Defaults, Mandates, and Taxes: Policy Design with Active and Passive Decision-Makers

Jacob Goldin and Nicholas Lawson*

June 1, 2016

Abstract

Growing evidence suggests that many people are surprisingly responsive to unconventional policy tools, such as defaults or choice-framing, yet unresponsive to conventional ones, such as taxes or subsidies. This paper studies the optimal choice of policy instrument in settings characterized by such features. We utilize a simple binary-choice model in which decision-makers are either active or passive; active choosers make their decisions by comparing perceived costs and benefits whereas passive choosers select whichever option is the default. From this simple model, a number of results emerge. First, manipulating the default option is preferable to imposing a mandate when active choosers tend to make correct decisions. Second, taxes and defaults are complements, not substitutes; employing the two types of instruments in conjunction can yield better results than utilizing either one alone. Finally, the optimal combination of taxes and defaults is typically preferable to a mandate even in settings where active choosers are prone to biases. The results establish important limits on the range of settings in which mandates are an efficient policy response to decision-maker errors.

*Goldin: Stanford Law School, email: jsgoldin@law.stanford.edu. Lawson: Aix-Marseille University (Aix-Marseille School of Economics), CNRS & EHESS, email: nicholas.lawson@univ-amu.fr. For helpful comments, we are grateful to Ian Ayres, Ryan Bubb, Christine Jolls, Louis Kaplow, Wojciech Kopczuk, Daniel Reck, Dmitry Taubinsky, and Abraham Wickelgren, along with two anonymous referees.

Behavioral economics has upended conventional wisdoms about which policy tools are effective at shaping behavior. Growing evidence suggests that price instruments, the policy lever traditionally favored by economists, tend to be overlooked or ignored by many decision-makers, muting their effectiveness at influencing people's choices (Chetty, Looney and Kroft, 2009; Feldman and Ruffle, 2015). On the other hand, recent studies suggest a surprisingly large role for unconventional policy instruments, such as defaults or choice-framing, which shape behavior without altering the monetary incentives that individuals face (Madrian and Shea, 2001; Homonoff, 2014). This paper aims to incorporate both of these insights into a single framework that can inform policymakers' choice of instrument in a range of settings.

To study these issues, we consider a simple model in which individuals make a binary decision between two available options. There are two types of decision-makers in society: active choosers and passive choosers. Active choosers make their decisions in the way that standard models assume, by comparing the perceived costs and benefits of the available options. In contrast, passive choosers simply pick whichever option is the default. For example, in the case of saving for retirement, active choosers would decide how much to save by taking income tax incentives into account (e.g., by forming an IRA or claiming the Saver's Credit) whereas passive choosers would simply follow the default savings plan offered by their employer. Though stark, this simple model captures the empirical findings that the behavior of some decision-makers is sensitive to the default (e.g., Madrian and Shea, 2001; Beshears et al., 2009) as well as the fact that other decision-makers do account for traditional price instruments like taxes (e.g., Chetty et al., 2014), albeit in a potentially biased manner. We assume that neither group's behavior generates externalities, and the government's only goal in setting policy is to make individuals as well off as possible. From these straightforward assumptions, several important results emerge.

We begin by considering the policy choice between defaults and mandates. This question has received substantial attention in the recent behavioral law and economics literature. Prominent academics from a range of disciplines have endorsed the use of defaults and similar nudges to solve all sorts of policy problems (Camerer et al., 2003; Thaler and Sunstein, 2008). On the other hand, detractors have noted that many of the arguments cited in favor of nudges actually support the imposition of mandates, once the arguments are followed through to their logical conclusion (Bubb and Pildes, 2014).[1]

Our first result provides an important baseline for evaluating this debate. The policy choice of mandate versus default only affects the decisions of the active choosers; the passive choosers end up with the same outcome under each policy. It therefore follows that in settings in which the active choosers make unbiased decisions, imposing a mandate can only make people worse off (relative to adjusting the default). For a mandate to be the most efficient policy, a necessary assumption is that the active choosers – those individuals who consistently select the same option regardless of the default – are systematically making incorrect decisions. That is, those individuals who select the non-default option must actually be better off on average with the default option.

Next, we turn away from mandates and focus on policy design when both conventional instruments (taxes and subsidies) and unconventional instruments (defaults) are available. The previous literature has focused on the choice between these two instrument types, exploring how policymakers should trade off between the costs and benefits of each. For example, Galle (2014) compares the economic efficiency of conventional price instruments with nudges, focusing on factors such as revenue, distributional effects, and behavioral distortions. He develops a set of general principles for choosing whether to employ a nudge *or* a price instrument, depending on the situation at hand. Similarly, Mullainathan, Schwartzstein and Congdon (2012) separately derive the optimal tax and the optimal nudge in a model with error-prone decision-makers, but do not consider how the two instruments can be used in conjunction (i.e., when the policy space includes both taxes and nudges).[2] In contrast to the previous literature, we argue that the two types of instruments should be thought of as complements rather than substitutes. Whereas defaults shape the behavior

[1]To be sure, neither side in this debate argues that one type of policy instrument is always better than the other. Sunstein (2013) notes that cases may exist in which mandates are optimal. Similarly, Bubb and Pildes (2014) don't advocate the reflexive use of mandates, they just call for them to be included in the policy analysis. The current project can thus be viewed as setting out a framework to implement the methodological approach called for by Bubb and Pildes (2014), in which the relative advantages of various policy instruments are compared to one another. What emerges from this analysis, however, is that the traditional skepticism towards mandates that Bubb and Pildes attribute to behavioral economics can often be justified on welfarist grounds.

[2]Farhi and Gabaix (2015) consider optimal policy with taxes and nudges, but their model treats nudges as a form of non-monetary psychological tax that directly affects decision-makers' welfare. While helpful for modeling many types of nudges, such as cigarette warning labels, this approach is less well-suited for modeling default effects, which may affect behavior through psychological biases rather than imposing welfare-relevant costs on decision-makers who select the non-default option. In our model, defaults affect decision-making without directly affecting utility. See Bernheim, Fradkin and Popov (2011) for a discussion of these issues.

of passive choosers, policymakers can utilize taxes to shape the behavior of active choosers, who make decisions based on the perceived costs and benefits of the available options. Using the two instruments in conjunction in binary choice settings allows policymakers to achieve higher social welfare by targeting the behavior of both types of decision-makers.

