The Welfare Economics of Default Options:

A Theoretical and Empirical Analysis of 401(k) Plans

Appendix

B. Douglas Bernheim, Stanford University and NBER

Andrey Fradkin, Stanford University Igor Popov, Stanford University

October 7, 2011

1 Proofs

Proof of Theorem 1

Let m denote a monetary transfer, and let $X(m)$ and f denote the individual's opportunity set and decision frame, respectively. For any alternative bundle x,[1]

$$EV_A(x) = \inf \{m \mid yP^*x \text{ for all } m' \geq m \text{ and } y \in C(X(m'), f)\}$$

and

$$EV_B(x) = \sup \{m \mid xP^*y \text{ for all } m' \leq m \text{ and } y \in C(X(m'), f)\}$$

First we show that if P_i^* is transitive, then zP_i^*x implies $EV_{Ai}(z) \geq EV_{Ai}(x)$ and $EV_{Bi}(z) \geq EV_{Bi}(x)$. Choose any $\varepsilon > 0$. By definition, yP_i^*z for all $m' \geq EV_{Ai}(z) + \varepsilon$ and $y \in C(X(m'), f)$. Thus, by transitivity, yP_i^*x for all $m' \geq EV_{Ai}(z)+\varepsilon$ and $y \in C(X(m'), f)$, which implies $EV_{Ai}(x) \leq EV_{Ai}(z)$. Similarly, by definition, xP_i^*y for all $m' \leq EV_{Ai}(x)-\varepsilon$ and $y \in C(X(m'), f)$. Thus, by transitivity, zP_i^*y for all $m' \leq EV_{Ai}(x)-\varepsilon$ and $y \in C(X(m'), f)$, which implies $EV_{Bi}(z) \geq EV_{Bi}(x)$.

Next choose any $x' \in X_M$. If x' is a weak generalized Pareto optimum we are done, so suppose it is not. Consider the (necessarily) non-empty set $U = \{y \in X \mid yP_i^*x' \text{ for all } i\}$.

[1] The defnitions given here are special cases of the definitions in Bernheim and Rangel (2009), in that here the alternative to the status quo is a specific bundle x, rather than an alternative opportunity set.

Choose any individual j and consider some z' and f such that $(U, f) \in \mathcal{G}$ and $z' \in C_j(U, f)$.[2] We claim that z' is a weak generalized Pareto optimum in X. If it were not, then there would be some w such that $wP_i^* z'$ for all i. By the transitivity of P_i^*, we would then have $w \in U$, which contradicts $z' \in C_j(U, f)$ (because in particular $wP_j^* z'$). From our first step, we then have $EV_{Ai}(z') \geq EV_{Ai}(x')$ and $EV_{Bi}(z') \geq EV_{Bi}(x')$ for all i, from which it follows that

$$\sum_i (\lambda_{Ai} EV_{Ai}(z') + \lambda_{Bi} EV_{Bi}(z')) \geq \sum_i (\lambda_{Ai} EV_{Ai}(x') + \lambda_{Bi} EV_{Bi}(x'))$$

Consequently, $z' \in X_M$. □

Proof of Theorem 2

Part 1: For period 0, offering a plan with $d = 0$ weakly Pareto improves upon offering no plan.

All employees receive the bundle $(e, x, z) = (0, 0, 1)$ without the plan. With the plan, the bundle $(0, 0, 1)$ remains available. Employees either choose it, in which case they are no worse off, or choose another bundle, in which case they are better off (strictly so for those with $x^*(\theta) > 0$ and γ sufficiently small). Under our assumptions, both sets have positive measure.

Part 2: For period 0, offering a plan with $d > 0$ does not weakly Pareto improve upon not offering a plan.

To prove this claim, we show that workers with $x^*(\theta) = 0$ are strictly worse off with the plan (recall that the set of such workers is assumed to have positive measure). Such workers either opt out to $x = 0$ and receive the bundle $(e', 0, 1)$, or fail to opt out and receive the bundle $(0, d, 1 - \tau(d))$. They strictly prefer $(0, 0, 1)$ to both of these bundles. □

Proof of Theorem 3

Plainly, the bundle obtained by any individual with zero opt-out costs does not depend on d. Thus, the optimal d maximizes aggregate surplus among those for whom opt-out costs

[2] Here we are employing the assumptions, stated in BR, that (i) $C(G)$ is non-empty for all $G \in \mathcal{G}^*$, and (ii) for every set Z there exists a frame f such that $(Z, f) \in \mathcal{G}$.

are positive. Because the solution to that problem is independent of κ (which simply scales the objective function), the default rate that maximizes aggregate EV does not depend on κ.

Next we show that there exists some $\overline{\kappa} \in (0,1)$ such that for $\kappa > \overline{\kappa}$, the opt-out minimizing default rate is either $d = 0$ or $d = \overline{x}$. We have assumed that the (induced) distribution of $x^*(\theta)$ has atoms at the extreme points of the feasible set, 0 and $\overline{x}$. Let $\phi_1 > 0$ denote the fraction of employees with $x^*(\theta) = 0$, and let $\phi_2 > 0$ denote the fraction with $x^*(\theta) = \overline{x}$. The fraction of individuals opting out with $d = 0$ is bounded above by $1 - \phi_1$, and the fraction opting out with $d = \overline{x}$ is bounded above by $1 - \phi_2$. Because the distribution of $x^*(\theta)$ is assumed to have no atoms on the interior of $[0, \overline{x}]$, the fraction opting out of any $d \in [0, \overline{x}]$ is bounded below by κ. Thus, the claim follows for any choice of $\overline{\kappa} > 1 - \max\{\phi_1, \phi_2\}$. $\square$

Proof of Theorem 4

For $x \in [0, \overline{x}]$. Define $\underline{m}(x)$ as the solution to

$$V(0, 1 + \underline{m}(x), \theta) = V(x, 1 - \tau(x), \theta) - \gamma$$

and $\overline{m}(x)$ as the solution to

$$V(0, 1 + \overline{m}(x), \theta) = V(x, 1 - \tau(x), \theta)$$

Under our assumptions, existence, uniqueness, and continuity are guaranteed. Hence $\underline{m}(x)$ has a minimum, $m_L > -1$, and $\overline{m}(x)$ has a maximum, m_H, on $[0, \overline{x}]$. Because V is continuously differentiable and $[m_L, m_H] \times \Theta$ is compact, $V_z(0, 1 + m, \theta)$ has a minimum, $v_L > 0$ (recall that V is strictly increasing in z) and a maximum, v_H, on $[m_L, m_H] \times \Theta$.

