Theory Notes to Accompany Active Decisions

James Choi, David Laibson, Brigitte Madrian, Andrew Metrick

1 The Model

Our model comprises a firm, which sets a default contribution rate for the company savings plan, and the workers who choose their behavior in response.

Consider a group of workers with quasi-hyperbolic preferences, so that they have discount function $1, \beta\delta, \beta\delta^2, ...$ where $0 < \beta \leq 1$. For simplicity, we assume that $\delta = 1$, eliminating long run discounting. We also assume that each worker has an exogenously determined and constant optimal savings rate, denoted s, that is known to the worker with certainty but not observed by the firm. The firm does know the probability distribution function for s, denoted $f_s(s)$.

In our model, the firm acts in period 0 by setting the default contribution rate, denoted d, at which all new employees are automatically enrolled in the company savings plan. In each period thereafter, every worker has the oppotunity to change her savings rate permanently from the default to her optimal level. Should she act, she incurs a stochastic cost c_t drawn from a uniform distribution on the interval $[\underline{c}, \bar{c}]$, where $\underline{c} < \bar{c}$. This cost is determined at the beginning of the period and thus is known to the worker at the time of her decision. After each period in which the agent saves at the default d, however, she faces a flow cost $L(s, d) \geq 0$; without loss of generality, we assume workers incur this cost at the beginning of the next period.

2 The Sophisticated Worker's Optimization Problem

Given d and c, workers attempt to minimize their current discounted loss function, denoted W. For this section, let c, c' and c'' denote the stochastic cost drawn in the current and two next periods, respectively. Also, s and d

are fixed here and do not matter outside of $L(s,d)$, and so we refer only to L. Thus,

$$W(L,c) = \begin{cases} c & \text{if act} \\ \beta \left[L + E[v(L,c')] \right] & \text{if not act} \end{cases} . \tag{1}$$

The function $v(\cdot)$, in turn, represents the agents' current perception of future loss when, having not acting today, they face the same decision tomorrow. Since agents have no long run discounting,

$$v(L,c') = \begin{cases} c' & \text{if act tomorrow} \\ L + E[v(L,c'')] & \text{if not act tomorrow} \end{cases} . \tag{2}$$

Notice, at this point, that the agent views tradeoffs between the current period and future periods differently than the tradeoff between future periods as a result of the quasi-hyperbolic discounting. Workers suffer from a dynamic inconsistency problem; they want to constrain future selves to act more patiently than they will act. We assume that our workers are sophisticated, however, and thus understand their own self control problems. Unlike naive agents, whose actions we characterize further below, sophisticated actors make current decisions based on rational expectations of future actions.

The equilibrium for this game takes the form of a "cutoff rule." Agents act if and only if the stochastic cost falls at or below some optimal point c_s^*. Since agents must be indifferent in the current period (but not in perceived future periods) between acting or not at the cutoff,

$$c_s^* = \beta \left[L + E[v(L,c')] \right]. \tag{3}$$

In combination with the fixed point condition for the future loss function

$$v(L,c') = \begin{cases} c' & \text{if } c' \leq c_s^* \\ L + E[v(L,c'')] & \text{otherwise} \end{cases} , \tag{4}$$

we solve for c_s^*.

Proposition 1 *The optimal cutoff rule for sophisticated agents can be expressed as*

$$c_s^* = \frac{\underline{c} + \sqrt{\underline{c}^2\left[1-(2-\beta)\beta\right] + 4\beta\left(1-\frac{\beta}{2}\right)(\bar{c}-\underline{c})\,L}}{2-\beta}. \quad (5)$$

Though mathematically complex, this expression is quite intuitive. Importantly, this function is increasing in L, the cost of deviation from a worker's optimal savings rate; the less ideally she is saving, the more willing she is to incur the cost of changing. Furthermore, c_s^* is increasing in β. As agents become less patient, they are less willing to incur c to reduce all future flow costs. Interestingly, while $c_s^* \geq \underline{c}$, it need not lie below $\bar{c}$. Agents can be so far from their optimal savings level that they are willing to incur any cost $c \in [\underline{c}, \bar{c}]$ in the first period in order to avoid future flow costs.