Finally, having established the benefits of employing taxes and defaults in conjunction, we return to the question of when mandates will be efficient. When policymakers can utilize both taxes and defaults, the case for mandates is even weaker than before. Whereas the previous results highlighted that mandates could be more efficient than defaults when active choosers make systematically suboptimal decisions, introducing taxes into the analysis implies that, for a mandate to be more efficient, the suboptimality of the active choosers' decision-making must persist as the size of the tax approaches infinity – i.e., as it becomes equivalent to a mandate. In particular, it must be the case that (on average) the individuals who would continue to choose an option as the tax on that option grows arbitrarily large would nonetheless be better off with the other option. In turn, this requires that active choosers' biases be strongly negatively correlated with their preferences, in a sense we make precise below. Outside of this limiting case, the combination of a default and a tax dominates a mandate; policymakers may utilize the tax to offset the active choosers' bias and ensure that most of the people selecting the non-default option are actually made better off by being allowed to do so. Thus even when mandating one of the two options would be preferable to adjusting the default (on its own), a better policy in such binary-choice settings would be to adjust the default while also imposing a tax.[3]

The remainder of the paper proceeds as follows. Section I describes the model and notation. Section II considers the policy choice between defaults and mandates and derives conditions under which each is desirable. Section III introduces taxation into the model and derives the optimal combination of taxes and defaults. Section IV revisits the question of when mandates are desirable given the availability of both defaults and taxes. Section V concludes.

[3]Outside of binary choice settings, this stark result need not hold. For example, suppose a retirement savings rate of 10% is roughly optimal for everyone with small variations due to individual circumstances. The tax necessary to get those with high biases close to 10% would likely push those with low biases well below 10%. In such cases, it might be preferable to simply mandate a 10% savings rate and sacrifice the small optimal adjustments for individual heterogeneity. This does not happen in our model because individuals are restricted to binary choices.

I. Model

Society is composed of individuals $i \in \mathbb{I}$, where $\mathbb{I}$ has measure one. Individuals make a binary decision from a fixed menu of alternatives, $X = \{x, y\}$, and have preferences over the available options represented by $u_i(x)$ and $u_i(y)$.[4] Let $c_i \in X$ denote the option chosen by individual i. The baseline choice environment in which agents make their decisions consists of the menu (X) and a default, $d \in \{d_x, d_y\}$, where d_x denotes that the default is x and d_y denotes that the default is y.[5]

There are two types of individuals in society, Active choosers (A) and Passive choosers (P). Let $\mathbb{A}$ and $\mathbb{P}$ denote the sets of agents in these two groups, so that $\mathbb{A}$ and $\mathbb{P}$ partition $\mathbb{I}$. To avoid degenerate results, we assume that both $\mathbb{A}$ and $\mathbb{P}$ are non-empty. Passive choosers make their decisions according to which option is the default: $i \in \mathbb{P} \implies \{c_i = y \iff d = d_y\}$. In contrast, active choosers select the option they perceive to be most consistent with their preferences, $i \in \mathbb{A} \implies \{c_i = y \iff u_i(y) - u_i(x) + b_i \geq 0\}$, where $b_i \in \mathbb{R}$ represents a bias in choosing between the available options; a positive value of b_i indicates that an individual is biased towards choosing y. The presence of b in the decision-rule for active choosers accommodates the possibility that even individuals who are not susceptible to default-effects may still exhibit systematic decision-making biases.[6] Note that individuals may differ in the direction in which they are biased. Although simple, this approach to modeling biased decision-making is sufficiently general to accommodate a range of psychological theories (Mullainathan, Schwartzstein and Congdon, 2012).[7] As a result,

[4] By "preferences" we mean the extent to which the available options further the decision-maker's welfare. Preferences are not *defined* by what decision-makers choose; this accommodates the possibility that decision-makers make mistakes. See Basu (2003) for a discussion of this issue. The distinction between preferences and choice corresponds to what some have referred to as "experienced utility" and "behavioral utility."

[5] More generally, one can view d as characterizing the manner in which a choice is "framed" to the decision-maker, with d_x denoting the frame that "nudges" the decision-maker towards x and d_y the frame that "nudges" the decision-maker towards y. The defining feature of a nudge in this framework is that d is not relevant from the perspective of individuals' preferences but nonetheless affects their choices.

[6] Note that for binary choices, a decision-maker whose choice is sensitive to the default is necessarily insensitive to other factors such as prices and taxes. Alternatively, the default could affect the degree to which active decision-makers are biased, but if the effect on their bias is large enough to change which option they select, those decision-makers would, by definition, be passive rather than active. The Online Appendix considers an extension in which the default is allowed to affect the extent to which active choosers are biased, and demonstrates that the results remain consistent with the intuition presented in the main body of the paper.

[7] In addition, our approach does not require specifying the exact positive model that drives sensitivity to default effects – e.g., as between a status quo bias or an anchoring effect. The main restriction on positive models imposed by this approach to modeling default effects is that the default does not directly enter into decision-makers' welfare – e.g., by imposing normatively relevant transaction costs on decision-makers who opt out of the default option. When choosing against the default does directly affect decision-makers' welfare, boundedly-rational decision-makers

we can draw conclusions about optimal policy without specifying the exact positive model that generates agents' behavior.

Combining the behavioral assumptions in the previous paragraph, we have:

$$c_i = \begin{cases} y & i \in \mathbb{P},\ d = d_y \\ x & i \in \mathbb{P},\ d = d_x \\ y & i \in \mathbb{A},\ u_i(y) - u_i(x) + b_i \geq 0 \\ x & i \in \mathbb{A},\ u_i(y) - u_i(x) + b_i < 0 \end{cases} \tag{1}$$

The government may be able to add a set of policies to the baseline choice environment, including altering the default, mandating that a particular option be chosen, or imposing a tax on a particular option. Denote the set of policies available to the government as Π, and the agent's choice under a particular policy $\pi \in \Pi$ as $c_i(\pi)$. We assume that utility U_i is quasi-linear in wealth, so that $U_i = u_i(c_i) + z_i$, where z_i denotes individual i's wealth. We normalize $\int_i z_i = 0$ so that it drops out of the social welfare function.[8] The government's objective is to maximize aggregate social welfare,

$$W = \int_i u_i(c_i).$$

An assumption implicit in this social welfare function is that the process by which a decision-maker achieves an outcome is not welfare relevant. That is, an individual's welfare depends only on which option she ends up selecting; it does not matter whether that option was or was not the default, or whether it was mandated.[9]

may decide to passively follow the default if and only if the perceived benefits to doing so outweigh the associated costs. We consider an extension along such lines in the Online Appendix.

[8] The assumption of quasi-linear utility allows us to abstract from distributional considerations associated with, for example, the imposition of a tax. When utility is not quasi-linear, the policy choice between taxes, defaults, and mandates will have distributional implications. See Galle (2014) for a discussion of these issues.

[9] Generalizing the model to account for these other possibilities would be a useful extension for future research. The extension to the case in which social welfare is given by a weighted average of individual utilities is straightforward, and so is ignored here for expositional simplicity.

