescription-drug-benefit).

areas with just a *single* plan (i.e., no real choice). Moreover, passivity does not change significantly when a plan enters or exits a region. We conclude that there is little evidence that choice overload is responsible for passive behavior in this context.

Inattention or Misunderstanding.—A second type of reason for passivity is that some people are inattentive or misunderstand the steps required to enroll in coverage.[52] If so, requiring an additional step of action—even a seemingly simple step —will lead some individuals to "fall through the cracks" and not enroll. We present three sets of facts consistent with a role for inattention and/or misunderstanding. These are discussed here, with the underlying analyses presented in Supplemental Appendix C.8.

- **"Lost in the Mail":** A natural reason for inattention is if some people do not receive the approval letter instructing them how to actively enroll. Anecdotally, address errors are a common problem in welfare programs, partly because of greater residential instability in low-income populations. To test for this, we construct a proxy for "address mismatches" based on observing different zip codes in CommCare's enrollment file (based on the address used in administrative mailings) versus on the enrollee's first observed medical claim (submitted by the medical provider, often based on paperwork filled out at a visit). As detailed in Appendix C.8, address mismatch is surprisingly common, occurring for about one-third of enrollees. Moreover, it is predictive of passive behavior. After conditioning on the sample with an observed claim in their first 6 months, the passive rate is 28 percent for mismatched, about 3 percentage points (or 13 percent) higher than for nonmismatched people. This pattern is robust to controlling for demographics, health, and timing of the first claim.
- **Special Barriers:** Misunderstanding may be more common in groups that face special barriers to interacting with the state and learning about take-up rules. This idea is consistent with the evidence, shown above, that socioeconomically disadvantaged groups are more likely to be passive. Another such group is immigrants, who likely face greater language and cultural barriers.[53] Consistent with this, passive rates are higher for immigrants (41 percent rate), about 7 percentage points (or 21 percent) higher than for nonimmigrants (34 percent).
- **Cross-Program Transitions:** Misunderstanding or inattention may be more common when people transition between public programs in which take-up rules differ. We observe two types of transitions in our data: (i) a large shift of enrollees from the state's uncompensated care pool to the CommCare exchange in early 2007 and (ii) regular transitions from Medicaid into CommCare (e.g., due to changes in income, age, or family status). Active plan choice was not required in either the UCP or Medicaid, so there may be greater confusion in

[52] There is substantial evidence of limited attention/understanding and other behavioral frictions for consumer choice *among* health plans (e.g., Abaluck and Gruber 2011; Handel 2013; Ericson 2014; Handel and Kolstad 2015). Thus, it is plausible to think that the same issues might affect whether people enroll in health insurance in the first place.

[53] Immigrants were excluded from our main analysis sample, as discussed in Section IIC. For this analysis, we augment the main sample to re-include them.

these groups about enrollment processes in CommCare. Consistent with this, passive rates are much higher for these transitions. People transitioning from the UCP had a 60 percent passive rate (versus 40 percent for other enrollees at the same time in early 2007). People transitioning from Medicaid have a 39 percent passive rate (versus 31 percent for non-Medicaid enrollees). The latter is partly driven by very high passivity for kids transitioning off of Medicaid at age 19 (Jácome 2020), but passive rates are higher for Medicaid transitions even controlling for age, gender, and health covariates.

V. Empirical Model and Policy Trade-offs

In this section, we empirically apply our model from Section I to our health insurance setting in Sections VA–VC, using a combination of our administrative data, the auto-enrollment natural experiment, and outside estimates. We use the estimates to assess the question with which we started the paper: How well do ordeals work to target enrollment in health insurance?

A. *Model Implementation*

Our ordeals welfare framework requires estimates of four objects for enrollees: (i) the direct medical cost of insurance, C_i; (ii) the enrollee value of insurance, W_i; (iii) social spillovers, E_i; and (iv) fiscal externalities, FE_i. Together, these let us calculate $V_i^{Soc} = \mu W_i + E_i$ (for various assumptions on the social welfare weight μ) and $C_i^{Net} = C_i - FE_i$, which together are sufficient for net social welfare, $\gamma_i = V_i^{Soc} - C_i^{Net}$.

Our natural experiment and rich insurance claims data let us directly measure the distribution of marginal (passive) and inframarginal (active) enrollees and their medical costs (C_i). We assume that the government either directly pays medical expenses (as in traditional Medicare and Medicaid) or engages in zero-profit contracting with private insurers (as we find is roughly true in Massachusetts).[54] In both cases, medical costs for individual i in the claims data are a reasonable estimate of the government's marginal cost when they enroll in insurance (i.e., C_i in the model).[55] With this assumption, our claims data give us a direct estimate of C_i and the average cost for active ($\bar{C}_1$) and passive ($\bar{C}_0$) enrollees.

[54] Supplemental Appendix Table A.9 shows evidence of this zero-profit contracting for the below-poverty population, for whom CommCare negotiated a separate set of payment rates directly with insurers (as opposed to the bidding system used for higher-income groups). The table compares the government's payment and insurer's cost for active and passive enrollees. Insurers earned small overall margins (of about 4 percent, or $16 per enrollee-month), despite overpaying for passive and underpaying for active enrollees. The table also shows that had the exchange paid using more sophisticated risk adjustment, this group-specific over-/underpayment would shrink, but overall profit margins would remain near zero. We interpret this as evidence that (i) CommCare was able to negotiate lower average prices for the below-poverty population as a whole because of the inclusion of healthier auto-enrollees, and (ii) average prices paid approximately reflect average costs.

[55] This relationship is immediate when the government directly pays claims. In the zero-profit contracting case, the relationship follows from the fact that the government's total payments equal insurers' total cost for all enrollees. When i is enrolled, insurers' total costs increase by C_i, and to maintain zero profits, the government's extra cost is also C_i. Note that this analysis abstracts from any nonmedical administrative costs (for either government or private insurers), which we cannot directly measure in our claims data.

To estimate the remaining items (ii)–(iv), we combine what we do observe with information from other studies and data sources. In what follows, we describe our strategy for estimating each term.

Uncompensated Care Costs.—The main component of social and fiscal externalities is uncompensated care, so we start with estimating it. In our data, we observe medical costs when insured, C_i.[56] To estimate uncompensated care costs that i would incur if *uninsured*, we proceed in two steps. First, the uninsured use less care than the insured because of moral hazard, which we assume increases costs by a constant factor, $1 + MH$. Second, the uninsured themselves pay only a share, $\phi < 1$, of their medical bills, with uncompensated care covering the other $1 - \phi$. Thus, uncompensated care costs equal

$$C_i^{UC} = \left(\frac{1-\phi}{1+MH}\right) \cdot C_i. \tag{14}$$

Estimating C_i^{UC} requires values for ϕ and MH. For our baseline estimates, we draw on the analysis of Finkelstein, Hendren, and Luttmer 2019 of the Oregon Health Insurance Experiment. They estimate a moral hazard effect of $MH = 33.3\%$ and an uninsured out-of-pocket share of bills of $\phi = 0.21$, both of which we treat as constant across enrollees.[57] Using this method, therefore, we estimate $C_i^{UC} = 0.59\, C_i$.