3 The Firm's Optimization Problem

Given the future equilibrium policies of the workers, the firm sets the default contribution rate in period 0. To avoid complication, we assume that this decision is made by a benign regulator acting in the interest of the agents. Thus, the firm sets the default rate to minimize the average loss function for the population of workers. The benevolent planner, however, does not discount quasi-hyperbolically and takes the long run perspective; its decision problem takes the form

$$d^* = \arg\min_d \int l_s(s,d) f_s(s) ds,$$

where

$$l_s(s,d) = \begin{cases} c & \text{if } c \leq c_s^* \\ L(s,d) + E[v(s,d,c')] & \text{otherwise} \end{cases}.$$

This function, though, is precisely the future loss function from the workers' decision problem. For analytical tractability, we now assume $L(s,d) = \kappa(s-d)^2$ as the functional form for the cost of suboptimal saving.[1] At this point, note that s and d always occur in the relation $s-d$ in the mathemetics, and that the function $L(s,d)$

[1] The symmetry of this loss function around the point $s = d$ may be an unrealistic representation of actual savings plans. Firms often match employee contributions to the plan up to a discreet cutoff point. Thus, a fixed deviation from one's optimal rate may be more costly if it falls below the ideal point than above it. This function does include convexity in deviations from the optimum, though, and so perhaps represents a good balance between realism and tractability.

is symmetric around the point $s - d = 0$. For ease of notation, we denote $\Delta = s - d$, so $L(s, d) = L(\Delta) = \kappa\Delta^2$. Plugging this function into the the general form derived in the proof of Proposition 1 yields the follwing result.

Proposition 2 *The expected loss function for workers from the firm's perspective can be expressed as*

$$l_s(\Delta) = \begin{cases} 0 & \text{if } \Delta = 0 \\ \frac{\sqrt{2(\bar{c}-\underline{c})\beta(2-\beta)\kappa\Delta^2+(1-\beta)^2\underline{c}^2}}{\beta(2-\beta)} - \kappa\Delta^2 + \frac{\underline{c}}{\beta(2-\beta)} & \text{if } 0 < |\Delta| < \sqrt{\frac{\bar{c}(2-\beta)-\underline{c}\beta}{2\beta}} \\ \frac{\bar{c}+\underline{c}}{2} & \text{otherwise} \end{cases} . \quad (6)$$

Figure 1 includes three examples of this function for different values of β (ignore the dotted and dashed lines, as well as the shaded area, for the moment). As one varies Δ, l_s has three ranges. Where $\Delta = 0$, the agent suffers no loss since she begins the model at her optimal contribution rate. Note that, in general, there is a discontinuity at $l_s(0)$. Because there is no long run discounting, any worker for whom $s \neq d$ *must* eventually act to avoid $W \to \infty$ from gradually accruing flow costs; when she does, she will incur some stochastic cost c, creating the discontinuity . On each extreme, if a worker is far enough from her optimal savings level, she automatically acts in the first period, and so her expected loss is constant at the average action cost. This case results from the situation discussed above in which $c_s^* \geq \bar{c}$. Finally, there are two intermediate ranges in which the loss depends on $|\Delta|$. Importantly, if $\beta < 1$, this dependence is *non-monotonic*, as displayed in the final two panels in Figure 1. If the default setting is close to a worker's optimal contribution rate, then she is unlikely to act to change the rate but also accrues few flow costs. As $|\Delta|$ increases, though, her cutoff rule, and thus her likelihood of acting, increases more slowly than the flow costs; her dynamic inconsistency problem causes the agent to delay acting even when it is in her long run interest. This creates "humps" in which, from the firm's long run perspective, agents are worse off than if forced to act immediately.