II. Defaults vs. Mandates

This section investigates the relative efficiency of defaults and mandates by comparing the social welfare attained under each. We assume that the government has four policy options from which to choose. First, it may "default" individuals into x, which in this model means setting the default to be d_x. Similarly, the government may default agents into y by setting the default as d_y.[10] Alternatively, the government may impose a mandate, requiring all agents to choose one option or the other. Let m_x denote a mandate of x and m_y denote a mandate of y. We assume that the mandate effectively eliminates the non-mandated option from agents' choice set, so that $c_i(m_x) = x$ and $c_i(m_y) = y \; \forall i$. The set of policy instruments from which the government may choose is thus given by $\Pi = \{d_x, d_y, m_x, m_y\}$.

Proposition 1 *Let $v_i = u_i(y) - u_i(x)$ denote the welfare advantage to individual i of ending up with option y rather than option x. A mandate generates higher social welfare than a default in the following cases:*

1. $W(m_x) \geq W(d_x) \iff E[v_i \,|\, v_i \geq -b_i, \; i \in \mathbb{A}] \leq 0$ *, and*

2. $W(m_y) \geq W(d_y) \iff E[v_i \,|\, v_i < -b_i \;, i \in \mathbb{A}] \geq 0$

Proof: The proof of Proposition 1, and all other results, is contained in the Appendix.

Corollary 1 *Suppose $b_i = 0 \; \forall i \in \mathbb{A}$. Then $W(d_k) \geq W(m_k)$ for $k = x, y$.*

Proposition 1 highlights that the policy choice between defaults and mandates comes down to the behavior of the active choosers – the individuals who make the same decisions regardless of which option is the default. Intuitively, passive choosers end up with the same option under both a mandate and a default. As a result, the welfare difference between the two policies hinges on the outcomes realized by those who are insensitive to the default: the active choosers. In the special case in which all individuals in this group are unbiased, the active choosers will select

[10] We do not consider the possibility that the government can entirely remove the default, forcing individuals to make an active choice. Although such policies can be desirable in certain contexts, they are not possible in many others and may be costly to implement (both administratively and from the perspective of decision-makers themselves). See, e.g., Sunstein (2014*a*).

the option that is most consistent with their preferences. By interfering with their ability to make such choices, a mandate reduces social welfare by frustrating the (unbiased) choices of this group.[11] More generally, in settings where some active choosers are biased, Proposition 1 shows that mandating x will be more efficient than defaulting people into x when the group of active choosers selecting y would (on average) be better off under x. That is, not only must the active choosers be systematically biased, that bias must be sufficiently large that the individuals who actively choose y are actually worse off for being allowed to do so.[12]

To highlight the factors that shape the expectation in Proposition 1, it is helpful to assume a constant degree of bias, $b_i = b$, and to rewrite the conditions in integral form:

$$W(m_x) \geq W(d_x) \iff E[v_i \,|\, v_i > -b,\ i \in \mathbb{A}] = \int_{-b}^{\infty} v\, df_A(v) \leq 0 \tag{2}$$

and:

$$W(m_y) \geq W(d_y) \iff E[v_i \,|\, v_i < -b\ , i \in \mathbb{A}] = \int_{-\infty}^{-b} v\, df_A(v) \geq 0 \tag{3}$$

where $f_A(.)$ denotes the probability density function for v across active choosers. The larger the bias towards y, i.e., the greater is b, the more negative the integral in (2) will be. Intuitively, when active choosers are heavily biased towards y, a higher fraction of those who select y will be doing so incorrectly. Conversely, when b is small, the only active choosers who incorrectly select y will have low-intensity preferences, and as a result, will not be much better off when forced by a mandate to choose x. Finally, when $f_A(.)$ has a thick right tail, relatively more active choosers will have strong preferences for y, bolstering the case for nudging towards x rather than mandating x.

When agents are heterogeneous in their bias b_i, the desirability of a mandate also depends on the empirical correlation between b and v in the population of active choosers. When the two are

[11]Thus Proposition 1 reveals that employing a mandate is not only less *liberal* than a default (in the sense of restricting choice), there is also an important sense in which it is more *paternalistic*. That is, in order to argue for a mandate, one has to be willing to assume that even those decision-makers who consistently choose the same option, regardless of the frame, are nonetheless "getting it wrong." In contrast, one can design and implement the optimal default while assuming that the only people making mistakes are those whose choices fail to reveal coherent preference information – i.e., the people whose choices vary according to which option is labeled as the default.

[12]Proposition 1 adds precision to a condition whose general nature has previously been recognized. For example, Bubb and Pildes (2014) write that "the core issue in choosing between a mandatory retirement savings plan and an opt-out one is whether the individuals who would opt out would be making good decisions." Similarly, Sunstein (2014*b*) acknowledges that when "people opt out of 401(k) plans for reasons that are bad ... the argument for a mandate gains force on welfare grounds."

positively correlated, individuals tend to be biased towards the option they prefer. Consequently, when b and v are positively correlated, fewer of those who select x under d_y would actually be better off choosing y, lessening the case for a mandate. Conversely, when b and v are negatively correlated among the active choosers, the case for a mandate will be stronger, because those who choose x are more likely to be better off choosing y and vice-versa.

The results thus far suggest a limited role for mandates relative to defaults when designing policy. The next sections show that when taxes are available as an additional policy instrument, the range of applications in which mandates can be justified on efficiency grounds is even narrower.

III. Defaults and Taxes

This section considers optimal policy design when the government has access to both conventional price instruments (i.e., taxes) as well as unconventional ones (i.e., defaults). We show that under our assumptions the optimal policy involves combining the two types of instrument. Additionally, the task of implementing the optimal level of each instrument is made simpler by the separability of the problem.

Suppose that the government cannot mandate a particular option, but that in addition to setting the default, it may tax decision-makers based on which option they choose. Let $\tau \in \mathbb{R}$ denote the size of the tax on y, which may be positive or negative. The policy space is thus given by $\Pi = \{d_x, d_y\} \times \mathbb{R}$, and a particular policy $\pi \in \Pi$ consists of a default and tax combination, (d, τ).