We consider two alternatives in sensitivity analysis. First, as extreme upper and lower bounds, we consider $\phi = 0$ (full uncompensated care) and $\phi = 1$ (implying $C_i^{UC} = 0$). Second, we construct new estimates using data from a Massachusetts program, the Health Safety Net (HSN), that covers a subset of medical expenses for uninsured low-income adults. The HSN is an uncompensated care pool that (unlike most similar programs) pays based on formal claims, which are observable in the state's APCD. We use these data, combined with estimates of total uninsurance from the ACS, to estimate uncompensated care costs by age-sex group, which we then project onto our CommCare data. The method involves several assumptions, which we detail in Supplemental Appendix E.

Social and Fiscal Externalities of Insurance.—Having estimated uncompensated care costs, we divide its incidence between the government (part of FE_i) and private

[56] Technically, we observe *realized* medical spending, which differs from *ex ante* expected costs due to the realization of an *ex post* health shock. We assume throughout that this shock is idiosyncratic and additively separable, so that it averages to zero in any sufficiently large group g (e.g., passive enrollees). Formally, let C_i be realized costs and $E[C_i]$ be expected costs. We assume that $C_i = E[C_i] + \omega_i$, with $E[\omega_i] = 0$ and ω_i independent of all other variables in the model including group membership. Under these assumptions, $\bar{C}_g = \frac{1}{N_g}\sum_{i \in g} C_i = \frac{1}{N_g}\sum_{i \in g}(E[C_i] + \omega_i) \rightarrow \frac{1}{N_g}\sum_{i \in g} E(C_i)$ for large enough N_g.

[57] Finkelstein, Hendren, and Luttmer (2019) estimate that in the Oregon experiment, health insurance increases annual medical spending by \$900, which is 33.3 percent of the control complier (uninsured) mean of \$2700. They estimate that control compliers (the uninsured) spend \$569 per year in out-of-pocket expenses, which implies $\phi = 569/2700 = 0.21$. We treat MH and ϕ as constant across enrollees, implying C_i^{UC} scales proportionally with insured costs, since it is unclear how to estimate heterogeneity. If anything, the evidence suggests that C_i^{UC} are disproportionately larger for passives, suggesting we may (conservatively) understate their relative efficiency.

providers (part of E_i). We assume that the government bears a fixed share, $\psi_G \in [0,1]$, of costs, which implies

$$(15) \qquad FE_i = \psi_G \cdot C_i^{UC} \quad \text{and} \quad E_i = (1 - \psi_G) C_i^{UC}.$$

Note that this assumes no other externalities of insurance besides uncompensated care, which is a conservative assumption.[58] To estimate ψ_G, we draw on the evidence from Garthwaite, Gross, and Notowidigdo (2018), who study the impact of uninsurance on hospital uncompensated care costs and profits. They find that for every $1 higher uncompensated care costs, hospitals absorb $0.60–$0.67 in lost profits. In our main estimates, we set $\psi_G = 0.635$, the midpoint of this range.

Enrollee Value of Insurance.—Estimating value (or WTP) is challenging in our main sample because of a lack of price variation—all plans are free. Moreover, the presence of frictions raises concerns about inferring low WTP directly from passive behavior, which may be a consequence of enrollees having high frictions (e.g., inattention or forgetfulness). To make progress, we follow the "rational consumer benchmark" approach described by Bernheim and Taubinsky (2018), which has also been implemented by Bronnenberg et al. (2015) and Allcott, Lockwood, and Taubinsky (2019). The approach involves estimating preferences among a well-informed reference population (the "benchmark") in order to impute the WTP of another group. We use price variation for higher-income CommCare enrollees (150–250 percent of poverty) who all pay positive prices, replicating and extending the demand estimation method of Finkelstein, Hendren, and Shepard (2019). We then project these demand estimates onto our below-poverty population at the level of detailed observables (age-sex-risk group cells).

This exercise rests on two assumptions: (i) that higher-income enrollees reveal their WTP when making active choices and (ii) that age-sex-risk observables are sufficient for projecting WTP onto lower-income groups. Assumption (i) is consistent with a model of pure inattention frictions (e.g., forgetting to act) that prevent passive types from enrolling but do not bias demand estimates for active choosers. This assumption implies that demand reveals true WTP *among the sample of higher-income active enrollees* (150–250 percent of poverty).[59] Assumption (ii) allows us to impute this WTP distribution onto our lower-income (0–100 percent of poverty) population of interest, conditional on age-sex-risk cells. However, it is vulnerable to concerns about selection on unobserved preferences. To address this, we examine robustness to alternative assumptions about unobserved sorting, described in greater detail below.

We summarize the method here, with details and estimates presented in Supplemental Appendix F. Finkelstein, Hendren, and Shepard (2019) use RD variation in subsidies and premiums to estimate a demand (WTP) curve for insurance.

[58] For instance, there is evidence that health insurance for kids leads to long-run economic gains that boost future tax revenue (Brown, Kowalski, and Lurie 2020) and that insurance for young adults reduces crime (Jácome 2020). We do not include these since it is unclear how to estimate their distribution for different types of enrollees.

[59] Of course, this benchmark may under-/overstate the value of insurance if higher-income active choosers suffer from behavioral biases or liquidity constraints. Our analysis that scales enrollee welfare by a range of social welfare weights, μ, can partly address this concern.

They observe three income thresholds at which premiums increase discretely: from \$0 to \$39 per month (at 150 percent of poverty), from \$39 to \$77 (at 200 percent of poverty), and from \$77 to \$116 (at 250 percent of poverty). By observing how much enrollment falls at each threshold, they infer points on an insurance demand curve. These can be linearly connected and extrapolated to generate a full demand curve $D(s)$, where $s \in [0,1]$ indexes people from highest to lowest WTP.

To adapt Finkelstein, Hendren, and Shepard's (2019) method to our problem, we make two adjustments. First, we use 2009–2011 data, matching our analysis period. Second, we use the micro-data to estimate demand separately by cell of $g = \{$age group, sex, risk score bin$\}$. We use roughly 5-year age bins and quintiles of HCC risk score, with an additional category for the sickest 5 percent of enrollees. With a demand curve for each cell, $D_g(s)$, we project WTP onto each enrollee i in our below-poverty sample using the average WTP for their g cell, that is, $W_i = E\left[D_{g(i)}(s)\right]$, where the average is over s.[60] This method lets us capture WTP heterogeneity via observable factors included in g (age, sex, and medical risk). We also consider several assumptions for *unobserved* sorting between active versus passive enrollees, including no sorting, perfect sorting, and (for our baseline specification) unobserved sorting of "equal magnitude" to observed sorting, in a sense formalized in Supplemental Appendix F.[61]

We consider several alternatives in sensitivity analysis. In addition to variations on the demand-based approach (e.g., no or perfect unobserved sorting), we consider mapping insured medical costs (which we observe) to enrollee WTP using simple relationships estimated in the literature. Specifically, Finkelstein, Hendren, and Luttmer (2019) find that low-income Medicaid enrollees value insurance at 20–48 percent of insured costs (i.e., $W_i = \kappa \cdot C_i$ for $\kappa \in [0.20, 0.48]$); we report estimates for the endpoints of this range. We also consider a plausible lower bound in which WTP equals expected uninsured out-of-pocket (OOP) costs (with no value for risk protection), based on the framework underlying equation (14). This implies $W_i = \left(\frac{\phi}{1+MH}\right)C_i = 0.16\,C_i$ given the values of $\phi = 0.21$ and $MH = 0.333$.