The quasi-hyperbolic discount factor β, through its influence on worker preferences, also has a large effect on the shape of the expected loss function. This impact can be seen easily by comparing the different panels in Figure 1. In the first panel, where $\beta = 1$, there is no "hump"; since agents are exponential discounters, they do not suffer from dynamic inconsistency, and so being closer to one's ideal saving rate is always better. In the second panel, β falls below one but is quite high. There are small "humps" from the quasi-hyperbolic discounting, but agents who are very close to their optimal savings rates are still the best off. In the third panel, however, β is very small and so the dynamic inconsistency problem is very large. This affects the loss function in two related but distinct ways. First, the loss in utility for individual agents, graphically represented by the height of the humps over the region in which workers automatically act, is extremely large. Second, the set of agents whose problems from quasi-hyberbolic discounting is very broad. In fact, β is small enough in this panel so that *any* deviation from a worker's optimal savings rate places her in such a range. In this case, every agent (except the lucky soul for whom $s = d$) is better off if forced to act. Such differences in β, and thus l_s, drastically affect the firm's equilibrium choice of d^*.

The final assumption required to determine the firm's choice of d is the distribution of s; here we assume that s is uniformly distributed on the interval $[\underline{s}, \bar{s}]$. Thus, the firm tries to minimize the expression in equation 6 over $\Delta \in [\underline{s} - d, \bar{s} - d]$, so that

$$d^* = \arg\min_d \int_{\underline{s}-d}^{\bar{s}-d} l_s(\Delta) d\Delta. \quad (7)$$

The firm's optimization problem boils down to finding the limits of intergration so as to minimize the area under a function, such as that in one of the panels from Figure 1, which is now fixed with respect to d. The two key elements in this problem are the shape of function, determined largely by β as discussed above, and the width of the integral bounds, measured by $\bar{s} - \underline{s}$. We now proceed to characterize the different optimal default contribution rate regimes as a function of these two variables in two subsections: First, we describe the different solutions graphically and intuitively. We then provide a more precise mathematical characterization in a theorem and proof heuristic. The full proof appears in the theoretical appendix.

4 Characterizing Optimal Default Policies

The optimal solution for the default savings policy is depicted in Figure 2 as a function of the quasi-hyperbolic discount rate, β, and the heterogeneity in optimal savings rates among the population of workers, $\bar{s} - \underline{s}$. There are three essential regions to this space which correspond to three different solution types. We discuss each in turn in the following paragraphs. The firm is indifferent between policies along the boundaries between the regions. Each of the three panels in Figure 1 graphically represent one of the solutions; in these figures, the solid line represents the expected loss function as a function of Δ, the dashed lines depict the range of optimal savings rates for workers (in terms of Δ), and the shaded region displays the total expected loss for the workers under the optimal default policy.

First, consider the southeast region of the diagram, in which agents have fewer dynamic inconsistency problems and also have more homogeneous optimal savings rates. The optimal default policy here is a "center" solution in which d is set in the middle of the range of possible worker optimal contribution rates. The intuition for this solution is straightforward and is displayed graphically in the first panel of Figure 1. Because agents have a high β, and thus few self control problems, there is no reason to keep agents farther from their own optimal rates than necessary. Furthermore, even as problems from quasi-hyperbolic discounting get worse, and the "humps" discussed above get larger, the range of optimal contribution rates of workers is small enough so that few workers fall in those bad regions. Once self control problems get too bad, however, the center default is no longer optimal, and we must consider another solution concept.