To incorporate taxes into the analysis, we must specify how they affect the behavior of each type of decision-maker. As before, we assume that the behavior of passive choosers is fully determined by the default; for example, individuals who simply choose according to the default may not account for factors that affect the relative costs and benefits of the available options. In contrast, taxes will shape the behavior of active choosers, who choose according to an option's perceived costs and benefits. With utility quasi-linear over wealth, we have $U_i = u_i(c_i) + z_i - \tau \times 1(c_i = y)$, and thus active choosers will select y if and only if $u_i(y) - u_i(x) + b_i - \tau \geq 0$.[13] Summarizing these

[13]This approach is in the spirit of Mullainathan, Schwartzstein and Congdon (2012) and Allcott, Mullainathan

assumptions, we have:

$$c_i = \begin{cases} y & i \in \mathbb{P},\ d = d_y \\ x & i \in \mathbb{P},\ d = d_x \\ y & i \in \mathbb{A},\ u_i(y) - u_i(x) + b_i - \tau \geq 0 \\ x & i \in \mathbb{A},\ u_i(y) - u_i(x) + b_i - \tau < 0 \end{cases} \tag{4}$$

As before, social welfare is given by the unweighted sum of utilities, $W(\pi) = \int_i U_i$. Assuming that tax revenue is distributed lump-sum to individuals in society, and continuing to normalize the sum of pre-tax wealth to zero we can write $W(\pi) = \int_i u_i(c_i) = E[u_i(c_i)]$. As above, let v_i denote the utility difference between the two options, $v_i = u_i(y) - u_i(x)$, and let $f_b(.)$ denote the pdf of v among those active choosers with bias equal to b (that is, $f_b(v) \equiv f_{v|b}(v|b, i \in \mathbb{A})$, to simplify notation).

Proposition 2 *The optimal default depends only on the preferences of the passive choosers. The optimal tax depends only on the preferences and biases of the active choosers. In particular:*

(2.1) *The optimal default is equal to d_x if and only if $E[u_i(x) \,|\, i \in \mathbb{P}] > E[u_i(y) \,|\, i \in \mathbb{P}]$*

(2.2) *If tax τ^* is optimal, then it satisfies $\tau^* = \frac{E[b\, f_b(\tau^* - b) \,|\, i \in \mathbb{A}]}{E[f_b(\tau^* - b) \,|\, i \in \mathbb{A}]}$*

Proposition 2 provides formulas for the optimal default and optimal tax. Although versions of (2.1) and (2.2) have been previously derived in separate settings in which only defaults or only taxes are available (see Goldin and Reck (2014) for defaults and Allcott and Taubinsky (2014) for taxes) the present result highlights that the problem of determining the optimal tax and default is separable in an important way: the optimal tax depends only on the preferences and behavior of the active choosers, whereas the optimal default depends only on those of the passive choosers. This result reflects the basic principle that the optimal design of a policy instrument depends on the characteristics of the agents who will be affected by it.

The separability of Proposition 2 has important practical implications. It could be that many decision-makers in the population are biased with respect to some decision, but if that bias is

and Taubinsky (2014). Note that non-tax cost differences between x and y are built into $u_i(x)$ and $u_i(y)$.

mostly concentrated among the passive choosers, there won't be a strong case for imposing a corrective "internality" tax. That is, if all of the biased decision-makers simply follow the default, imposing a corrective price instrument would have no beneficial effect. Conversely, the extent to which passive decision-makers are biased doesn't affect how the government should set the default – all that matters for answering that question is which option the passive choosers prefer. This point is frequently lost in the literature, such as when commentators cite decision-making biases as justifications for nudges. In the Online Appendix, we presents two extensions to our model in which this separability is lessened – in particular, when defaults affect the biases of the active choosers or the determination of whether a given decision-maker is active or passive. In such cases we show that although the optimal default and optimal tax are interdependent, the basic point remains that each instrument is set to maximize the welfare of the (endogenously determined) group of decision-makers affected by it.

Next, consider the solution for the optimal tax. First suppose that the bias among active choosers is uniform, so that $b_i = b$. In this case the optimal tax formula reduces to $\tau^* = b$ – the optimal tax is to exactly offset the bias, so that decision-makers select a particular option if and only if they actually prefer to do so.[14] As a result, social welfare is just as high under the optimal tax as if no one had been biased to begin with.

The story grows more complicated when active choosers vary in the extent to which they are biased. Let g_b denote the density of active choosers with bias b and preferences such that they are marginal between x and y at the optimal tax, $g_b = f_b(\tau^* - b)$. For example, at a high b, the only active choosers who are marginal are those with low values of v. Using this notation, we can rewrite the optimal tax formula as $\tau^* = \frac{E[b\, g_b]}{E[g_b]}$, where the optimal tax is simply a weighted average of the active choosers' biases, and the weight on a value of b depends on the fraction of active choosers who are marginal at that level of b. The optimal tax, τ^*, thus equals the average bias for those active choosers who are on the brink between choosing x and y. Consequently, the optimal tax doesn't depend much on levels of b for which few active choosers are marginal (i.e., because their values of v are either too high or too low) because relatively few of these decision-makers will

[14]The fact that the optimal tax in response to a uniform bias is to fully offset that bias is well-recognized in the literature (e.g., Mullainathan, Schwartzstein and Congdon, 2012).

behave differently as a result of a slight increase or decrease in the tax rate.[15]

Thus, the optimal combined tax and default takes an intuitive form: passive choosers should be defaulted into the option that maximizes their average utility, while active choosers should be subjected to a tax equal to the average bias among those who are marginal between x and y. Each instrument is aimed at a separate portion of the population, and they are therefore complementary, each solving its own portion of the behavioral market failure that would exist absent government intervention.

IV. Comparing Mandates to Defaults and Taxes

This section utilizes the optimal policy results from the prior section to determine conditions under which mandates are the best way to maximize social welfare, given that defaults and taxes are also available policy instruments. The policy space in this section consists of mandates as well as any combination of defaults and taxes: $\Pi = \{d_x, d_y\} \times \mathbb{R} \cup \{m_x, m_y\}$. Although a mandate and an infinitely large tax are equivalent in this model, we restrict our focus to finite taxes, since any real-world tax would have to be finite.

The following proposition establishes an important limit on the range of settings in which mandates are an efficient policy response to decision-maker errors. When only defaults are available, mandates can be desirable because they provide a way of targeting the behavior of the active choosers, who might otherwise choose suboptimally under the default. Conversely, if taxes were the only instrument that was available, mandates might be a better option because they provide a way of shaping the behavior of the passive choosers, who wouldn't respond to price incentives. But when both taxes and defaults are available, policymakers can target the behavior of both active and passive choosers in a more fine-tuned way than simply requiring everyone to choose the same option.

Proposition 3

[15] Equation (2.2) expresses a necessary condition for the optimal tax, which may not be unique in certain cases. Additionally, if v is bounded, it is possible that (2.2) could be degenerate if nobody is indifferent at a particular tax rate; in that case, the optimal tax may well mimic a mandate.