Finally, we examine implied WTP for full insurance from a simple model of homogeneous risk aversion, under a benchmark assumption of no moral hazard or uncompensated care. Specifically, we simulate the value of insurance using observed

[60] Calculating average WTP (the conceptually correct statistic) requires using the linearly extrapolated portion of the demand curve, which comprises about the bottom 30–40 percent of demand. As robustness, we also examine the median and seventy-fifth percentiles of WTP, which are much less likely to be extrapolated. These generate smaller estimates of WTP but similar implications for the *relative* WTP and MVPF for active versus passive enrollees.

[61] Briefly, unobserved sorting relates to the range of s over which we average to calculate $W_i = E\left[D_{g(i)}(s)\right]$. For no sorting, we average over $s \in [0,1]$ for both actives and passives; therefore, WTP is equal for everyone *within* a g cell. For perfect sorting, we assume that within each g cell, actives comprise the highest 67 percent of WTP types ($s \in [0, 0.67]$), while passives comprise the lowest 33 percent of WTP types ($s \in [0.67, 1.00]$), where 33 percent is the overall share of passives in our data. For our baseline specification, we assume "equal" sorting on unobservables and observables. Formally, we calculate the probability that a random active enrollee is in a g cell with higher estimated WTP than a random passive enrollee. This is 56 percent in our data. We then set the averaging ranges of s so that this probability is also 56 percent *within* each g cell (i.e., unobserved sorting), which we show corresponds to $s \in [0, 0.96]$ for actives and $s \in [0.08, 1.00]$ for passives.

medical claims and an exponential utility function with coefficient of absolute risk aversion of $\alpha = 8.6 \times 10^{-5}$ taken from Handel and Kolstad (2015).[62]

Social Welfare Weight (μ).—Our key value statistic is the social value of insurance, $V_i^{Soc} = \mu W_i + E_i$, which scales enrollee WTP (W_i) by a social welfare weight, μ (and adds externalities, E_i). For simplicity, we use a constant μ for all eligible individuals, but we consider a range of values to capture distributional goals. Our baseline calculations use $\mu = 1$ (i.e., Kaldor-Hicks efficiency), but we consider a range of $\mu \in [0.5, 3.0]$ for robustness, where $\mu > 1$ allows for a social value of redistribution, while $\mu < 1$ captures tight public budgets.

Direct Cost of Ordeals, $L(\sigma)$.—Throughout this exercise, we focus only on the ordeal's targeting implications, that is, the "gains from targeting" piece of their welfare impact in equation (6). Implicitly, we ignore any *direct costs* of the ordeals ($L(\sigma)$), which we do not have a good way to estimate and which we believe are small in our setting. Because direct costs would only reinforce our finding that ordeals do not work well, we view this as a conservative assumption. However, measuring direct costs may be important in other settings where these are likely to be larger.

B. *Results: Model Estimates and Targeting*

Figure 8 shows our model's baseline estimates and the selection properties of auto-enrollment, comparing active versus passive enrollees in our main sample (as used in Table 1). Figure 8, panel A shows selection on social value, which includes both enrollee value and uncompensated care savings to private providers. Both the mean and the distribution of social value is lower for passive enrollees. On average, passive enrollees have both a lower private value of insurance (about 28 percent less than active enrollees) and use less uncompensated care when uninsured since they are healthier. Their average social benefit is \$143 per month, about 34 percent less than for active enrollees at \$217 per month. This finding that passive enrollees have lower (private and social) value of insurance than actives holds across every sensitivity analysis we consider, including different assumptions for demand estimation and alternate measures of uncompensated care (see Supplemental Appendix Table A.10). Our estimates, therefore, robustly suggest the active enrollment ordeal screens out low-value types, consistent with self-targeting and favorable sorting on value.

While there is favorable sorting on value, value and costs are also strongly correlated. Figure 8, panel B is a binned scatterplot showing the relationship between social value and net public costs, again comparing active and passive enrollees.[63]

[62] We compute expected utility, $\bar{u}_{g(i)} = E\left[\frac{-1}{\alpha} \cdot \exp(\alpha C_i)\right]$, separately by cells of $g = \{$age group, sex, risk score bin, passive versus active$\}$, taking the expectation over the observed distribution of monthly medical spending C_i within each cell. WTP for individuals in each cell is defined as the certainty equivalent, $W_i = \frac{1}{\alpha} \cdot \log(-\alpha \cdot \bar{u}_{g(i)})$.

[63] At the individual level, we observe realized—not expected—costs. We estimate expected medical costs by taking the mean of monthly realized costs (weighted by number of months enrolled) by cell of $g = \{$age group, sex, risk score bin$\}$ interacted with whether the individual was passive or active. Panel B of Figure 8 can therefore be thought of as displaying the joint distribution of social value and expected medical cost at the $\{$age group, sex, risk score bin, active versus passive status$\}$-cell level.

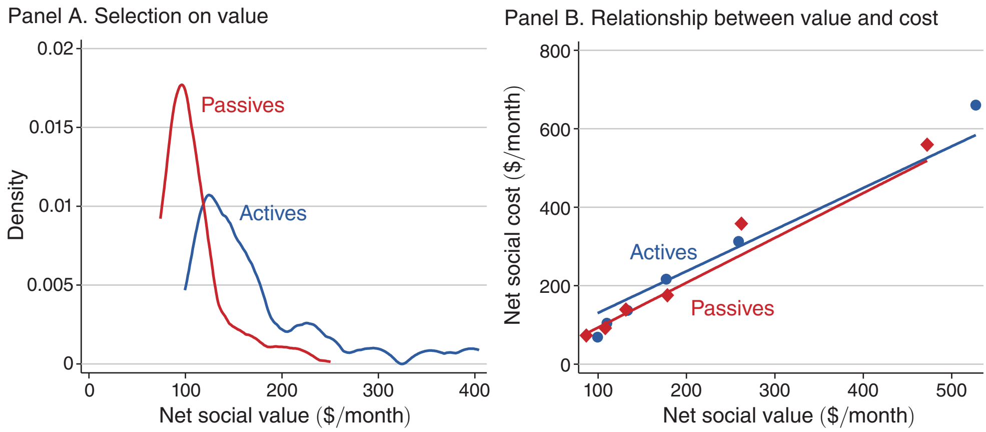

FIGURE 8. MODEL ESTIMATES: SELECTION ON VALUE AND COST

Notes: Panel A plots the density of our estimates of social value separately for both active (in blue) and passive (in red) enrollees, under our baseline demand and uncompensated care assumptions. For ease of visualization, only the bottom 90 percent of each distribution is shown in panel A. Panel B illustrates the joint distribution of social value and net costs for active (blue circles) versus passive (red diamonds) enrollees, along with respective best-fit lines. The sample for both figures is our main 2008–2009 new enrollee sample in the below-poverty group, just as in Table 1. See Section VA for the model estimation method. Both figures plot the distribution of estimates (mean WTP and mean costs per month, weighted by number of months enrolled) at the {age group, sex, risk score bin, active versus passive status}-cell level.

There is a strong positive correlation between value and cost that holds similarly for both active and passive enrollees. Moreover, the two best-fit lines are nearly on top of each other, suggesting that the ordeal achieves little sorting on residual costs (ω_i) conditional on value. Instead, passive enrollees are simply low-value types who also have (proportionally) lower costs. In contrast to the standard case considered in the ordeals literature, screening out low-benefit types is insufficient to make the ordeal well targeted.