For this second optimal default regime, consider the west region of the figure in which agents exhibit serious dynamic inconsistency in their behavior. Here, the optimal policy is a "active decision" policy, in which d is set so far above or below the range of optimal savings rate such that every agent for whom $s \neq d$ acts in the first period. Because agents discount future periods so severely, if left alone many would set their cutoff rule very low and thus accrue large flow costs while waiting to act. By forcing universal action in the first period, the firm throws away the option value for workers of waiting for a low switching cost but, in effect, saves them from their own tendency to "procrastinate." Note that, for sufficiently low values of β, the "active decision" regime is optimal for *all* values of $\bar{s} - \underline{s}$. Panel 3 in Figure 1 specifically displays this situation. All workers (for whom $\Delta \neq 0$) are happier if forced to act; thus, no matter how small the range of ideal savings rates, an "active decision" policy is best. In the part of the figure where β is a bit higher, though, some workers are better off waiting to act, and thus this default system will only be optimal when agent heterogeneity is sufficiently high to rule out a "center default." Generally, as agents discount the future less, a greater spread among ideal agent rates is needed for the "active decision" policy to be optimal; thus the boundary between these regions, defined by a relation $(\bar{s} - \underline{s})(\beta)$ such that the firm is indifferent between the two methods, is increasing in β.

Finally, there is a third solution type, which we denote "offset," for the region in which both $\bar{s} - \underline{s}$ and β are large. Under this policy, the firm fixes the default within the range of possible worker ideal rates but moved towards one end of this range. Panel 2 in Figure 1 illustrates this optimal default regime. In these situations, the loss function has small “humps" since dynamic inconsistency problems are small but still present, but the range of optimal savings rates is very wide. The offset is essentially a compromise between the “active decision” and “center” solutions. Because agents show little dynamic inconsistency, it is not efficient to force everyone to switch in the first period since many could gain from the option value of waiting. On the other hand, the range of optimal savings rates is large enough that a “center” solution would place too many workers in the pessimal intermediate ground. By offsetting the default rate below the center, for instance, the firm can beneficially force some workers (those who would have fallen in the upper “hump”) to act while still letting others (those near the new default rate) take advantage of the option value of waiting.

The boundaries of this region are also interesting. First, notice in Figure 2 that the line along which the firm is indifferent between the “offset” and “active decision” regions is vertical. In each solution, if one were to slightly increase the heterogeneity of agents, the marginal saver would be forced to act immediately. Thus, one canot effect the relation between the total expected utility of agents under these two plans by changing $\bar{s} - \underline{s}$; only β can effect this tradeoff.

Next, consider the boundary to this region when $\beta = 1$. In this case, the agent has no quasi-hyperbolic discounting, and so the expected loss function is strictly increasing in $|\Delta|$ as in panel 1 in Figure 1. In this case, the “center” solution is *always* a solution. However, suppose $\bar{s} - \underline{s}$ is large enough so that some agents act immediately under the “center” default rate regime; if the firm were to move the default rate within a small region such that the marginal saver is still acting immediately, the total expected loss for agents is uneffected. Thus, the precise solution set along this boundary includes not only the “center” policy but also a neighborhood thereabouts; and while this region does include the limit of the “offset” as $\beta \to 1$, the intuition is more akin to the “center” solution. Unfortunately, the boundary between the “offset” and “center” default regions is less tractable.

Another interesting property of these equilibrium policies is the existence of a “global indifference point” at which all three default systems produce equal total expected loss for workers.

We have now completely characterized the optimal default rate as a function of β and $\bar{s} - \underline{s}$; in the following subsection, we treat the above analysis with somewhat more rigor. If the reader is interested in the mathematical specifics of the optimal default rate policies, this section will be of interest. Otherwise, we suggest the reader skip this section.

4.1 Optimal Default Policies: A Mathematical Characterization

The following theorem describes the optimal default policies for different values of $\bar{s} - \underline{s}$ and β.

Theorem 3 *Suppose agents are sophisticated, and the optimal savings rates are uniformly distributed between $\underline{s}$ and $\bar{s}$. Then the optimal default policies fall into the following three "classes":*

DEFAULT POLICY	CONDITIONS
center (d_c)	$d^* = \frac{\bar{s}+\underline{s}}{2}$
offset (d_o)	$d^* = \bar{s} - \sqrt{\frac{\bar{c}\beta - \underline{c}(2-\beta)}{2(2-\beta)\kappa}}$ *or* $d = \underline{s} + \sqrt{\frac{\bar{c}\beta - \underline{c}(2-\beta)}{2(2-\beta)\kappa}}$
active decision (d_a)	$d^* \leq \underline{s} - \sqrt{\frac{\bar{c}(2-\beta) - \underline{c}\beta}{2\beta\kappa}}$ *or* $d^* \geq \bar{s} + \sqrt{\frac{\bar{c}(2-\beta) - \underline{c}\beta}{2\beta\kappa}}$