(3.1) *For m_x to be optimal, it must be the case that*

(i) $E[u_i(x) \,|\, i \in \mathbb{A}] \geq E[u_i(y) \,|\, i \in \mathbb{A}]$,

(ii) $E[u_i(x) \,|\, i \in \mathbb{P}] \geq E[u_i(y) \,|\, i \in \mathbb{P}]$, *and*

(iii) $E_b \left[\int_{v=\tau-b}^{\infty} v\, f_b(v)\, dv \,|\, i \in \mathbb{A} \right] < 0$ *for all finite* $\tau \in \mathbb{R}$.

(3.2) *For m_y to be optimal, it must be the case that:*

(i) $E[u_i(y) \,|\, i \in \mathbb{A}] \geq E[u_i(x) \,|\, i \in \mathbb{A}]$

(ii) $E[u_i(y) \,|\, i \in \mathbb{P}] \geq E[u_i(x) \,|\, i \in \mathbb{P}]$, *and*

(iii) $E_b \left[\int_{-\infty}^{\tau-b} v\, f_b(v)\, dv \,|\, i \in \mathbb{A} \right] > 0$ *for all finite* $\tau \in \mathbb{R}$.

Corollary *No mandate is optimal when the support of b is bounded.*

Proposition 3 provides three necessary conditions for a mandate to be optimal; to simplify the discussion consider the conditions for a mandate of x in 3.1. First, the active choosers must (on average) prefer x; otherwise a default of x combined with a very large tax on x could achieve higher welfare. In that case, the passive choosers would be equally well off (they would continue selecting x) and the active choosers would be better off because they would motivated by the tax to select y (which, by assumption, they prefer on average). Second, the passive choosers must (on average) prefer x; otherwise a mandate would be worse than a default of y combined with a large tax on y. Finally, since the default/tax combination allows the active choosers to opt out, the average utility gain from doing so (i.e., the average v for those opting out from x) must be negative for all finite taxes. In that case the optimal tax must be infinite, which in our model means that the optimal policy takes the form of a mandate.

Note that when the bias is bounded, the optimal tax must be finite. In that case, the government can simultaneously implement the optimal default for passive choosers and the optimal tax for active choosers, and achieve higher welfare than any mandate.[16] The story grows more complicated

[16]Of course, when the optimal tax is so large that virtually no one continues to choose the non-default option, then the effects of the tax and default combination will be similar to the effects of the mandate. In such cases it is likely that factors outside the model, such as administrative costs, will determine which policy better promotes

when the distribution of b is unbounded. In that case, it is possible that no finite tax will be optimal, if some active choosers continue to select the taxed option regardless of the size of the tax but would actually have been better off selecting the untaxed option (even absent the tax). In that case, a mandate will be optimal, and the following result illustrates one setting in which this can occur.

Illustration *Suppose that among active choosers, b and v are characterized by a bivariate normal distribution,* $\begin{bmatrix} v \\ b \end{bmatrix} \sim N\left(\begin{bmatrix} \mu_v \\ \mu_b \end{bmatrix}, \begin{bmatrix} \sigma_v^2 & \rho\sigma_v\sigma_b \\ \rho\sigma_v\sigma_b & \sigma_b^2 \end{bmatrix}\right)$. *If the average preferences of active and passive choosers are in the same direction – that is, $E[u_i(x) \,|\, i \in \mathbb{A}] \geq E[u_i(y) \,|\, i \in \mathbb{A}]$ and $E[u_i(x) \,|\, i \in \mathbb{P}] \geq E[u_i(y) \,|\, i \in \mathbb{P}]$, or vice-versa – then a mandate will be optimal if and only if the correlation between b and v is sufficiently negative: $\rho < -\frac{\sigma_v}{\sigma_b}$.*

The Illustration considers a special case of the model in which the joint distribution of b and v is described by a bivariate normal distribution. For the mandate to be optimal, the correlation between b and v must be strongly negative. To understand why, consider Proposition 3.1; for m_x to be optimal, it must be the case that the decision-makers who are marginal at arbitrarily high values of the tax (i.e., those with high $b + v$), must on average prefer x to y (i.e., $v < 0$). For v to tend to be negative when $b + v$ is large and positive requires that negative v be associated with large and positive b; in other words, a strongly negative correlation.

An alternative intuition for the case when a mandate is optimal can be provided by considering what it means for the bias b to be infinite. Such an individual is, by assumption, active with respect to the default in that they always choose the same option regardless of which default is selected. However, they are passive with respect to the tax: an individual with an infinitely positive b always chooses y at any finite level of the tax. Thus, Proposition 3 leaves open the possibility that a mandate could be superior to a default/tax combination in the case in which there is a significant fraction of the population whose behavior cannot be affected by either a default or any finite tax.

social welfare. One case in which the optimal tax functions exactly the same as a mandate is when the distribution of v is bounded, and the optimal tax, though finite, is sufficiently large that it induces all active decision-makers to choose the untaxed option.

Proposition 3 is a natural extension of Proposition 1, which showed that a mandate of x is optimal only when those individuals who choose y are on average wrong to do so. When a tax is also allowed, this condition must hold even as the tax approaches infinity, so that individuals with an infinitely strong propensity to choose y are nonetheless wrong to do so. Thus unless active choosers' biases are strongly negatively correlated with their preferences, a mandate will not be the optimal policy.

V. Conclusion

This article has investigated the range of settings in which findings from behavioral economics support the use of mandates. Drawing on a simple but plausible model of decision-making, we argue that the answer depends critically on which other policy instruments are available as alternatives. We find that in settings in which decision-makers are choosing between two alternatives, the combination of a default and a tax is superior to a mandate except in a limiting case in which those individuals who continue to choose a particular option are wrong, on average, to do so even as the tax on that option goes to infinity.

Even when taxes are unavailable, the assumption required to support the use of a mandate – that the active choosers are systematically and strongly biased – is likely to be suspect in many applications; the fact that some decision-makers make suboptimal choices (i.e., by always following the default) does not imply anything about the quality of the choices made by decision-makers who do not behave in this way. To take one example, consider the fact that some purchasers of auto insurance may be sensitive to whether full or limited liability insurance is presented as the default (Johnson et al., 1993). Without more information, there is little reason to believe that those who consistently select limited liability (even when the default is full liability) would actually be better off purchasing full liability.

Given the simplicity of the model and the starkness of the results, a natural question is whether the assumptions relied on by the model are plausible. As described above, the key assumption driving the model's results is that the behavior of all individuals may be influenced by either a default or a tax. This assumption strikes us as being a plausible one in many contexts: a decision-maker who is sufficiently engaged in a particular decision to consistently pick the same option (even when some other option is the default) seems likely to account for the available options' costs and benefits (even if she ends up doing so in a biased manner). Suppose, however, that in contrast to this assumption, some fraction of decision-makers were insensitive to the default but did not choose actively with respect to the tax. For example, such decision-makers might choose an option at random because they do not understand the decision or are unwilling to consider the default. Neither defaults nor taxes would induce random choosers to select the option they most prefer. Hence if there are many individuals who behave this way, a mandate may in fact be the optimal policy.