Value-Cost Correlation and the Adverse Selection "Tax".—As discussed theoretically in Section I, a positive value-cost correlation, ρ, reduces the social gains from screening out low-value types since they also have low costs. The extent of sorting on cost relative to sorting on value is captured by the term $\hat{\beta} = \rho \cdot \sigma_C/\sigma_V$, which we call the "adverse selection tax" on targeting efficiency. In the classic ordeals case with constant or uncorrelated costs ($\rho = 0$), targeting efficiency is purely a function of value sorting. But as the value-cost correlation and the variance of costs increases, this tax becomes larger, which reduces targeting efficiency relative to sorting on value. Overall, the correction term for cost sorting—or the *rate of selection on cost* ($= \Delta C^{Net}/\Delta V^{Soc}$)—equals the sum of the adverse selection tax and any selection on residual costs (ω_i) uncorrelated with value (see equation 11).

Table 2 shows how this plays out using our estimates of social benefit and cost for both our baseline specification and several alternatives, using $\mu = 1$ for the social welfare weight on beneficiaries. Robustly across all specifications, we find a substantial positive value-cost correlation, ρ, which is 0.69 in our main specification. Correspondingly, we find substantial rates of selection on cost for the ordeal, exceeding 100 percent in both our baseline and 3 of the remaining 4 specifications.

Table 2—Value-Cost Correlation and Targeting

	Value and cost specification				
		Sensitivity analyses			
	Baseline (1)	No unobserved sorting (2)	Perfect unobserved sorting (3)	WTP = OOP costs (4)	Baseline w/ HSN uncomp. care estimates (5)
Panel A. Joint distribution					
Value-cost correlation (ρ)	0.70	0.69	0.67	1.00	0.21
SD of net cost (σ_C)	\$246	\$246	\$246	\$246	\$392
SD of net cost (σ_V)	\$156	\$155	\$183	\$147	\$115
Panel B. Effect of value-cost correlation					
Adverse selection tax ($\rho \times (\sigma_C/\sigma_V)$)	110%	110%	90%	167%	72%
Selection on residual cost ($= \Delta\omega$)	42%	103%	−31%	0%	283%
Total effect ($\Delta C^{Net}/\Delta V^{Soc}$)	152%	213%	59%	167%	354%

Notes: Column 1 shows results from our baseline model estimates, while columns 2–5 show sensitivity to alternative specifications. The sample is our main 2008–2009 new enrollee sample in the below-poverty group, just as in Table 1. See Section VA for the model estimation method. Panel A shows properties of the joint distribution of our estimates of social value V^{Soc} and expected net cost C^{Net}, computed at the level of demographic cells defined in Section VA. Panel B shows the implication of the joint distribution for targeting of an ordeal that screens on V^{Soc}, under a baseline assumption of Kaldor-Hicks efficiency ($\mu = 1$). The adverse selection tax, defined as the regression coefficient $\rho \cdot \sigma_C/\sigma_V$, gives the rate at which screening on value also generates screening on cost. We also estimate $\Delta\tilde{\omega}$, the extent to which the enrollment ordeal selects on residual costs (unexplained by social value), which is relative to ΔV^{Soc}.

The lone exception is the "perfect sorting" specification, which reflects an extreme assumption on how well ordeals sort on unobserved value. But even in the perfect sorting case, we estimate a rate of selection on cost of 58 percent; that is, the social gains from targeting are limited to $1 - 0.58 = 42\%$ of the active-passive difference in value. Thus, our results suggest that adverse selection tends to reduce, and in many cases overturns, the gains from screening out low-value enrollees.

Value-Cost Ratios and Targeting.—When the government pays the full cost of insurance, as in CommCare, the value-cost ratio for active and passive enrollees ($\bar{R}_g = \bar{V}_g^{Soc}/\bar{C}_g^{Net}$ for group g)—or social benefit per dollar of net government spending—is informative for targeting efficiency. Table 3 shows the value-cost ratios for both active and passive enrollees in our main sample. In our baseline model (with $\mu = 1$, shown in columns 1–2), we find a higher social value-cost ratio for passive enrollees at 1.00, compared to 0.85 for actives. Mechanically, this reflects the correction for value-cost correlation described above: passive enrollees' proportional cost difference (−44 percent) exceeds their difference in social value (−34 percent). Thus, under our baseline specification, the ordeal targets ineffectively ($\Delta\gamma = \bar{\gamma}_1 - \bar{\gamma}_0 < 0$) and results in backward sorting. In principle, it would be optimal to exclude the active enrollees and enroll the passives, but the ordeal does the opposite.

Columns 3–4 of the table show what happens when we allow for distributional concerns by increasing the social welfare weight μ to 3.0, thus scaling up the social value of enrollee welfare. In this case, it is optimal to cover *both* active and passive enrollees because both their value-cost ratios exceed one. Thus, with $\mu = 3$, we are in the "optimal universality" case discussed in the theory.

Table 3—Targeting Impact of Auto-Enrollment

	Baseline ($\mu = 1.0$)		Higher welfare weight ($\mu = 3.0$)	
Value or cost variable (\$/month)	Active enrollees (1)	Passive enrollees (2)	Active enrollees (3)	Passive enrollees (4)
Social benefits				
WTP of enrollees (demand estimate, W_i)	\$129	\$93	\$386	\$280
Spillovers: Private uncomp. care savings (E_i)	\$88	\$49	\$88	\$49
Total benefits	\$217	\$143	\$474	\$330
Public costs				
Medical spending (gross costs)	\$408	\$228	\$408	\$228
Fiscal externality: Public uncomp. care savings (FE_i)	−\$154	−\$86	−\$154	−\$86
Net public cost (C_i^{Net})	\$255	\$142	\$255	\$142
Value-cost ratio (R_i)	0.85	1.00	1.86	2.32
	(Backward sorting)		(Enrolling both groups optimal)	

Notes: Columns 1 and 2 show our baseline model estimates of the social benefits and costs of insurance for active versus passive enrollees (or inframarginal versus marginal enrollees due to auto-enrollment), while column 3 shows the estimates where enrollee private valuations have been scaled by a social welfare weight of $\mu = 3.0$. The sample is our main 2008–2009 new enrollee sample in the below-poverty group, just as in Table 1. See Section VA for the model estimation method. Enrollee value comes from our demand estimates, using the specification with unobserved sorting equal to observed sorting on WTP.

Supplemental Appendix Table A.10 reports a variety of sensitivity analyses on these targeting results, using different estimates of enrollee value and uncompensated care. As already noted, the finding that (private and social) value is lower for passive enrollees is highly robust, holding in every specification. We also generally find that passive enrollees have similar or larger value-cost ratios, though this finding reverses if sorting on WTP is strong enough (this happens under the "perfect unobserved sorting," and "exponential utility" specifications).

Robustness: Varying Social Preferences for Equity.—How do different social preferences for equity change the implications of these targeting findings? Figure 9 examines the net social welfare of different policies for varying values of the social welfare weight on enrollees, μ (on the x-axis). As noted, a higher μ indicates a stronger value for distributional equity, given that enrollees are low income. The graph plots social welfare for three policies: (i) the ordeal, (ii) full enrollment, and (iii) no enrollment. If ordeals were optimal—that is, if there were positive gains from targeting—the value of SW^{Ordeal} (dashed blue) would need to be higher than both $SW^{FullEnroll}$ (solid red) and $SW^{NoEnroll} = 0$ (solid black). However, this is never the case: the ordeal is dominated by full enrollment for $\mu > 1.3$, by no enrollment for $\mu < 1.0$, and by both policies for $\mu \in [1.0, 1.3]$.[64]

[64] Supplemental Appendix Table A.11 reports sensitivity of this analysis across different demand and externality assumptions. Across most specifications, we find that the ordeal is never optimal at any value of μ. The exceptions are (i) with perfect unobserved sorting, where the ordeal is assumed to sort extremely well on unobservables and so is optimal for a wide range of μ, and (ii) with the simulated exponential utility for a narrow range of $\mu \in (0.55, 0.73)$.