Then, let $\beta^* = 1 - \frac{\bar{c}-\underline{c}}{\bar{c}+\underline{c}}$, and define $\widehat{\beta}$ implicitly with the equation

$$\int_{\sqrt{\frac{\bar{c}\beta - \underline{c}(2-\beta)}{2(2-\beta)\kappa}}}^{\sqrt{\frac{\bar{c}(2-\beta) - \underline{c}\beta}{2\beta\kappa}}} \left(\frac{\bar{c} + \underline{c}}{2} - l_s(\Delta) \right) d\Delta = 0.$$

Also, implicitly define $[\bar{s} - \underline{s}]^{a/c}(\beta)$ by the equality

$$\frac{1}{\bar{s} - \underline{s}} \int_{-\frac{\bar{s}-\underline{s}}{2}}^{\frac{\bar{s}-\underline{s}}{2}} \left(\frac{\bar{c} + \underline{c}}{2} - l_s(\Delta) \right) d\Delta = 0$$

and $[\bar{s} - \underline{s}]^{o/c}(\beta)$ with

$$\int_{\frac{\bar{s}-\underline{s}}{2}}^{\sqrt{\frac{\bar{c}(2-\beta) - \underline{c}\beta}{2\beta\kappa}}} \left(l_s(\Delta) - \frac{\bar{c} + \underline{c}}{2} \right) d\Delta - \int_{-\frac{\bar{s}-\underline{s}}{2}}^{-\sqrt{\frac{\bar{c}\beta - \underline{c}(2-\beta)}{2(2-\beta)\kappa}}} \left(l_s(\Delta) - \frac{\bar{c} + \underline{c}}{2} \right) d\Delta = 0.$$

Then the optimal d^* are:

d^*	CONDITIONS
active decision	*when* $\beta \in (0, \beta^*]$ *or* $\beta \in (\beta^*, \hat{\beta})$ *and* $\bar{s} - \underline{s} > (\bar{s} - \underline{s})^{a/c}(\beta)$
offset default	*when* $\beta \in (\hat{\beta}, 1)$ *and* $\bar{s} - \underline{s} > (\bar{s} - \underline{s})^{o/c}(\beta)$
center default	*when* $\beta \in (\beta^*, \hat{\beta}]$ *and* $\bar{s} - \underline{s} < (\bar{s} - \underline{s})^{a/c}(\beta)$ *or* $\beta \in (\hat{\beta}, 1)$ *and* $\bar{s} - \underline{s} < (\bar{s} - \underline{s})^{o/c}(\beta)$

All three indifference curves meet at a global indifference point $(\hat{\beta}, [\bar{s} - \underline{s}]^{o/c}\left(\hat{\beta}\right))$ *in the parameter space. Finally, if* $\beta = 1$ *and* $d^* = \bar{s} - \underline{s} > (\bar{s} - \underline{s})^{o/c}(1)$*, then* $d^* \in [\underline{s} + \sqrt{\frac{\bar{c}-\underline{c}}{2\kappa}}, \bar{s} - \sqrt{\frac{\bar{c}-\underline{c}}{2\kappa}}]$.

The theoretical appendix contains the (quite lengthy) proof of this theorem, but we give the heuristic for the proof here.