In addition, our results should not be extrapolated outside of the binary-choice framework we assume here. When choices are non-binary, the analysis grows more complicated; imposing a large tax, for example, though necessary for offsetting the biases of the most biased active choosers, may have the undesirable consequence of pushing less biased decision-makers to reduce their consumption of the taxed good by too great an amount. Similarly, the set of active and passive choosers may vary based on which value of the choice variable is selected as the default; for example, some decision-makers may only follow the default when the default option is close to the option they truly prefer. In contrast, as long as choices are binary, the qualitative results presented here continue to hold even when the determination of whether decision-makers are active or passive is endogenized to depend on the benefit from active choice, as described in the Online Appendix.

Other situations in which mandates may be the most efficient instrument choice are when implementing or enforcing a tax involves higher administrative costs than a mandate, or in certain settings in which decision-makers are prone to behave in non-systematic ways that are hard to correct with taxes (e.g., by making random mistakes). Especially in settings where the optimal tax is so large that it approximates the effects of a mandate, factors such as these might be decisive. Finally, if there is some constraint on the size of a feasible tax - for instance, political feasibility - a mandate might be better than the constrained optimal tax. Thus, our conclusion is not that a mandate can never be the optimal policy. Rather, the results here shed light on the types of conditions that can (and cannot) support the use of a mandate, and bolster the case for employing a default or other form of nudge in the settings where such conditions are unlikely to hold.

References

Allcott, Hunt, and Dmitry Taubinsky. 2014. "The Lightbulb Paradox: Using Behavioral Economics for Policy Evaluation." *Working Paper.*

Allcott, Hunt, Sendhil Mullainathan, and Dmitry Taubinsky. 2014. "Energy Policy with Externalities and Internalities." *Journal of Public Economics*, 112: 72–88.

Basu, Kaushik. 2003. *Prelude to Political Economy.*

Bernheim, B Douglas, Andrey Fradkin, and Igor Popov. 2011. "The welfare economics of default options in 401 (k) plans." National Bureau of Economic Research.

Beshears, John, James J Choi, David Laibson, and Brigitte C Madrian. 2009. "The importance of default options for retirement saving outcomes: Evidence from the United States." In *Social security policy in a changing environment.* 167–195. University of Chicago Press.

Bubb, Ryan, and Richard Pildes. 2014. "How Behavioral Economics Trims Its Sails and Why." *Harvard Law Review*, 127: 1593.

Camerer, Colin, Samuel Issacharoff, George Loewenstein, Ted O'Donoghue, and Matthew Rabin. 2003. "Regulation for Conservatives: Behavioral Economics and the Case for Asymmetric Paternalism." *University of Pennsylvania Law Review*, 151(3): 1211–1254.

Chetty, Raj, Adam Looney, and Kory Kroft. 2009. "Salience and Taxation: Theory and Evidence." *American Economic Review*, 99(4): 1145–1177.

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

Farhi, Emmanuel, and Xavier Gabaix. 2015. "Optimal Taxation with Behavioral Agents." *Working Paper*.

Feldman, Naomi, and Bradley Ruffle. 2015. "The Impact of Including, Adding and Subtracting a Tax on Demand." *American Economic Journal: Economic Policy*, 7(1): 95–118.

Galle, Brian. 2014. "Tax, Command...or Nudge? Evaluating the New Regulation." *Texas Law Review*, 92(4): 837–894.

Goldin, Jacob, and Daniel Reck. 2014. "Preference Identification Under Inconsistent Choice: A Reduced-Form Approach." *Working Paper*.

Homonoff, Tatiana. 2014. "Can Small Incentives Have Large Effects? The Impact of Taxes versus Bonuses on Disposable Bag Use." *Working Paper*.

Johnson, Eric J, John Hershey, Jacqueline Meszaros, and Howard Kunreuther. 1993. "Framing, probability distortions, and insurance decisions." *Journal of Risk and Uncertainty*, 7(1): 35–51.

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): 1149–1187.

Mullainathan, Sendhil, Joshua Schwartzstein, and William Congdon. 2012. "A Reduced-Form Approach to Behavioral Public Finance." *Annual Review of Economics*, 4: 511–540.

Sunstein, Cass. 2013. "The Storrs Lectures: Behavioral Economics and Paternalism." *Yale Law Journal*, 122: 1826.

Sunstein, Cass. 2014*a*. "Choosing Not to Choose." *Duke Law Journal*, 64: 1–52.

Sunstein, Cass. 2014*b*. "Nudges vs. Shoves: The Benefits of Preserving Choice." *Harvard Law Review*, 127: F210.

Thaler, Richard H, and Cass R Sunstein. 2008. *Nudge: Improving decisions about health, wealth, and happiness.* Yale University Press.

Appendix: Proofs

Proof of Proposition 1

First, because $\mathbb{I}$ has measure 1, we can write $W(\pi) = E[u_i(c_i(\pi))]$. Applying the law of iterated expectations:

$$E[u_i(c_i(\pi))] = \phi_P \, E[u_i(c_i(\pi)) \,|\, i \in \mathbb{P}] + \phi_A \, E[u_i(c_i(\pi)) \,|\, i \in \mathbb{A}] \tag{5}$$

where ϕ_P and ϕ_A denote the fraction of the population that are (respectively) active and passive: $\phi_P = p(i \in \mathbb{P})$, $\phi_A = 1 - \phi_P$. Applying the same theorem again, to just the active choosers, we can write:

$$\begin{aligned} E[u_i(c_i(\pi)) \,|\, i \in \mathbb{A}] = & \; E[u_i(c_i(\pi)) \,|\, i \in \mathbb{A},\, c_i(d_x) = x] \; p(c_i(d_x) = x \,|\, i \in \mathbb{A}) \\ & + \; E[u_i(c_i(\pi)) \,|\, i \in \mathbb{A},\, c_i(d_x) = y] \; p(c_i(d_x) = y \,|\, i \in \mathbb{A}) \end{aligned} \tag{6}$$

Collecting results, we obtain:

$$\begin{aligned} W(\pi) = & \; \phi_P \, E[u_i(c_i(\pi)) \,|\, i \in \mathbb{P}] + \phi_A \, E[u_i(c_i(\pi)) \,|\, i \in \mathbb{A},\, c_i(d_x) = x] \; p(c_i(d_x) = x \,|\, i \in \mathbb{A}) \\ & + \; \phi_A \, E[u_i(c_i(\pi)) \,|\, i \in \mathbb{A},\, c_i(d_x) = y] \; p(c_i(d_x) = y \,|\, i \in \mathbb{A}) \end{aligned} \tag{7}$$