Let $\phi(x)$ denote the fraction of individuals for whom $x^*(\theta) = x$. Note that $\phi(x)$ is strictly positive for $x \in \mathcal{A}$ and zero otherwise. Let $\phi^* \equiv \max_{d \in \mathcal{A}} \phi(d)$.

Consider any $d \in \mathcal{A}$, and any individual for whom $x^*(\theta) = d$. In light of the fact that

$$V(0, 1 + m^0(d, \theta), \theta) = V(d, 1 - \tau(d), \theta) \tag{1}$$

and

$$V(0,1+m^1(\theta,\gamma),\theta) = V(x^*(\theta),1-\tau\left(x^*(\theta)\right),\theta)-\gamma, \tag{2}$$

for those individuals we have

$$V(0,1+m^0(d,\theta),\theta) - V(0,1+m^1(\theta,\gamma),\theta) = \gamma.$$

It follows that

$$\left[m^0(d,\theta) - m^1(\theta,\gamma)\right] v_H \geq \gamma.$$

Consequently,

$$\int_{D(d)} \left[m^0(d,\theta) - m^1(\theta,\gamma)\right] dH^\theta(\theta) dH_k^\gamma(\gamma) \geq \frac{\phi(d)\gamma_k}{v_H}. \tag{3}$$

Now consider any $d \notin \mathcal{A}$. From equations (1) and (2), we see that, for all $(\gamma,\theta) \in D(d)$,

$$V(0,1+m^0(d,\theta),\theta) - V(0,1+m^1(\theta,\gamma),\theta) \leq \gamma$$

(where we have used the fact that $V(x^*(\theta),1-\tau\left(x^*(\theta)\right),\theta) \geq V(d,1-\tau\left(d\right),\theta)$). It follows that

$$\left[m^0(d,\theta) - m^1(\theta,\gamma)\right] v_L \leq \gamma.$$

Consequently,

$$\int_{D(d_k)} \left[m^0(d,\theta) - m^1(\theta,\gamma)\right] dH^\theta(\theta) dH_k^\gamma(\gamma) \leq \frac{\overline{\gamma}_k}{v_L} \int_{D(d_k,\overline{\gamma}_k)} dH^\theta(\theta). \tag{4}$$

where $D(d,\gamma) \subset \Theta$ denotes the opt-in set for a fixed value of γ.

Now suppose the theorem is false. Then there is some sequence H_k^γ with $\overline{\gamma}_k \to 0$ and $\gamma_k/\overline{\gamma}_k > e^* > 0$, and an associated sequence of optimal defaults $d_k \notin \mathcal{A}$ with $d_k \to d^* \notin \mathcal{A}$. Plainly, from (3) and (4),we must have, for all k,

$$\int_{D(d_k,\overline{\gamma}_k)} dH^\theta(\theta) \geq \frac{v_L}{v_H}\phi^* e^* > 0.$$

Accordingly, we will introduce a contradiction by demonstrating that $\int_{D(d_k,\overline{\gamma}_k)} dH^\theta(\theta) \to 0$.

We claim that, if $d_k \to d^*$, then for all $\varepsilon > 0$ there exists K^ε such that for $k > K^\varepsilon$ all those with ideal points outside $(d^* - \varepsilon, d^* + \varepsilon)$ opt out. We prove this claim in four steps.

Step 1: With a default of $d^* - \frac{\varepsilon}{2}$, there exists K_L^ε such that for $k > K_L^\varepsilon$, all workers for whom $x^*(\theta) \le d^* - \varepsilon$ opt out.

Because $x^*(\theta)$ is continuous and Θ compact, we know that $\{\theta \mid x^*(\theta) \le d^* - \varepsilon\}$ is compact. Thus, we can define

$$\vartheta_L = \max_{\theta \in \{\theta' \mid x^*(\theta') \le d^* - \varepsilon\}} \left[V(x^*(\theta), 1 - \tau\left(x^*(\theta)\right), \theta) - V(d^* - \frac{\varepsilon}{2}, 1 - \tau\left(d^* - \frac{\varepsilon}{2}\right), \theta) \right].$$

Furthermore, because $x^*(\theta)$ is unique, we necessarily have $\vartheta_L > 0$ (otherwise we would have $x^*(\theta) = d^* - \frac{\varepsilon}{2}$ for some $\theta \in \{\theta' \mid x^*(\theta') \le d^* - \varepsilon\}$). Step 1 then follows from the fact that there exists K_L^ε such that $\overline{\gamma}_k < \vartheta_L$ for all $k > K_L^\varepsilon$.

Step 2: With a default of $d^* + \frac{\varepsilon}{2}$, there exists K_H^ε such that for $k > K_H^\varepsilon$, all workers for whom $x^*(\theta) \ge d^* + \varepsilon$ opt out.

The proof mirrors that of Step 1. The set $\{\theta \mid x^*(\theta) \ge d^* + \varepsilon\}$ is also compact, so we define

$$\vartheta_H = \max_{\theta \in \{\theta' \mid x^*(\theta') \ge d^* + \varepsilon\}} \left[V(x^*(\theta), 1 - \tau\left(x^*(\theta)\right), \theta) - V(d^* + \frac{\varepsilon}{2}, 1 - \tau\left(d^* + \frac{\varepsilon}{2}\right), \theta) \right],$$

and observe that $\vartheta_H > 0$. Step 2 then follows from the fact that there exists K_H^ε such that $\overline{\gamma}_k < \vartheta_H$ for all $k > K_H^\varepsilon$.

Step 3: With any default $d \in \left[d^* - \frac{\varepsilon}{2}, d^* + \frac{\varepsilon}{2}\right]$ and $k > \max\{K_L^\varepsilon, K_H^\varepsilon\}$, all workers for whom $x^*(\theta) \notin (d^* - \varepsilon, d^* + \varepsilon)$ opt out.