The proof begins by establishing basic properties for $l_s(\Delta)$, including symmetry, continuity, and differentiability at all points except the knots. Next, we characterize those regions of $l_s(\Delta)$ which lie above $\frac{\bar{c}+\underline{c}}{2}$ and determine the extent of the "humps" in which agents would prefer being forced to act immediately. This proof the region where $\beta \leq \beta^*$, in which case $l_s(\Delta) \geq \frac{\bar{c}+\underline{c}}{2}$ for all $\Delta \neq 0$. We also characterize the derivatives of l_s and prove that $l_s(\Delta)$ has global maxima at $\Delta = \pm\sqrt{\frac{(\bar{c}-\underline{c})^2 - \underline{c}^2(1-\beta)^2}{2(\bar{c}-\underline{c})\beta(2-\beta)\kappa}}$, as well as a global minimum at $\Delta = 0$. The next step shows that, if $l_s(a) = l_s(b)$, a and b fall on different sides of the same "hump," and $|a| < |b|$, then $|l'_s(a)| < |l'_s(b)|$. In other words, the "humps" are steeper on the outside portion than on the portion closer to $\Delta = 0$. This fact is critical because it states, intuitively, that, starting from the maximum of l_s at the top of each "hump," the expected loss for workers falls faster as they move away from the default rate than as they move towards it. After showing that the relationship between $l_s(\Delta)$ and β is $\frac{\partial l}{\partial \beta} \leq 0$, we complete the preliminary work by solving for the first and second order conditions implied in the firm's maximization problem and identifying the three potential solution policies for the firm's optimization problem.

We then chacterize the optimal solutions. In the range $0 < \beta < \beta^*$, we prove that the "active decision" solution is the only possible solution. Because $l_s(\Delta) > \frac{\bar{c}+\underline{c}}{2}$ for all $\Delta \neq 0$, the proof is trivial. Similarly, in the extreme case where $\beta = 1$, then the center default must always work, by the monotonicity of $l_s(\Delta)$ in $|\Delta|$, as well as those defaults in the immediate neighborhood for large enough $\bar{s} - \underline{s}$. In the large intermediate range, where $\beta^* < \beta < 1$, each solution is optimal in certain ranges of the space $[\bar{s} - \underline{s}] \times \beta$.

The final step of the proof is to choose between the multiple solutions in the range where $\beta^* < \beta < 1$. This part of the proof involves the rigorous determination of the boundaries between the regions; we implicitly define a $\hat{\beta}$ which delineates the vertical boundary between the "active decision" and "offset" policies, as well as functions which mark the other boundaries. The proof concludes with the formalization of several properties of these implicit bounderies. We formalize the intuition given above for the positive monotonicity of the boundary between the "active decision" and "center default" regions, as well as prove the existence of the "global indifference point" where the firm is indifferent between all three solutions.

5 The Case of Naive Workers

Our analysis to this point has assumed that workers are sophisticated and understand their own dynamic inconsistency. The force of our results is not diminished if we assume the opposite, though, and instead model workers as completely unaware of their own self control problems.

While the mathematics behind the optimal default policy for naive workers are different from those for sophisticated agents, the intuition is very much the same. Workers who exhibit quasi-hyberbolic discounting may benefit from a default rate that forces them to act in the first period regardless of the stochastic switching cost. The effects of such a policy should be even greater for naifs than for the sophisticated workers, though, because agents who do not account properly for their future actions are even less likely than sosphiticates to overcome the urge to delay the cost of switching. The following theorem summarizes this result.

Theorem 4 *For a given parameter value, if an "active decision" rule chracterizes the optimal default policy for sophisticated agents, then d^* for naive agents is also an "active decision" policy.*

The intuition for this theorem is a direct result of the lack of foresight in the behavior of naifs. Though both types of agents use a "cutoff rule" for determining whether or not to act in a given period, naifs foolishly believe that their future optimization proceeds according to current valuations of future periods. Consequently, in addition to the cutoff rule for the current period, they choose a second cutoff rule for use in those future periods, denoted $\hat{c}_n^*$, which minimizes future loss from the current perspective. Because of this unrealistic expectation of future optimization, the naive perception of expected future losses is less than that for sophisticates. Believing that she is faced with a rosier tomorrow, the naif is less willing to incur the stochastic action cost in the current period to change her savings rate, and thus her current cutoff c_n^* is lower than for a similar sophisticate. In expectation, from the perspective of the benevolent firm, she is worse off in the long run than the sophisticate. Now, suppose the parameter values are such that an "active decision" rule is optimal for the firm. This implies that enough sophisticates suffer from severe dynamic inconsistency that immediate action is optimal. The naifs, however, are worse off than the sophisticates, so an "active decision" default rule must be optimal in this case as well.