We can use (7) to evaluate social welfare under defaults and mandates. In particular, we will compare $W(m_x)$ to $W(d_x)$ (the comparison of $W(m_y)$ to $W(d_y)$ is analogous). Under m_x, both active and passive agents select x: $c_i(m_x) = x \; \forall i$. Hence social welfare is given by:

$$\begin{aligned} W(m_x) = & \; \phi_P \, E[u_i(x) \,|\, i \in \mathbb{P}] + \phi_A \, E[u_i(x) \,|\, i \in \mathbb{A},\, c_i(d_x) = x] \; p(c_i(d_x) = x \,|\, i \in \mathbb{A}) \\ & + \phi_A \, E[u_i(x) \,|\, i \in \mathbb{A},\, c_i(d_x) = y] \; p(c_i(d_x) = y \,|\, i \in \mathbb{A}) \end{aligned} \tag{8}$$

Under d_x, passive choosers select x, $c_i(d_x) = x \; \forall i \in \mathbb{P}$. Hence we can write:

$$\begin{aligned} W(d_x) = & \; \phi_P \, E[u_i(x) \,|\, i \in \mathbb{P}] + \phi_A \, E[u_i(x) \,|\, i \in \mathbb{A},\, c_i(d_x) = x] \; p(c_i(d_x) = x \,|\, i \in \mathbb{A}) \\ & + \; \phi_A \, E[u_i(y) \,|\, i \in \mathbb{A},\, c_i(d_x) = y] \; p(c_i(d_x) = y \,|\, i \in \mathbb{A}) \end{aligned} \tag{9}$$

Subtracting (8) from (9) yields the conditions under which the default induces higher social welfare than the mandate:

$$W(d_x) - W(m_x) = \phi_A \, E[u_i(y) - u_i(x) \,|\, i \in \mathbb{A},\, c_i(d_x) = y] \, p(c_i(d_x) = y \,|\, i \in \mathbb{A}) \tag{10}$$

As probabilities, $p(c_i(d_x) = y \,|\, i \in \mathbb{A})$ and ϕ_A are both weakly greater than 0; meanwhile, (1) implies that $c_i(d_x) = y$ for an active chooser if and only if $v_i \geq -b_i$, which means that the expectation term can be rewritten as $E[v_i \,|\, v_i \geq -b_i,\, i \in \mathbb{A}]$, and the first part of the proposition follows immediately. The proof for the second part of the proposition is symmetric.

■

Proof of Corollary 1

If $b_i = 0$ for all active choosers, $E[v_i \,|\, v_i \geq -b_i,\, i \in \mathbb{A}]$ must necessarily be weakly positive, which proves the corollary for mandates or defaults in favor of x; the proof for y is symmetric. ■

Proof of Proposition 2.1[17]

First, note that for an arbitrary (d, τ) combination, the law of iterated expectations allows us to write social welfare as:

$$W(d, \tau) = \phi_P \, E[u_i(c_i(d, \tau)) \,|\, i \in \mathbb{P}] + \phi_A \, E[u_i(c_i(d, \tau)) \,|\, i \in \mathbb{A}] \tag{11}$$

To determine the optimal default, we can use (4) and (11) to write:

$$W(d_x, \tau) = \phi_P \, E[u_i(x) \,|\, i \in \mathbb{P}] + \phi_A \, E[u_i(c_i(d_x, \tau)) \,|\, i \in \mathbb{A}] \tag{12}$$

and similarly:

$$W(d_y, \tau) = \phi_P \, E[u_i(y) \,|\, i \in \mathbb{P}] + \phi_A \, E[u_i(c_i(d_y, \tau)) \,|\, i \in \mathbb{A}] \tag{13}$$

Finally, because active decision-makers choose according to their perceived benefits, rather than the default, we have $i \in \mathbb{A} \implies c_i(d_x, \tau) = c_i(d_y, \tau)$. Using this fact, along with (12) and

[17] This result comes from Goldin and Reck (2014). For clarity, the proof is repeated here using the current notation.

(13), we can write the advantage to d_x over d_y as:

$$W(d_x, \tau) - W(d_y, \tau) = \phi_P \, E[u_i(x) - u_i(y) \,|\, i \in \mathbb{P}] \tag{14}$$

Hence, the optimal default depends upon whether passive choosers, on average, would be better off under x or y, $W(d_x, \tau) > W(d_y, \tau) \iff E[u_i(x) \,|\, i \in \mathbb{P}] > E[u_i(y) \,|\, i \in \mathbb{P}]$. ■

Proof of Proposition 2.2

Note that, among active choosers, i chooses y (i.e., $c_i = y$) if and only if $v_i + b_i - \tau \geq 0$. Social welfare depends on the tax only to the extent that the tax shapes the choices of the active choosers: $W(d, \tau) = \phi_P \, E[u_i(d) \,|\, i \in \mathbb{P}] + \phi_A \, E[u_i(c_i(d, \tau)) \,|\, i \in \mathbb{A}]$, or:

$$W(d, \tau) = \phi_P \, E[u_i(d) \,|\, i \in \mathbb{P}] + \phi_A \, E[u_i(x) \,|\, i \in \mathbb{A}] + \phi_A \, p(c_i = y \,|\, i \in \mathbb{A}) \, E[v_i \,|\, c_i = y, \, i \in \mathbb{A}]$$

Hence choosing τ to maximize social welfare is equivalent to choosing τ to maximize $w(\tau) \equiv p(c_i = y \,|\, i \in \mathbb{A}) \, E[v_i \,|\, c_i = y, \, i \in \mathbb{A}]$.

Define $f_{v,b}$ to be the joint density of v and b among active choosers. We can use this definition along with the fact that $c_i = y \iff v_i + b_i - \tau \geq 0$ to write the maximand $w(\tau)$ as:

$$\begin{aligned} w(\tau) &= p(v + b \geq \tau) \int_{b=-\infty}^{\infty} \int_{v=-\infty}^{\infty} v \, f_{v,b}(v, b \,|\, v + b \geq \tau) \, dvdb \\ &= \int_{b=-\infty}^{\infty} \int_{v=\tau-b}^{\infty} v \, f_{v,b}(v, b) \, dvdb \\ &= E_b \left[\int_{v=\tau-b}^{\infty} v \, f_b(v) \, dv \right] \end{aligned} \tag{15}$$

where $f_b(v)$ is the distribution of v conditional on b, as described in section III, and where for simplicity of notation we omit the implied conditioning on $i \in \mathbb{A}$. We can write useing on $i \in \mathbb{A}$. Differentiating $w(\tau)$ yields $w'(\tau) = E_b \left[-(\tau - b) \, f_b(\tau - b) \,|\, i \in \mathbb{A} \right]$. Setting this first-order condition equal to zero and solving for τ yields the result. ■

Proof of Proposition 3

We present the proof for the necessary conditions for m_x to be optimal; the proof for m_y is exactly analogous if x and y are reversed. Thus, it must be the case that $E[u_i(x) \,|\, i \in \mathbb{A}] \geq E[u_i(y) \,|\, i \in \mathbb{A}]$; if not, m_x can never be optimal, because higher welfare can be achieved by a very large but finite negative tax on y combined with a default in whichever direction maximizes the average utility of passive choosers.[18] Then there are two possible cases: either (i) $E[u_i(x) \,|\, i \in \mathbb{P}] < E[u_i(y) \,|\, i \in \mathbb{P}]$ or (ii) $E[u_i(x) \,|\, i \in \mathbb{P}] \geq E[u_i(y) \,|\, i \in \mathbb{P}]$.