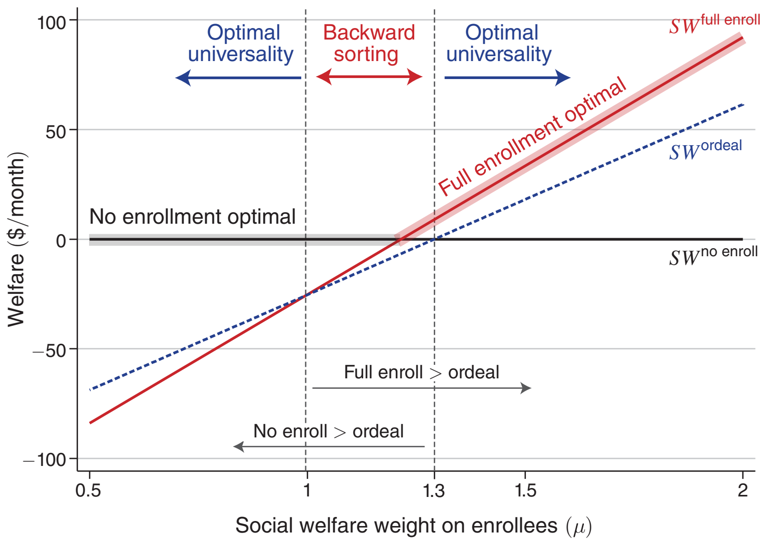

Figure 9. Optimal Policy under Varying Social Values of Equity (μ)

Notes: The figure plots net social welfare of ordeals (blue dashed line) versus full enrollment (red solid) and no enrollment (black solid, which is normalized to zero) under different values for the social welfare weight μ (the x-axis). Social welfare is average net welfare $(= V^{Soc} - C^{Net})$ per eligible person per month. The graph shows that the ordeal is not optimal for any value of μ; it is dominated by no enrollment for lower values $(\mu < 1.3)$ and by full enrollment for higher values $(\mu > 1.0)$ and by both policies for $\mu \in [1.0, 1.3]$, which is the region of backward sorting.

The figure illustrates the reasons why ordeals are nonoptimal, as outlined in Section I. When μ is sufficiently high (above 1.3), the ordeal is undesirable because society wants to cover *both* active and passive enrollees; that is, this illustrates what we called "*optimal universality*." This is likewise true for $\mu < 1.0$, where it is optimal to not enroll both actives and passives. For the small range $\mu \in [1.0, 1.3]$, it would in theory be desirable to exclude the active enrollees, while covering the passives, but ordeals do the opposite. Thus, this case illustrates *backward sorting*.

C. *Policy Comparison: Auto-Enrollment versus Subsidies*

While the main focus of our paper is on the targeting properties of ordeals, we can also use our estimates to compare the trade-offs of two different take-up policies: auto-enrollment versus subsidies. We think of this as a guide for an insurance policymaker who has extra funds and can choose whether to expand coverage via auto-enrollment (for zero-premium enrollees) or larger subsidies (for higher-income groups). This analysis is relevant to understanding trade-offs under the ACA today, in which 40–50 percent of the uninsured likely qualify for free coverage (Cox and McDermott 2020), while many middle-income uninsured Americans owe premiums that could be reduced via larger subsidies. It also reflects (in reverse) Massachusetts's 2010 situation when it chose to eliminate auto-enrollment, rather than cutting subsidies.

Table 4—Policy Comparison: Auto-Enrollment versus Subsidies

	Auto enrollment	Subsidy increase (↓ premiums)		
	0–100% FPL (1)	\$39 to \$0 150% FPL (2)	\$77 to \$39 200% FPL (3)	\$116 to \$77 250% FPL (4)
Panel A. Marginal enrollees				
Enrollment impact	32%	34%	36%	32%
Social benefit $(W_i + E_i)$	\$143	\$62	\$116	\$157
Medical costs	\$228	\$196	\$268	\$281
Gross subsidy (= costs − premiums paid)	\$228	\$196	\$229	\$204
Net public cost (= gross subsidy − FE)	\$142	\$122	\$128	\$98
Value-cost ratio (marginals)	1.00	0.51	0.90	1.60
Panel B. Transfers to inframarginals				
Premium discount (\$/month)	–	\$39	\$38	\$39
× inframarginals per marginal	3.12	2.92	2.80	3.14
= transfer spending per marginal	\$0	\$114	\$106	\$123
Value-cost ratio (inframarginals)	–	1.00	1.00	1.00
Panel C. Cost-effectiveness and MVPF cost-effectiveness				
Net public cost per newly insured	\$142	\$236	\$235	\$221
ΔInsured per \$1 million	7,024	4,238	4,261	4,530
Overall MVPF of policy	1.00	0.74	0.95	1.27

Notes: The table compares auto-enrollment with three subsidy changes generated by premium RDs at three income thresholds: a premium decrease from \$39 to \$0 per month at 150 percent of poverty (FPL) (column 2), from \$77 to \$39 at 200 percent of FPL (column 3), and from \$116 to \$77 at 250 percent of FPL (column 4). For auto-enrollment, results come from our model estimates (Section VA) using the reduced-form variation studied in this paper. For subsidies, estimates come from our calculations using the WTP and cost results reported in Finkelstein, Hendren, and Shepard 2019. Demand for marginal enrollees is assumed to equal the midpoint of the higher and lower premium amounts, and uncompensated care estimates come from applying our model in Section VA to marginal enrollees' costs. Cash transfers are assumed to have an MVPF of 1.0.

For auto-enrollment, we use our model estimates, as just discussed. For subsidies, we use the results of Finkelstein, Hendren, and Shepard (2019). We consider the three subsidy changes in their analysis: reducing premiums from \$39 per month to \$0 (for enrollees at 150 percent of poverty), from \$77 to \$39 (at 200 percent of poverty), and \$116 to \$77 (at 250 percent of poverty).

This analysis yields two main results, shown in Table 4. First, all four take-up policies involve similar enrollment impacts of +32–36 percent. They also all enroll a similar set of low-cost marginal enrollees, with medical costs of \$196–\$281 per month (well below the market average of \$370). Indeed, after subtracting premiums paid, the "gross subsidy" for marginal enrollees is remarkably similar across policies, ranging from \$196 to \$229. The same is true of the net public cost, after subtracting uncompensated care savings. Overall, this suggests that auto-enrollment and the three subsidy expansions have relatively similar take-up impacts and targeting properties.

Second, however, the two policies differ markedly in their expenditures on inframarginal enrollees. Auto-enrollment spends nothing on inframarginal (active) enrollees, while the subsidies all spend $>$ \$100 per marginal enrollee on transfers (the \$38–\$39 monthly subsidy increase times the ≈ 3 inframarginals per marginal enrollee). As a result, auto-enrollment is a much more *cost-effective* policy for expanding take-up. Auto-enrollment's net public cost per newly insured is 36–40 percent lower than for subsidies. This implies that each \$1 million in public spending covers 55–66 percent more people if used for auto-enrollment rather than subsidies. Therefore, a budget-constrained government wishing to maximize take-up would want to prioritize auto-enrollment over subsidies.