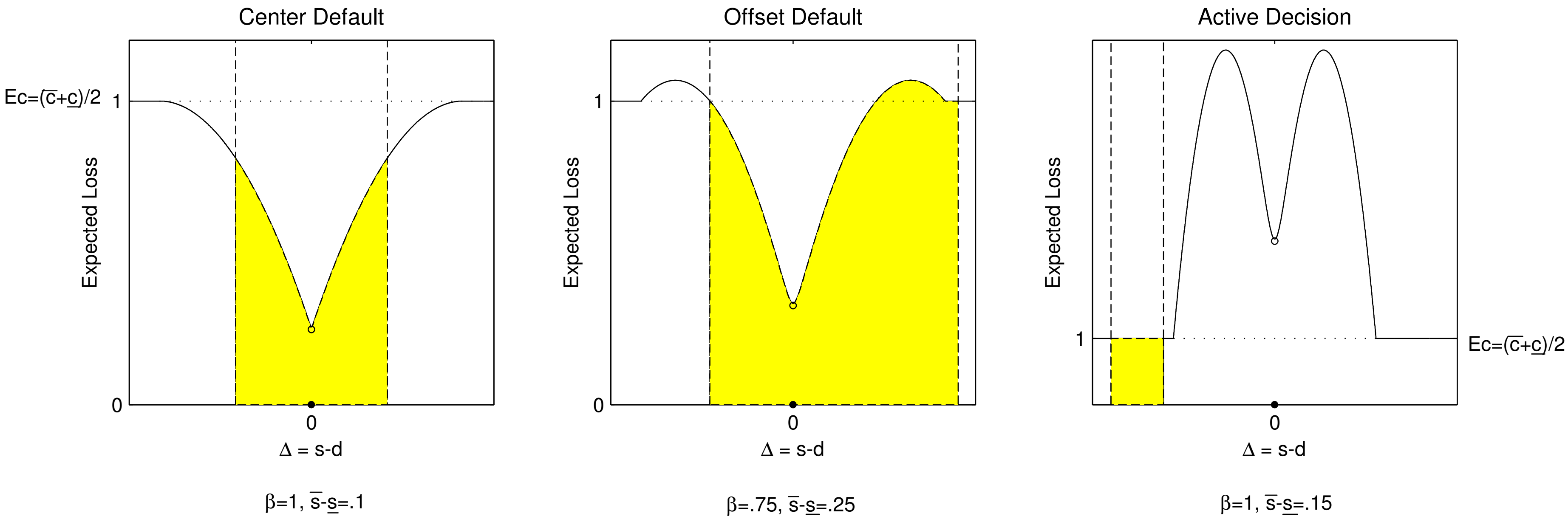

Fig. 1. Possible optimal default regimes. The panels illustrate respective parameter settings that support the three class es of optima studied in this paper: center default, offset default and active decision. The shaded area in each panel represents the social welfare losses generated by the correspoding default regime. The parameter values common to the three panels are the average action cost $(\overline{c}+\underline{c})/2=1$, the range of costs $\overline{c}-\underline{c}=1.5$, and the loss function scaling factor $\kappa=100$. The parameters specific to each panel, i.e., the quasi-hyperbolic discount factor β and the range of optimal saving rates $\overline{s}-\underline{s}$, appear below each figure.