Consider a worker for whom $x^*(\theta) \le d^* - \varepsilon$. By Step 1, for $k > K_L^\varepsilon$ we know that

$$V(x^*(\theta), 1 - \tau\left(x^*(\theta)\right), \theta) - \overline{\gamma}_k > V(d^* - \frac{\varepsilon}{2}, 1 - \tau\left(d^* - \frac{\varepsilon}{2}\right), \theta) \tag{5}$$

With $d \in \left[d^* - \frac{\varepsilon}{2}, d^* + \frac{\varepsilon}{2}\right]$, we also have

$$V(d^* - \frac{\varepsilon}{2}, 1 - \tau\left(d^* - \frac{\varepsilon}{2}\right), \theta) \ge V(d, 1 - \tau\left(d\right), \theta) \tag{6}$$

To see why, let $q \in (0,1)$ satisfy $qx^*(\theta) + (1-q)d = d^* - \frac{\varepsilon}{2}$, and define $\widetilde{z} = 1 - q\tau(x^*(\theta)) - (1-q)\tau(d)$. Because V is quasiconcave,

$$V(d^* - \frac{\varepsilon}{2}, \widetilde{z}, \theta, 0) \geq \min\{V(x^*(\theta), 1 - \tau(x^*(\theta)), \theta), V(d, 1 - \tau(d), \theta)\} = V(d, 1 - \tau(d), \theta)$$

Because τ is convex, $V(d^* - \frac{\varepsilon}{2}, 1 - \tau\left(d^* - \frac{\varepsilon}{2}\right), \theta) \geq V(d^* - \frac{\varepsilon}{2}, \widetilde{z}, \theta, 0)$. Combining these inequalities yields (6). Combining (5) and (6), we obtain

$$V(x^*(\theta), 1 - \tau(x^*(\theta)), \theta) - \overline{\gamma}_k > V(d, 1 - \tau(d), \theta),$$

which implies that the worker opts out of d, as desired.

The case of any worker for whom $x^*(\theta) \geq d^* + \varepsilon$ is completely analogous, but employs Step 2 instead of Step 1.

Step 4: Now we prove the claim. Because $d_k \to d^*$, there exists K_I^ε such that, for $k > K_I^\varepsilon$, we have $d_k \in \left[d^* - \frac{\varepsilon}{2}, d^* + \frac{\varepsilon}{2}\right]$. Defining $K^\varepsilon = \max\{K_L^\varepsilon, K_H^\varepsilon, K_I^\varepsilon\}$, we see that for $k > K^\varepsilon$ and with a default rate of d_k, all workers for whom $x^*(\theta) \notin (d^* - \varepsilon, d^* + \varepsilon)$ opt out.

Having established the claim, we now complete the proof of the theorem. If $d^* \notin \mathcal{A}$, then the measure of workers with ideal points in $(d^* - \varepsilon, d^* + \varepsilon)$, call it $y(\varepsilon)$, converges to zero along with ε. But plainly $y(\varepsilon) \geq \int_{D(d_k, \overline{\gamma}_k)} dH^\theta(\theta)$ for $k > K^\varepsilon$. Consequently, we have $\int_{D(d_k, \overline{\gamma}_k)} dH^\theta(\theta) \to 0$, and thus the desired contradiction. $\square$

Proof of Theorem 5

Throughout this proof, we use i to denote a particular individual. BR define the relation R_i^* as follows: xR_i^*y iff $y \in C_i(X, f)$ implies $x \in C_i(X, f)$ for all $(X, f) \in \mathcal{G}$. Also, x is a weak Pareto improvement over y iff xR_i^*y for every individual and xP_i^*y for some individual.

Part 1: Regardless of whether the welfare-relevant domain is restricted or unrestricted, offering a plan with the $d = 0$, where choices are made in frame $f \geq f_M$ for the cases of time inconsistency and inattentiveness, yields a weak generalized Pareto improvement over no plan.

Partition the set of employees into two groups, those who opt out and those who do not (both of which have positive measure under our assumptions). Those who do not opt out receive the bundle $(e, x, z) = (0, 0, 1)$ both with and without the plan. By definition, $(0, 0, 1)R_i^*(0, 0, 1)$. A worker who opts out chooses some bundle (e', x', z'), where $x' > 0$ and $z' < 1$, over the bundle $(0, 0, 1)$. With anchoring, the choice is made in frame $f = d = 0$, and our monotonicity assumption implies that the same worker would choose (e', x', z') over $(0, 0, 1)$ in any frame $f > 0$. With time inconsistency, if frame 0 is welfare relevant, the choice is made in frame 0, and $\beta_0^i < 1 = \beta_{-1}^i$ implies that the same worker would choose (e', x', z') over $(0, 0, 1)$ in frame -1; if frame 0 is not welfare relevant, choices can be made in either frame but are evaluated from the perspective of frame -1, and the same implication follows from the fact that $\beta_f^i < 1 = \beta_{-1}^i$ where $f \in \{-1, 0\}$ is the decision frame. With inattentiveness, the choice is made in some $f \geq f_M$, and $\chi^i(f') \leq \chi^i(f_M) \leq \chi^i(f)$ for any welfare-relevant frame f' implies that the same worker would choose (e', x', z') over $(0, 0, 1)$ in f'. Thus, we have $(e', x', z')P_i^*(0, 0, 1)$ for those who opt out.[3] The desired conclusion follows directly.

Part 2: Regardless of whether the welfare-relevant domain is restricted or unrestricted, and regardless of the prevailing choice frame for the cases of time inconsistency and inattentiveness, offering a plan with $d > 0$ does not yield a weak generalized Pareto improvement over no plan.

Consider the set of workers for whom $x^*(\theta^i, \overline{x}) = 0$ in the case of anchoring, and $x^*(\theta^i) = 0$ in the case of time inconsistency or inattentiveness (both of which have positive measure under our assumptions). In the prevailing choice frame, call it f', such workers either opt out to $x = 0$ and receive the bundle $(e', 0, 1)$ (in the case of anchoring, this statement follows because, by our monotonicity requirement, $x^*(\theta^i, \overline{x}) = 0$ implies $x^*(\theta^i, f) = 0$ for all f, including f'), or fail to opt out and receive the bundle $(0, d, 1 - \tau(d))$. In the first case $(0, 0, 1)P_i^*(e', 0, 1)$, and in the second $(0, 0, 1)P_i^*(0, d, 1 - \tau(d))$ (in the cases of time inconsistency and inattentiveness because $x^*(\theta^i) = 0$, and in the case of anchoring because because $x^*(\theta^i, \overline{x}) = 0$ implies $x^*(\theta^i, f) = 0$ for all f). The desired conclusion follows directly.

[3] The same reasoning implies that, for those who are willing to either opt out or choose the default, we have $(e', x', z')R_i^*(0, 0, 1)$.

Part 3: For the cases of time inconsistency and inattentiveness, a plan with $d = 0$ does not achieve a weak generalized Pareto improvement over no plan if choices are made in some frame $f' < f_M$.