On the other hand, if the government wishes to implement the highest-MVPF policy, the analysis also depends on the relative MVPF of insurance versus cash transfers since subsidies combine the two.[65] Cash transfers have an MVPF of 1 in our model (since we do not include labor supply distortions), while the social value-cost ratio of insurance for marginal enrollees (with $\mu = 1$) ranges from 0.51 to 1.60 for subsidies and is (coincidentally) 1.00 for auto-enrollment. As a result, we find that auto-enrollment's MVPF ($= 1.00$) lies within the range of the three subsidy changes (from 0.74 to 1.24).

VI. Conclusion

Enrollment ordeals are a pervasive and controversial feature of many public programs, especially safety net programs for the poor. There is a longstanding debate and tension between two views. On the one hand, ordeals are barriers to poverty alleviation programs, which may undermine their goal of helping the poor. In this view, ordeals are inherently harmful, and particularly so when they reduce take-up a lot.

On the other hand, the classic economic ideas of Nichols and Zeckhauser (1982) show how ordeals can *target* public assistance toward those who need or value it most, saving money that can be redeployed toward those in greatest need. In this view, ordeals are harmful only if they fail to target well. Because the "self-targeting" case for ordeals relies on revealed preferences, standard critiques have largely focused on *behavioral frictions* as the main reason ordeals may not target well (Bertrand, Mullainathan, and Shafir 2004; Finkelstein and Notowidigdo 2019).

This paper argues that there is another big-picture reason ordeals self-targeting may not work well: *adverse selection*. We start by observing that in many public programs, enrollees vary in not just their *value* of assistance but also their *cost*. In other words, many programs—including but not limited to those providing insurance—share the key feature of "selection markets" that have been widely studied in the economics literature (Einav, Finkelstein, and Mahoney 2021). We then show

[65] MVPFs are calculated as follows. For auto-enrollment, we assume (conservatively) that the ordeal involves no real welfare costs ($L(\sigma) = 0$), so its MVPF is simply the social value-cost ratio of marginal (passive) enrollees, as in Table 3. For subsidies, the MVPF combines the social value of insurance (for the ΔD_S marginal enrollees) plus the value of cash discounts to inframarginals ($= \Delta S$ times D_0 inframarginals), divided by the total fiscal cost, or

$$(16)\qquad MVPF_S = \frac{\overbrace{\Delta D_S \bar{V}_S^{Soc}}^{\text{Insurance for marginals}} + \overbrace{D_0 \Delta S}^{\text{Cash for marginals}}}{\Delta D_S \bar{C}_S^{Net} + D_0 \Delta S} = \underbrace{\kappa_M \times \left(\frac{\bar{V}_S^{Soc} + \bar{E}_S}{\bar{C}_S^{Net}}\right)}_{\text{MVPF of marginals}} + \underbrace{(1 - \kappa_M) \times 1}_{\text{Transfer to inframarginals}},$$

where $\bar{X}_S$ is the average of each variable X for subsidy-marginals and $\kappa_M \equiv \Delta D_S \bar{C}_S^{Net} / \left(\Delta D_S \bar{C}_S^{Net} + D_0 \Delta S\right)$ is the share of extra spending on marginal enrollees. The equation shows that the MVPF of a subsidy is a weighted average of the MVPF of covering marginal enrollees and the MVPF of a cash transfer to inframarginals (which is 1.0).

that adverse selection tends to undermine the classic self-targeting logic for ordeals. When low-value types—those whom ordeals are designed to screen out—also have low costs (e.g., because they are lower-risk types), targeting gains from excluding them may be minimal or even negative. The key question in selection markets is not whether ordeals screen on value but whether they screen *more strongly* on value than on costs.

We develop a general framework to formalize this idea, visualized using the graphical selection markets model of Einav, Finkelstein, and Cullen (2010) and measured using a parameter we call the "adverse selection tax." We then test it empirically using a natural experiment in a subsidized health insurance program in Massachusetts. We find that eliminating auto-enrollment and adding a small ordeal leads to major 33 percent declines in enrollment. Ordeals differentially exclude precisely the young, healthy, and low-risk types one would expect under adverse selection. These individuals have lower value for insurance (consistent with self-targeting), but they are also much lower-cost. Our model estimates suggest that they are not less efficient, implying that ordeals induced "backward sorting" into insurance, analogous to the findings of Marone and Sabety (2022) for price-based sorting. This occurs because adverse selection is very strong, with a "tax" exceeding 100 percent in our baseline estimates. With distributional equity concerns, health insurance is socially optimal, but it is optimal for all enrollees, including passive types screened out by ordeals, consistent with our idea of "optimal universality."

These findings have broader implications for how policymakers think about enrollment ordeals in social programs. In terms of *take-up* impact, our results suggest that ordeals are a first-order important barrier in health insurance. Even when coverage is free, a large share of people do not enroll when doing so is a hassle. Completely removing ordeals via auto-enrollment has an order of magnitude larger take-up impact than lower-touch "nudges" like reminders and outreach (Domurat, Menashe, and Yin 2021; Goldin, Lurie, and McCubbin 2021; Ericson et al. 2023; Banerjee et al. 2021). Reaching universal coverage in the United States, therefore, may require automatic enrollment in some form.

In terms of *targeting*, our results suggest that the standard case for ordeals is less likely to work well in settings with adverse selection, that is, strongly correlated value and costs. This is clearly relevant for insurance programs, but it may also be relevant more broadly in transfer programs that pay varying benefit amounts to different groups. Fundamentally, adverse selection (like behavioral biases) interrupts the revealed preference link between demand and efficiency that is key to self-targeting. While ordeals are useful tools in some settings, they may not be well suited to health insurance and other adversely selected markets.

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.

Abaluck, Jason, and Jonathan Gruber. 2023. "When Less Is More: Improving Choices in Health Insurance Markets." *Review of Economic Studies* 90 (3): 1011–40.

Aizawa, Naoki, and You Suk Kim. 2020. "Public and Private Provision of Information in Market-Based Public Programs: Evidence from Advertising in Health Insurance Marketplaces." NBER Working Paper 27695.

Aizer, Anna. 2007. "Public Health Insurance, Program Take-Up, and Child Health." *Review of Economics and Statistics* 89 (3): 400–415.

Alatas, Vivi, Ririn Purnamasari, Matthew Wai-Poi, Abhijit Banerjee, Benjamin A. Olken, and Rema Hanna. 2016. "Self-Targeting: Evidence from a Field Experiment in Indonesia." *Journal of Political Economy* 124 (2): 371–427.

Allcott, Hunt, Daniel Cohen, William Morrison, and Dmitry Taubinsky. 2022. "When Do 'Nudges' Increase Welfare?" NBER Working Paper 30740.

Allcott, Hunt, Benjamin B. Lockwood, and Dmitry Taubinsky. 2019. "Regressive Sin Taxes, with an Application to the Optimal Soda Tax." *Quarterly Journal of Economics* 134 (3): 1557–1626.

Arbogast, Iris, Anna Chorniy, and Janet Currie. 2022. "Administrative Burdens and Child Medicaid Enrollments." NBER Working Paper 30580.