Take case (i) first. Since y is better than x for the average passive agent, the optimal default is d_y, and the welfare advantage of the default/tax combination over m_x is:

$$W(d_y, \tau) - W(m_x) = \phi_P \, \left[E[u_i(y) \,|\, i \in \mathbb{P}] - E[u_i(x) \,|\, i \in \mathbb{P}]\right] + \phi_A \, E_b \left[\int_{v=\tau-b}^{\infty} v \, f_b(v) \, dv \right] \tag{16}$$

The first term, multiplied by the ϕ_P, is positive by assumption, and the second term goes to zero as τ goes to infinity, which implies that some finite tax exists that is large enough that (16) is positive.

Therefore, a mandate of m_x may only be superior to a default/finite tax combination in case (ii), where both active and passive agents are at least weakly better off on average with x. Then the optimal mandate and the optimal default are both in favor of x (in the case of indifference, x is equally as good as y), and the welfare advantage of the default/tax combination is:

$$W(d_x, \tau) - W(m_x) = \phi_A \, E \left[\int_{v=\tau-b}^{\infty} v \, f_b(v) \, dv \right] \tag{17}$$

The mandate is superior only when this is negative for all finite taxes; that is, since we know that (17) converges to zero as τ approaches infinity, the mandate is superior only when the optimal tax is infinite, i.e. when for any finite tax, a larger tax can always improve welfare. ■

[18] The passive choosers can be made at least as well off with a default as with a mandate, and if active choosers prefer y on average, then an infinite negative tax - that is, an infinite subsidy - on y replicates a mandate of y for active choosers. Some finitely large negative tax can then come arbitrarily close to a mandate of y, thus achieving higher average utility than a mandate of x if $E[u_i(y) \,|\, i \in \mathbb{A}] > E[u_i(x) \,|\, i \in \mathbb{A}]$.

Proof of Corollary to Proposition 3

Consider the test for a mandate m_x in (3.1). If the bias b towards y is bounded, there exists some maximum value $\bar{b}$. If τ is set equal to $\bar{b}$, the lower limit of integration in $\int_{v=\tau-b}^{\infty} v\, f_b(v)\, dv$ is greater than or equal to zero for all b, and thus $E_b\left[\int_{v=\tau-b}^{\infty} v\, f_b(v)\, dv\right]$ must be weakly positive at the finite tax $\tau = \bar{b}$. The proof is symmetric for m_y.

Proof of Illustration of Proposition 3

Finally, we prove the sufficient condition given in the Illustration in section IV. We want to find the condition on the parameters that ensures that a mandate is optimal when b and v are distributed according to a bivariate normal distribution among active choosers. That is, assume that:

$$\begin{bmatrix} v \\ b \end{bmatrix} \sim N\left(\begin{bmatrix} \mu_v \\ \mu_b \end{bmatrix}, \begin{bmatrix} \sigma_v^2 & \rho\sigma_v\sigma_b \\ \rho\sigma_v\sigma_b & \sigma_b^2 \end{bmatrix}\right)$$

which implies that:

$$f_{v,b}(v,b) = \frac{1}{2\pi\sigma_b\sigma_v\sqrt{1-\rho^2}} \exp\left(\frac{-1}{2(1-\rho^2)}\left[\frac{(b-\mu_b)^2}{\sigma_b^2} + \frac{(v-\mu_v^2)}{\sigma_v^2} - \frac{2\rho(b-\mu_b)(v-\mu_v)}{\sigma_b\sigma_v}\right]\right) \tag{18}$$

We assume that the average preferences of both active and passive choosers are in favor of the same option; otherwise, as stated in the proof to Proposition 3, a default in favor of the passive choosers' best option and a sufficiently large tax in favor of the active choosers' best option can always do better than a mandate. Next, since setting the tax to maximize welfare is equivalent to maximizing $w(\tau)$ as defined in (15), we will consider the value of $w'(\tau)$:

$$w'(\tau) = \int_{-\infty}^{\infty} -(\tau - b)\, f_{v,b}(\tau - b, b)\, db$$

and using the density defined in (18), this integral can be solved analytically; the problem is very complex, so we used Wolfram Mathematica's symbolic integration capabilities to solve for:

$$w'(\tau) = \left[\frac{\exp\left(\frac{-(\mu_b+\mu_v-\tau)^2}{2(\sigma_b^2+\sigma_v^2+2\rho\sigma_b\sigma_v)}\right)}{\sqrt{2\pi}(\sigma_b^2+\sigma_v^2+2\rho\sigma_b\sigma_v)^{\frac{3}{2}}}\right](\sigma_v(\sigma_v+\rho\sigma_b)(\mu_b-\tau) - \mu_v\sigma_b(\sigma_b+\rho\sigma_v))$$

The set of terms inside the large square brackets are always positive, so the sign of $w'(\tau)$ depends on the final set of terms. Suppose that $\sigma_v + \rho\sigma_b > 0$; then it is clear that when τ is sufficiently negative, $w'(\tau)$ must be positive, and when τ is sufficiently positive, $w'(\tau)$ must be negative. In that case, the welfare derivative changes sign only once, from positive to negative, and thus there must be an interior optimal tax rate.

However, if $\sigma_v + \rho\sigma_b < 0$, the opposite is true: the welfare derivative is initially negative, then switches to become positive, and the optimal tax rate must be infinite. This is the sufficient condition provided in section IV for a mandate to be optimal, and it can be rewritten as $\rho < \frac{-\sigma_v}{\sigma_b}$; the correlation between the bias b and the true utility value v must not only be negative, but strongly negative, so that as $b + v$ goes to infinity, the average v among indifferent individuals must remain negative. Note also that this condition is easier to satisfy when the variance of v is small but the variance of b is large: in that case, there are many people who have a small negative value of v but a very strong bias towards y, so that they will continue to choose y erroneously even as the tax on y gets very large. $\blacksquare$