Suppose $d = 0$ and that choices are made in some frame $f' < f_M$. For the case of time inconsistency, where this inequality plainly implies $f' = -1$ and $f_M = 0$, we consider the set of workers for whom $\gamma^i \in \left(\beta_0^i \Delta(\theta^i, d, \pi), \Delta(\theta^i, d, \pi)\right)$, which has positive measure (because the interval is open for all β_0^i and θ^i). For the case of inattentiveness, consider the set of workers for whom $\gamma^i \in \left(\Delta(\theta^i, d, \pi) - \chi^i(f_M), \Delta(\theta^i, d, \pi) - \chi^i(f')\right)$, which also has positive measure (again because the interval is open for all χ^i and θ^i, and because, by assumption, for all f the set of workers with $\Delta(\theta^i, d, \pi) - \chi^i(f) > 0$ has strictly positive measure). Because the choice frame is f', any such worker opts out and receives the bundle $\left(e', x^*(\theta^i), 1 - \tau(x^*(\theta^i))\right)$. But the same worker would choose the bundle $(0, 0, 1)$ over $\left(e', x^*(\theta^i), 1 - \tau(x^*(\theta^i))\right)$ in frame f_M. Thus, we do *not* have $\left(e', x^*(\theta^i), 1 - \tau(x^*(\theta^i))\right) R_i^* (0, 0, 1)$.

Part 4: For the cases of time inconsistency and inattentiveness, fixing $d = 0$, a plan with choices made in frame f_M achieves a weak generalized Pareto improvement over any plan with choice made in frame $f' > f_M$.

Suppose $d = 0$ and consider the choice frames f' and f_M with $f' > f_M$. For the case of time inconsistency, where this inequality plainly implies $f' = 0$ and $f_M = -1$, we partition the set of workers as follows: for group L, $\gamma^i < \beta_0^i \Delta(\theta^i, d, \pi)$; for group I, $\gamma^i \in \left[\beta_0^i \Delta(\theta^i, d, \pi), \Delta(\theta^i, d, \pi)\right]$; and for group H, $\gamma^i > \Delta(\theta^i, d, \pi)$. For the case of inattentiveness, we partition the set of workers as follows: for group L, $\gamma^i < \Delta(\theta^i, d, \pi) - \chi^i(f')$; for group I, $\gamma^i \in \left(\Delta(\theta^i, d, \pi) - \chi^i(f'), \Delta(\theta^i, d, \pi) - \chi^i(f_M)\right)$; and for group H, $\gamma^i > \Delta(\theta^i, d, \pi) - \chi^i(f_M)$. (We will consider workers at the boundaries between these groups separately below.) For the same reasons as in Part 3, each of these groups has positive measure. Those in group L opt out and receive the bundle $\left(e', x^*(\theta^i), 1 - \tau(x^*(\theta^i))\right)$ in both frames, and by definition $\left(e', x^*(\theta^i), 1 - \tau(x^*(\theta^i))\right) R_i^* \left(e', x^*(\theta^i), 1 - \tau(x^*(\theta^i))\right)$. Those in group H end up with $(0, d, 1 - \tau(d))$ in both frames because they do not opt out, and by definition

$(0, d, 1-\tau(d)) R_i^*(0, d, 1-\tau(d))$. Those in group I opt out in frame f_M, receiving bundle $\left(e', x^*(\theta^i), 1-\tau(x^*(\theta^i))\right)$, and do not opt out in frame f', receiving bundle $(0, d, 1-\tau(d))$. Moreover, all such workers would choose $\left(e', x^*(\theta^i), 1-\tau(x^*(\theta^i))\right)$ over $(0, d, 1-\tau(d))$ in all frames $f < f_M$. Thus, $\left(e', x^*(\theta^i), 1-\tau(x^*(\theta^i))\right) P_i^*(0, d, 1-\tau(d))$. We treat workers at the boundary between groups L and I the same as members of group I if they opt out in frame f', and the same as members of group L if they do not opt out in frame f'. We treat workers at the boundary between groups I and H the same as members of group H if they do not opt out in frame f_M; if they do opt out in frame f_M, we still have $\left(e', x^*(\theta^i), 1-\tau(x^*(\theta^i))\right) R_i^*(0, d, 1-\tau(d))$ because they choose $\left(e', x^*(\theta^i), 1-\tau(x^*(\theta^i))\right)$ over $(0, d, 1-\tau(d))$ strictly in all frames $f < f_M$, and weakly in frame f_M. The desired conclusion follows directly. $\square$

Proof of Theorem 6

With zero opt-out costs, EV evaluated in frame f is given by the value of m_A^1 satisfying

$$V(0, 1+m_A^1, \theta, f) = V(x^*(\theta, d), 1-\tau\left(x^*(\theta, d)\right), \theta, f)$$

Because V is strictly increasing in z, the value of d that maximizes the RHS also maximizes EV evaluated in frame f. By definition, the solution to $\max_{x \in X} V(x, 1-\tau(x), \theta, f)$ is $x = x^*(\theta, f)$. It follows immediately that the solution to $\max_{d \in X} V(x^*(\theta, d), 1-\tau\left(x^*(\theta, d)\right), \theta, f)$ is $d = f$. Thus, because EV_A is evaluated from the perspective of frame $f = \overline{x}$, it is maximized by setting $d = \overline{x}$, and because EV_B is evaluated from the perspective of frame $f = 0$, it is maximized by setting $d = 0$.

We complete the proof by showing that EV_A and EV_B are respectively non-decreasing and non-increasing in d on $[0, \overline{x}]$. First observe that, as a consequence of our monotonicity assumption, $x^*(\theta, d)$ is non-decreasing in d. Second, note that $V(x, 1-\tau(x), \theta, 0)$ is non-increasing and $V(x, 1-\tau(x), \theta, \overline{x})$ non-decreasing in x on $[0, \overline{x}]$. To see why, consider any x', x'' with $x'' > x' > 0$. Let $z' = 1-\tau(x')$, $z'' = 1-\tau(x'')$, and $\widetilde{z} = (1-\tau(x''))\frac{x'}{x''}$. Because V is quasiconcave, $V(x', \widetilde{z}, \theta, 0) \geq \min\left\{V(0, 1, \theta, 0), V(x'', 1-\tau(x''), \theta, 0)\right\} = V(x'', 1-$