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.

Bernheim, B. Douglas, and Dmitry Taubinsky. 2018. "Behavioral Public Economics." In *Handbook of Behavioral Economics*, edited by B. Douglas Bernheim, Stefano DellaVigna, and David Laibson, 381–516. Elsevier.

Bertrand, Marianne, Sendhil Mullainathan, and Eldar Shafir. 2004. "A Behavioral-Economics View of Poverty." *American Economic Review* 94 (2): 419–23.

Besley, Timothy, and Stephen Coate. 1992. "Workfare versus Welfare: Incentive Arguments for Work Requirements in Poverty-Alleviation Programs." *American Economic Review* 82 (1): 249–61.

Bhargava, Saurabh, George Loewenstein, and Justin Sydnor. 2017. "Choose to Lose: Health Plan Choices from a Menu with Dominated Option." *Quarterly Journal of Economics* 132 (3): 1319–72.

Bhargava, Saurabh, and Dayanand Manoli. 2015. "Psychological Frictions and the Incomplete Take-Up of Social Benefits: Evidence from an IRS Field Experiment." *American Economic Review* 105 (11): 3489–3529.

Bronnenberg, Bart J., Jean-Pierre Dubé, Matthew Gentzkow, and Jesse M. Shapiro. 2015. "Do Pharmacists Buy Bayer? Informed Shoppers and the Brand Premium." *Quarterly Journal of Economics* 130 (4): 1669–1726.

Brot-Goldberg, Zarek C., Samantha Burn, Timothy Layton, and Boris Vabson. 2023. "Rationing Medicine through Bureaucracy: Authorization Restrictions in Medicare." NBER Working Paper 30878.

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): 2718–58.

Brown, David W., Amanda E. Kowalski, and Ithai Z. Lurie. 2020. "Long-Term Impacts of Childhood Medicaid Expansions on Outcomes in Adulthood." *Review of Economic Studies* 87 (2): 792–821.

Bundorf, M. Kate, Jonathan Levin, and Neale Mahoney. 2012. "Pricing and Welfare in Health Plan Choice." *American Economic Review* 102 (7): 3214–48.

Chandra, Amitabh, Jonathan Gruber, and Robin McKnight. 2011. "The Importance of the Individual Mandate—Evidence from Massachusetts." *New England Journal of Medicine* 364 (4): 293–95.

Chandra, Amitabh, Jonathan Gruber, and Robin McKnight. 2014. "The Impact of Patient Cost-Sharing on Low-Income Populations: Evidence from Massachusetts." *Journal of Health Economics* 33: 57–66.

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." *Quarterly Journal of Economics* 129 (3): 1141–1219.

Choukhmane, Taha. 2021. "Default Options and Retirement Saving Dynamics." Unpublished.

Cox, Cynthia, and Daniel McDermott. 2020. *Millions of Uninsured Americans Are Eligible for Free ACA Health Insurance*. Kaiser Family Foundation.

Currie, Janet. 2006. "The Take-Up of Social Benefits." In *Public Policy and the Distribution of Income*, edited by Alan J. Auerbach, David Card, John M. Quigley, 80–148. Russell Sage Foundation.

Dague, Laura. 2014. "The Effect of Medicaid Premiums on Enrollment: A Regression Discontinuity Approach." *Journal of Health Economics* 37: 1–12.

Deshpande, Manasi, and Yue Li. 2019. "Who Is Screened Out? Application Costs and the Targeting of Disability Programs." *American Economic Journal: Economic Policy* 11 (4): 213–48.

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.

Drake, Coleman, David Anderson, Sih-Ting Cai, and Daniel W. Sacks. 2023. "Financial Transaction Costs Reduce Benefit Take-Up: Evidence from Zero-Premium Health Insurance Plans in Colorado." *Journal of Health Economics* 89: 102752.

Dupas, Pascaline, Vivian Hoffmann, Michael Kremer, and Alix Peterson Zwane. 2016. "Targeting Health Subsidies through a Nonprice Mechanism: A Randomized Controlled Trial in Kenya." *Science* 353 (6302): 889–95.

Einav, Liran, and Amy Finkelstein. 2011. "Selection in Insurance Markets: Theory and Empirics in Pictures." *Journal of Economic Perspectives* 25 (1): 115–38.

Einav, Liran, Amy Finkelstein, and Mark R. Cullen. 2010. "Estimating Welfare in Insurance Markets Using Variation in Prices." *Quarterly Journal of Economics* 125 (3): 877–921.

Einav, Liran, Amy Finkelstein, and Neale Mahoney. 2021. "The IO of Selection Markets." In *Handbook of Industrial Organization*, edited by Kate Ho, Ali Hortaçsu, and Alessandro Lizzeri, 389–426. Elsevier.

Ericson, Keith M. Marzilli. 2014. "Consumer Inertia and Firm Pricing in the Medicare Part D Prescription Drug Insurance Exchange." *American Economic Journal: Economic Policy* 6 (1): 38–64.

Ericson, Keith M. Marzilli. 2020. "When Consumers Do Not Make an Active Decision: Dynamic Default Rules and Their Equilibrium Effects." *Games and Economic Behavior* 124: 369–85.

Ericson, Keith Marzilli, Timothy J. Layton, Adrianna McIntyre, and Adam Sacarny. 2023. "Reducing Administrative Barriers Increases Take-Up of Subsidized Health Insurance Coverage: Evidence from a Field Experiment." NBER Working Paper 30885.

Fang, Hanming, Michael P. Keane, and Dan Silverman. 2008. "Sources of Advantageous Selection: Evidence from the Medigap Insurance Market." *Journal of Political Economy* 116 (2): 303–50.

Finkelstein, Amy, Nathaniel Hendren, and Erzo F.P. Luttmer. 2019. "The Value of Medicaid: Interpreting Results from the Oregon Health Insurance Experiment." *Journal of Political Economy* 127 (6): 2836–74.

Finkelstein, Amy, Nathaniel Hendren, and Mark Shepard. 2019. "Subsidizing Health Insurance for Low-Income Adults: Evidence from Massachusetts." *American Economic Review* 109 (4): 1530–67.

Finkelstein, Amy, Neale Mahoney, and Matthew J. Notowidigdo. 2018. "What Does (Formal) Health Insurance Do, and for Whom?" *Annual Review of Economics* 10: 261–86.

Finkelstein, Amy, and Kathleen McGarry. 2006. "Multiple Dimensions of Private Information: Evidence from the Long-Term Care Insurance Market." *American Economic Review* 96 (4): 938–58.

Finkelstein, Amy, and Matthew J. Notowidigdo. 2019. "Take-Up and Targeting: Experimental Evidence from SNAP." *Quarterly Journal of Economics* 134 (3): 1505–56.

Frean, Molly, Jonathan Gruber, and Benjamin D. Sommers. 2017. "Premium Subsidies, the Mandate, and Medicaid Expansion: Coverage Effects of the Affordable Care Act." *Journal of Health Economics* 53: 72–86.

Garthwaite, Craig, Tal Gross, and Matthew J. Notowidigdo. 2018. "Hospitals as Insurers of Last Resort." *American Economic Journal: Applied Economics* 10 (1): 1–39.

Geruso, Michael, Timothy J. Layton, and Jacob Wallace. 2020. "Are All Managed Care Plans Created Equal? Evidence from Random Plan Assignment in Medicaid." NBER Working Paper 27762.