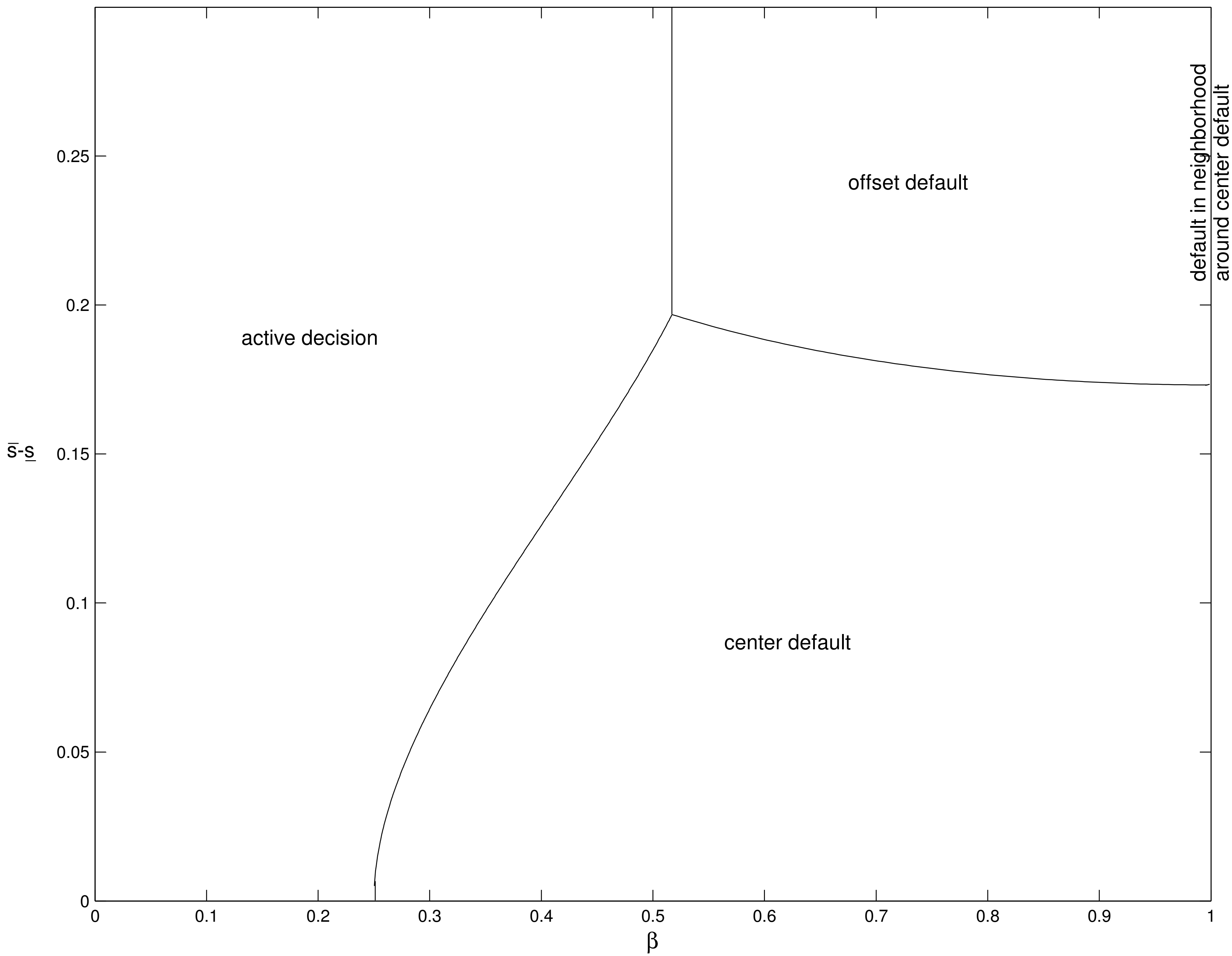

Fig. 2. Characterization of optimal default regimes. This figure characterizes the boundaries of the optimal default regimes as a function of the hyperbolic discount factor β and the range of optimal saving rates $\bar{s}$-$\underline{s}$ for the following set of parameters: average action cost $(\bar{c}+\underline{c})/2=1$, range of costs $\bar{c}-\underline{c}=1.5$, loss function scaling factor $\kappa=100$.