$\tau(x''), \theta, 0)$. Because τ is convex, $V(x', 1-\tau(x'), \theta, 0) \geq V(x', \widetilde{z}, \theta, 0)$. Combining these inequalities, we have $V(x', 1-\tau(x'), \theta, 0) \geq V(x'', 1-\tau(x''), \theta, 0)$, as desired. An analogous argument establishes $V(x', 1-\tau(x'), \theta, \overline{x}) \leq V(x'', 1-\tau(x''), \theta, \overline{x})$. Third, it follows as a consequence of the first two steps that $V(x^*(\theta, d), 1-\tau\left(x^*(\theta, d)\right), \theta, 0)$ is non-increasing and $V(x^*(\theta, d), 1-\tau\left(x^*(\theta, d)\right), \theta, \overline{x})$ non-decreasing in d on $[0, \overline{x}]$. The desired properties then follow from the fact that $V(0, 1+m_A^1, \theta, f)$ is non-decreasing in m_A^1. $\square$

2 Additional simulation results

In this section we provide the following supplementary figures. Figure A.1 corresponds to Figure 5 in the text, except it depicts simulations of employer contributions and lost government revenues for the model with anchoring effects. Figures A.2 through A.7 correspond to Figures 8, 9, and 11 through 14 in the text, except the horizontal axis is extended to display outcomes for very high default rates.

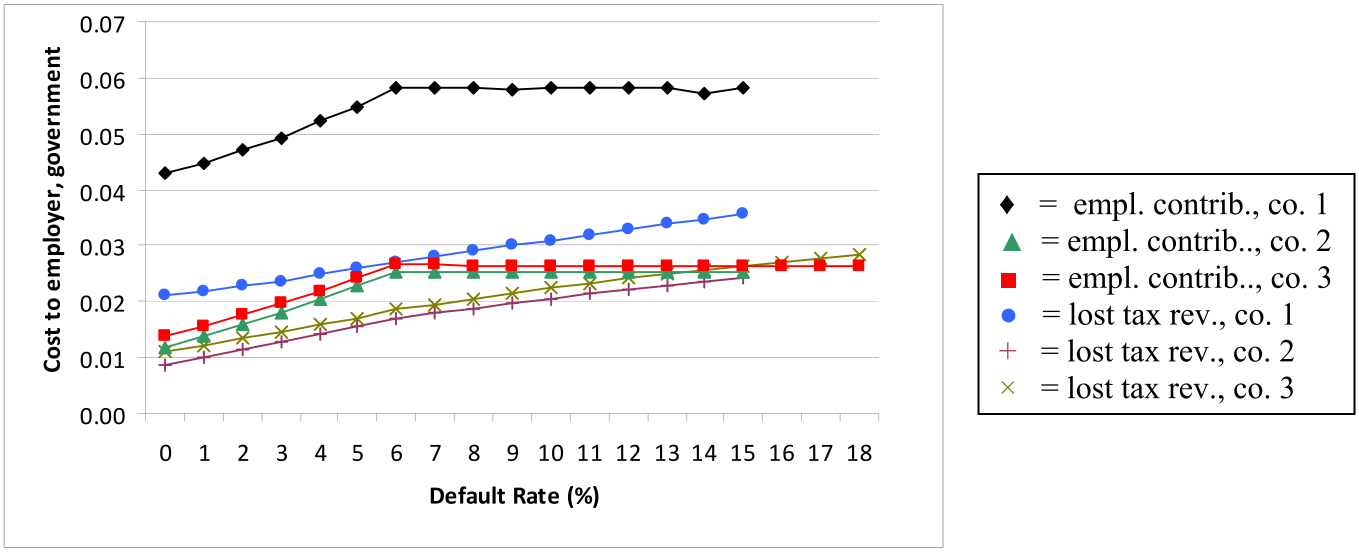

Figure A.1(a): Employer contributions and lost government revenue versus default rate, with anchoring effects and an employer match

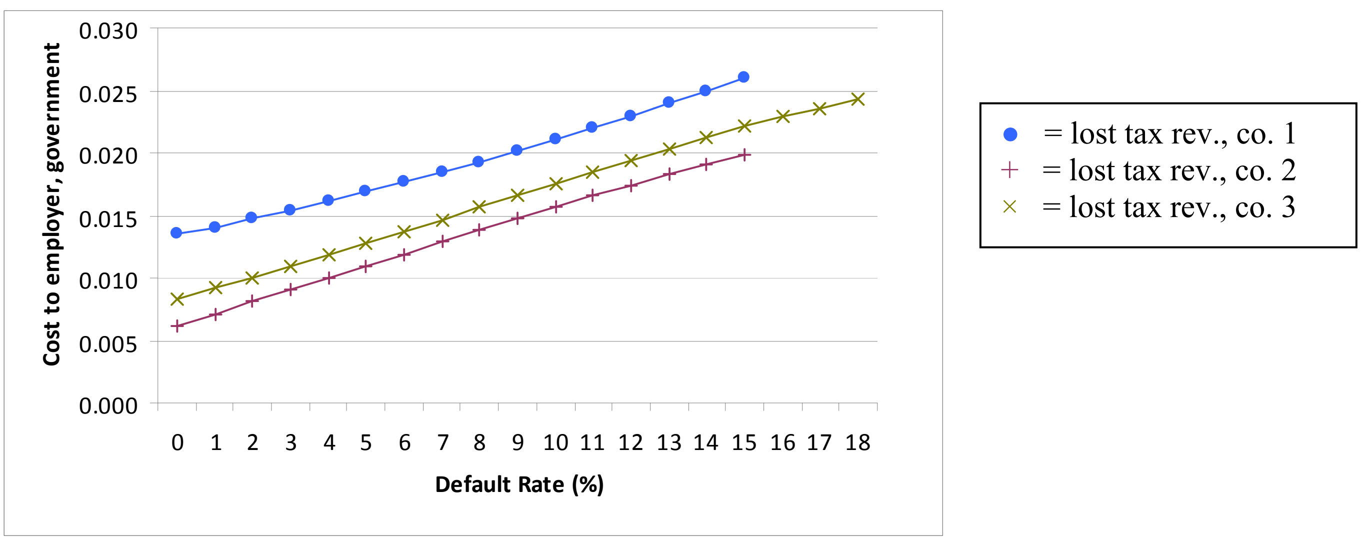

Figure A.1(b): Lost government revenue versus default rate, with anchoring effects and no employer match