Goldin, Jacob, Ithai Z. Lurie, and Janet McCubbin. 2021. "Health Insurance and Mortality: Experimental Evidence from Taxpayer Outreach." *Quarterly Journal of Economics* 136 (1): 1–49.

Google Maps API. 2013. "Driving Distances from Massachusetts Zip Codes to Healthcare Facilities." Derived from *Google Maps API* and shared with authors by Prof. Amanda Starc (accessed 2013).

Gruber, Jonathan. 2008. "Covering the Uninsured in the United States." *Journal of Economic Literature* 46 (3): 571–606.

Handel, Benjamin R. 2013. "Adverse Selection and Inertia in Health Insurance Markets: When Nudging Hurts." *American Economic Review* 103 (7): 2643–82.

Handel, Benjamin R., and Jonathan T. Kolstad. 2015. "Health Insurance for 'Humans': Information Frictions, Plan Choice, and Consumer Welfare." *American Economic Review* 105 (8): 2449–2500.

Hendren, Nathaniel. 2016. "The Policy Elasticity." *Tax Policy and the Economy* 30 (1): 51–89.

Herd, Pamela, and Donald P. Moynihan. 2018. *Administrative Burden: Policymaking by Other Means*. Russell Sage Foundation.

Homonoff, Tatiana, and Jason Somerville. 2021. "Program Recertification Costs: Evidence from SNAP." *American Economic Journal: Economic Policy* 13 (4): 271–98.

Iyengar, Sheena S., and Emir Kamenica. 2010. "Choice Proliferation, Simplicity Seeking, and Asset Allocation." *Journal of Public Economics* 94 (7-8): 530–39.

Jácome, Elisa. 2020. "Mental Health and Criminal Involvement: Evidence from Losing Medicaid Eligibility." Unpublished.

Jaffe, Sonia, and Mark Shepard. 2020. "Price-Linked Subsidies and Imperfect Competition in Health Insurance." *American Economic Journal: Economic Policy* 12 (3): 279–311.

Kleven, Henrik Jacobsen, and Wojciech Kopczuk. 2011. "Transfer Program Complexity and the Take-Up of Social Benefits." *American Economic Journal: Economic Policy* 3 (1): 54–90.

Lieber, Ethan M. J., and Lee M. Lockwood. 2019. "Targeting with In-Kind Transfers: Evidence from Medicaid Home Care." *American Economic Review* 109 (4): 1461–85.

Lurie, Ithai Z., Daniel W. Sacks, and Bradley Heim. 2019. "Does the Individual Mandate Affect Insurance Coverage? Evidence from Tax Returns." *American Economic Journal: Economic Policy* 13 (2): 378–407.

Madrian, Brigitte C., and Dennis F. Shea. 2001. "The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior." *Quarterly Journal of Economics* 116 (4): 1149–87.

Marone, Victoria R., and Adrienne Sabety. 2022. "When Should There Be Vertical Choice in Health Insurance Markets?" *American Economic Review* 112 (1): 304–42.

Massachusetts Center for Health Information and Analysis (CHIA). 2014. "Massachusetts All-Payer Claims Database." Accessed via DUA with the Massachusetts CHIA Agency.

Massachusetts Health Connector. 2014. "Commonwealth Care Program: Administrative Enrollment and Health Insurance Claims Records." Accessed via DUA with the Massachusetts Health Connector.

McIntyre, Adrianna, Mark Shepard, and Timothy J. Layton. 2024 "Small Marketplace Premiums Pose Financial and Administrative Burdens: Evidence from Massachusetts, 2016–17." *Health Affairs* 43 (1): 80–90.

McIntyre, Adrianna, Mark Shepard, and Myles Wagner. 2021. "Can Automatic Retention Improve Health Insurance Market Outcomes?" *AEA Papers and Proceedings* 111: 560–66.

Mullainathan, Sendhil, and Eldar Shafir. 2013. *Scarcity: Why Having Too Little Means So Much*. Times Books.

Myerson, Rebecca, Nicholas Tilipman, Andrew Fehrer, Hongli Li, Wesley Yin, and Isaac Menashe. 2021. "The Impact of Personalized Telephone Outreach on Health Insurance Take-Up: Evidence from a Randomized Controlled Trial." Unpublished.

Nichols, Albert L., and Richard J. Zeckhauser. 1982. "Targeting Transfers through Restrictions on Recipients." *American Economic Review* 72 (2): 372–77.

Polyakova, Maria. 2016. "Regulation of Insurance with Adverse Selection and Switching Costs: Evidence from Medicare Part D." *American Economic Journal: Applied Economics* 8 (3): 165–95.

Robert Graham Center. 2005–2009. *Social Deprivation Index (SDI)*. Washington, DC: Robert Graham Center. https://www.graham-center.org/maps-data-tools/social-deprivation-index.html (accessed November 2021).

Ruggles, Steven, Katie Genadek, Ronald Goeken, Josiah Grover, and Matthew Sobek. 2015. *Integrated Public Use Microdata Series: Version 6.0 Machine-readable database*. Minneapolis: University of Minnesota.

Shepard, Mark. 2022. "Hospital Network Competition and Adverse Selection: Evidence from the Massachusetts Health Insurance Exchange." *American Economic Review* 112 (2): 578–615.

Shepard, Mark, and Ethan Forsgren. 2023. "Do Insurers Respond to Active Purchasing? Evidence from the Massachusetts Health Insurance Exchange." *Journal of Risk and Insurance* 90 (1): 9–31.

Shepard, Mark, and Myles Wagner. 2025. *Data and Code for: "Do Ordeals Work for Selection Markets? Evidence from Health Insurance Auto-Enrollment."* Nashville, TN: American Economic Association; distributed by Inter-university Consortium for Political and Social Research, Ann Arbor, MI. https://doi.org/10.3886/E211741V1.

Smith, Vernon K., Kathleen Gifford, Eileen Ellis, Robin Rudowitz, Laura Snyder, and Elizabeth Hinton. 2015. *Medicaid Reforms to Expand Coverage, Control Costs and Improve Care: Results from a 50-State Medicaid Budget Survey for State Fiscal Years 2015 and 2016*. Kaiser Family Foundation.

Spinnewijn, Johannes. 2015. "Unemployed but Optimistic: Optimal Insurance Design with Biased Beliefs." *Journal of the European Economic Association* 13 (1): 130–67.

Spinnewijn, Johannes. 2017. "Heterogeneity, Demand for Insurance, and Adverse Selection." *American Economic Journal: Economic Policy* 9 (1): 308–43.

Tebaldi, Pietro. Forthcoming. "Estimating Equilibrium in Health Insurance Exchanges: Price Competition and Subsidy Design under the ACA." *Review of Economic Studies*.

Thaler, Richard H. 2018. "From Cashews to Nudges: The Evolution of Behavioral Economics." *American Economic Review* 108 (6): 1265–87.

Thaler, Richard H., and Cass R. Sunstein. 2008. *Nudge: Improving Decisions about Health, Wealth, and Happiness*. Yale University Press.

Tolbert, Jennifer, Sammy Cervantes, Clea Bell, and Anthony Damico. 2024. *Key Facts about the Uninsured Population*. Kaiser Family Foundation.

Wu, Derek, and Bruce D. Meyer. 2023. "Certification and Recertification in Welfare Programs: What Happens When Automation Goes Wrong?" NBER Working Paper 31437.