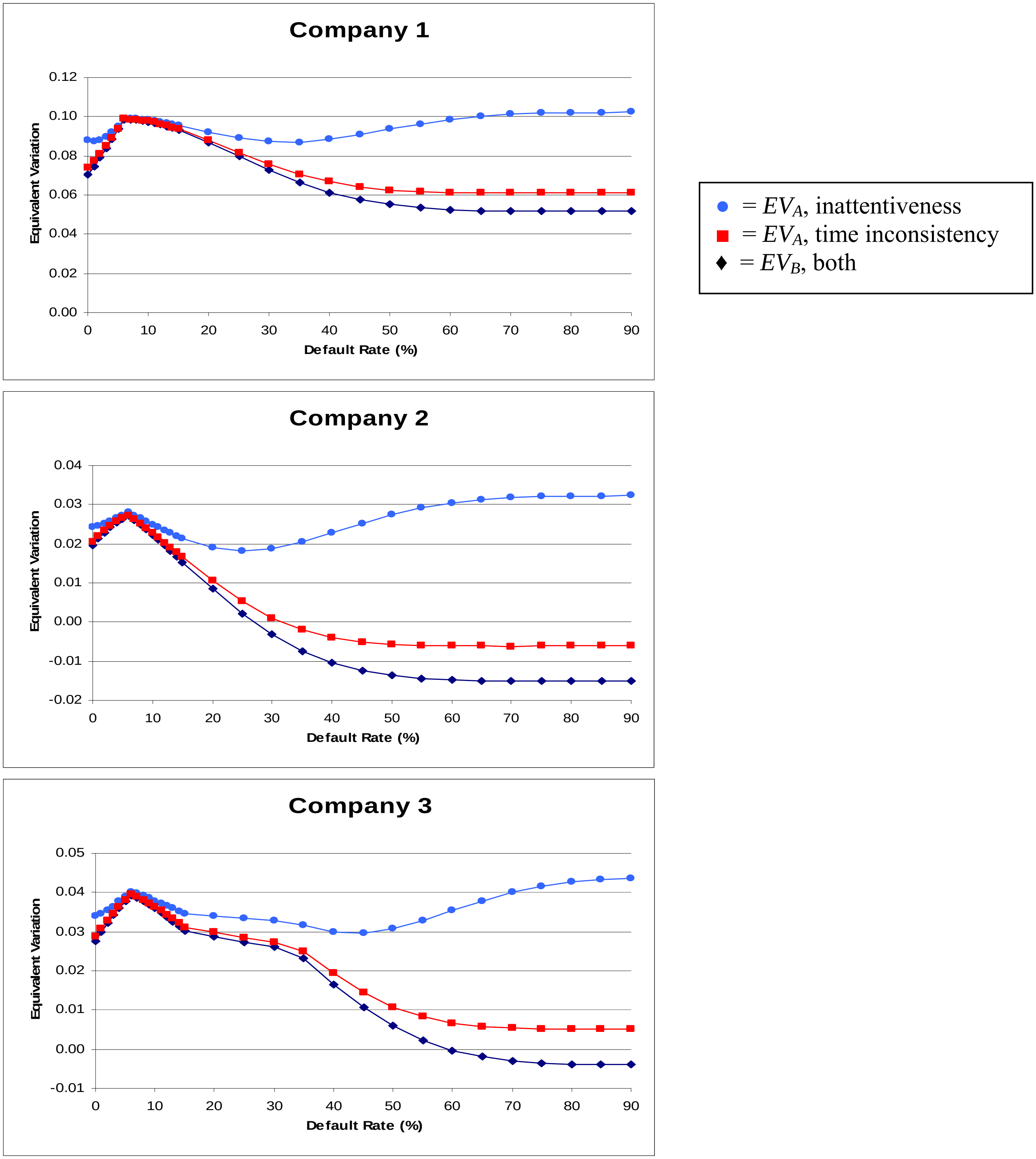

Figure A.2: Average equivalent variation with time inconsistency or inattentiveness and an employer match

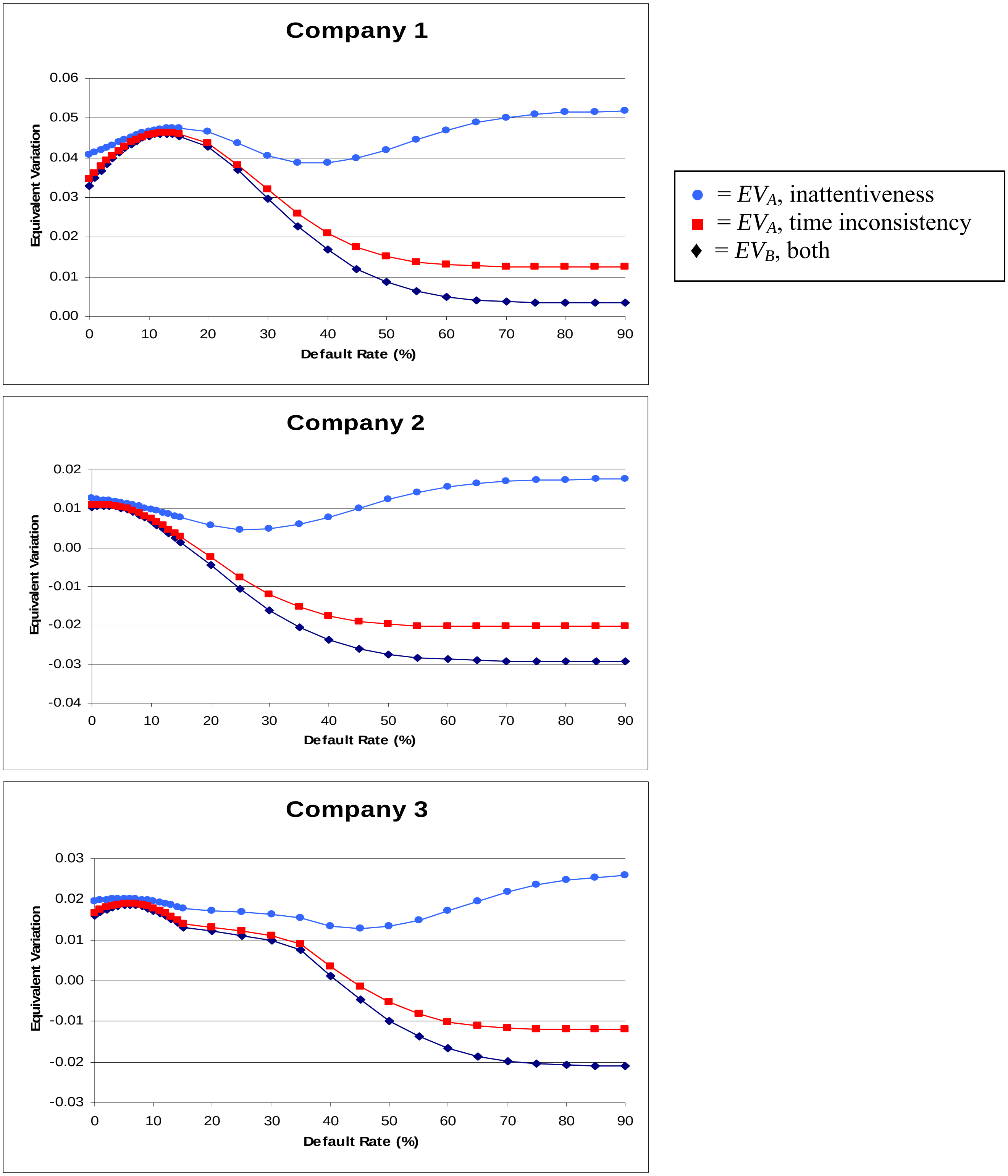

Figure A.3: Average equivalent variation with time inconsistency or inattentiveness and no employer match

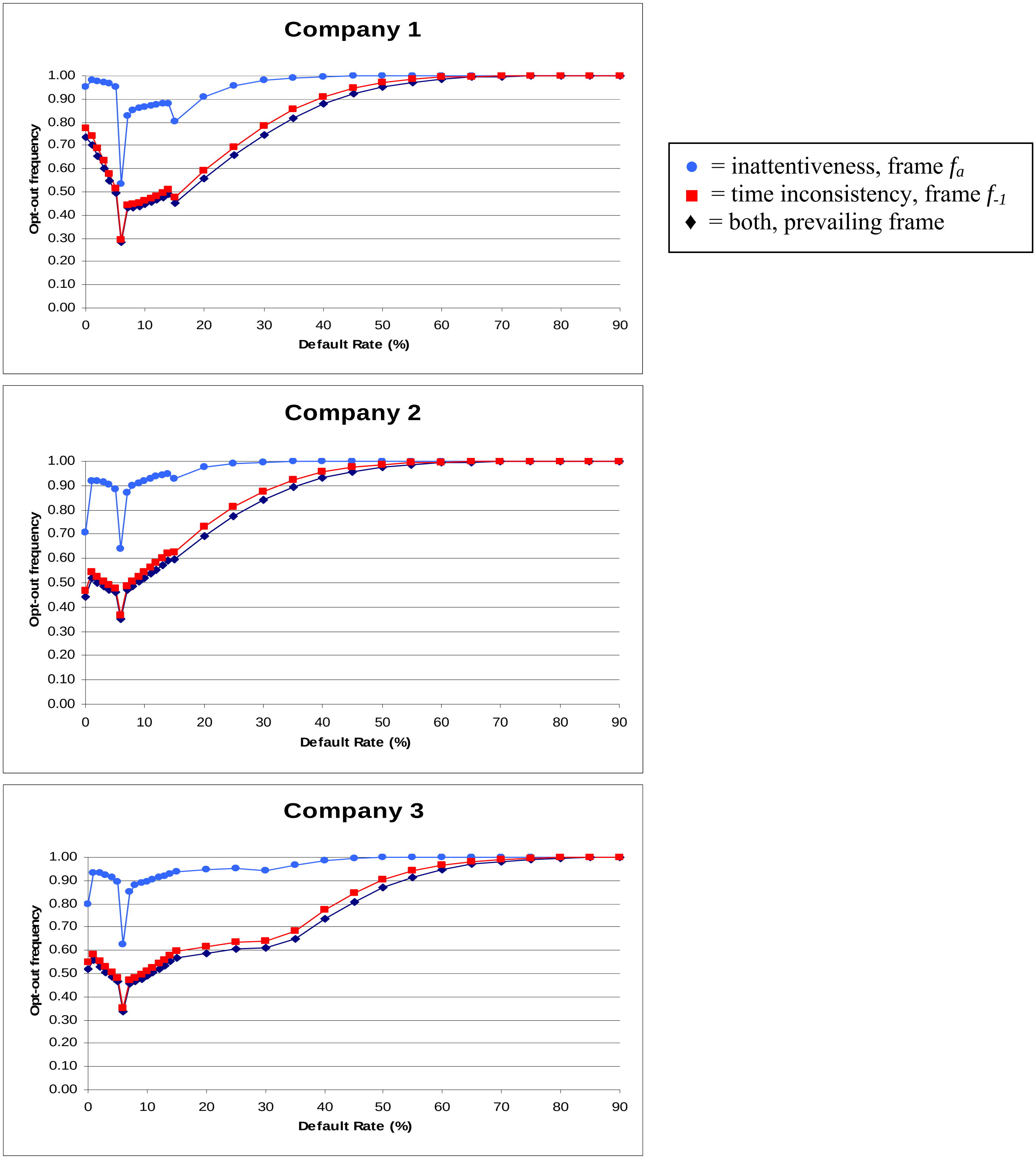

Figure A.4: Opt-out frequencies for various decision frames, with an employer match

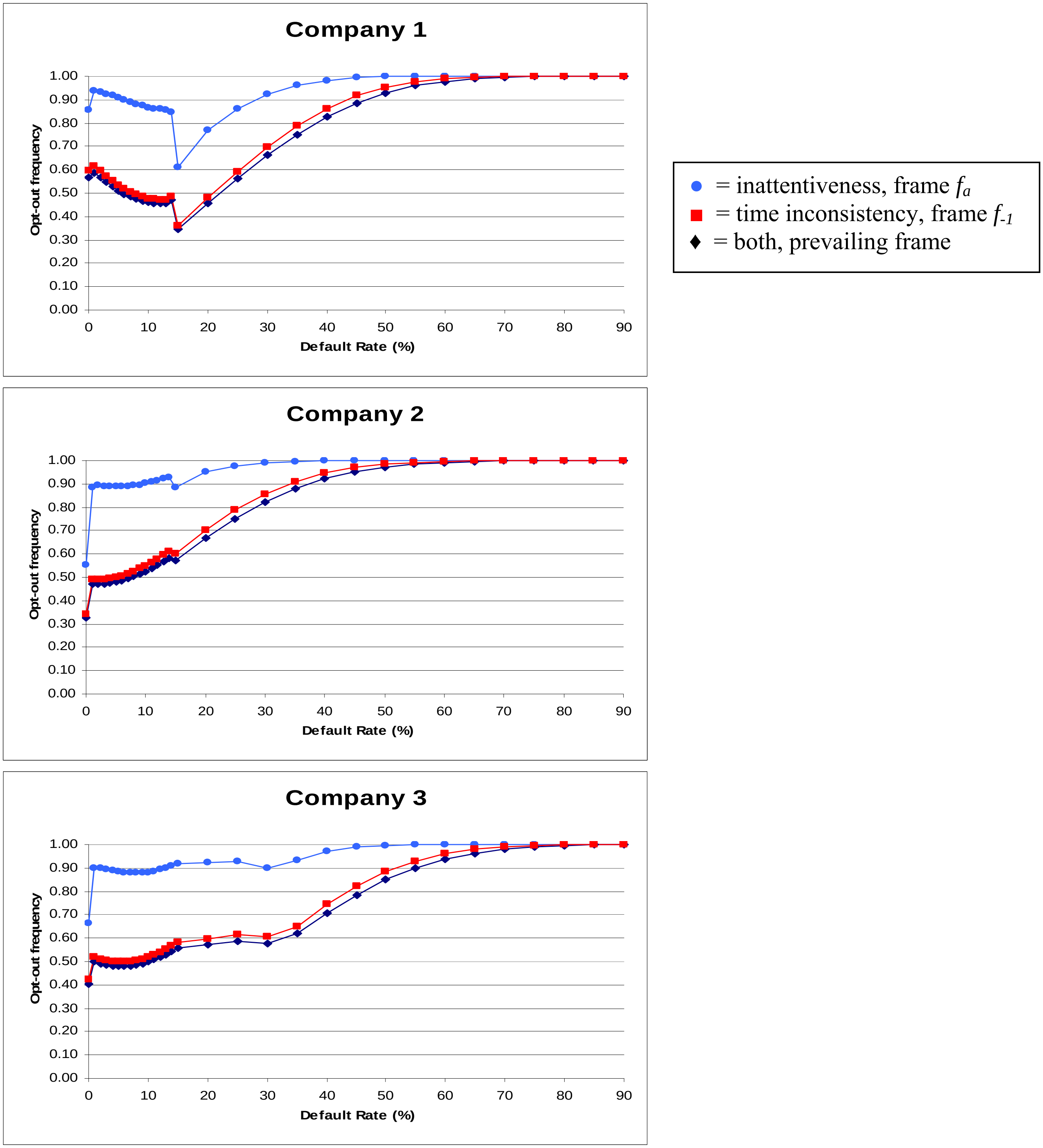

Figure A.5: Opt-out frequencies for various decision frames, without an employer match

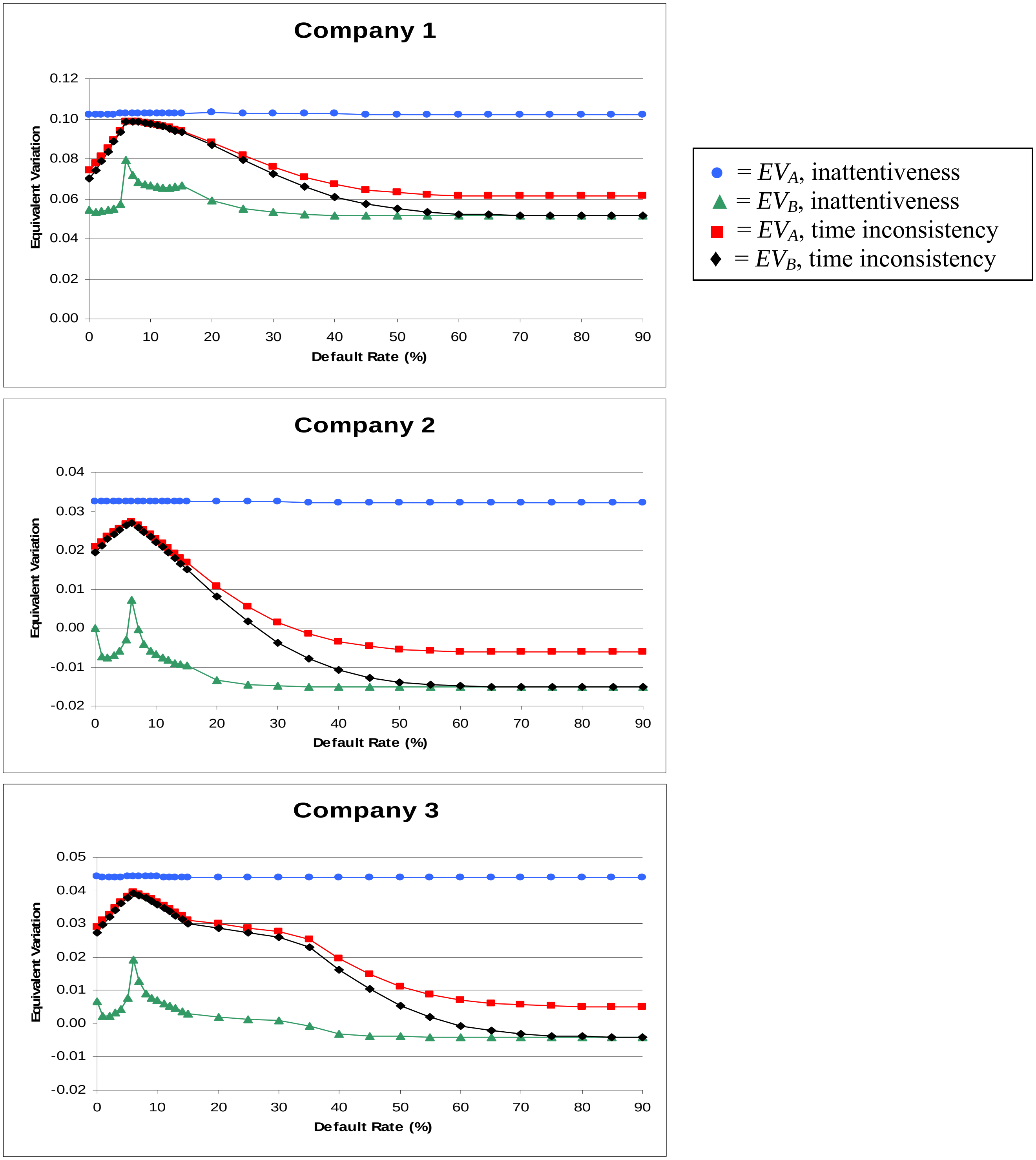

Figure A.6: Average equivalent variation, decisions made in the forward-looking frame for time inconsistency and in the fully attentive frame for inattentiveness, with an employer match

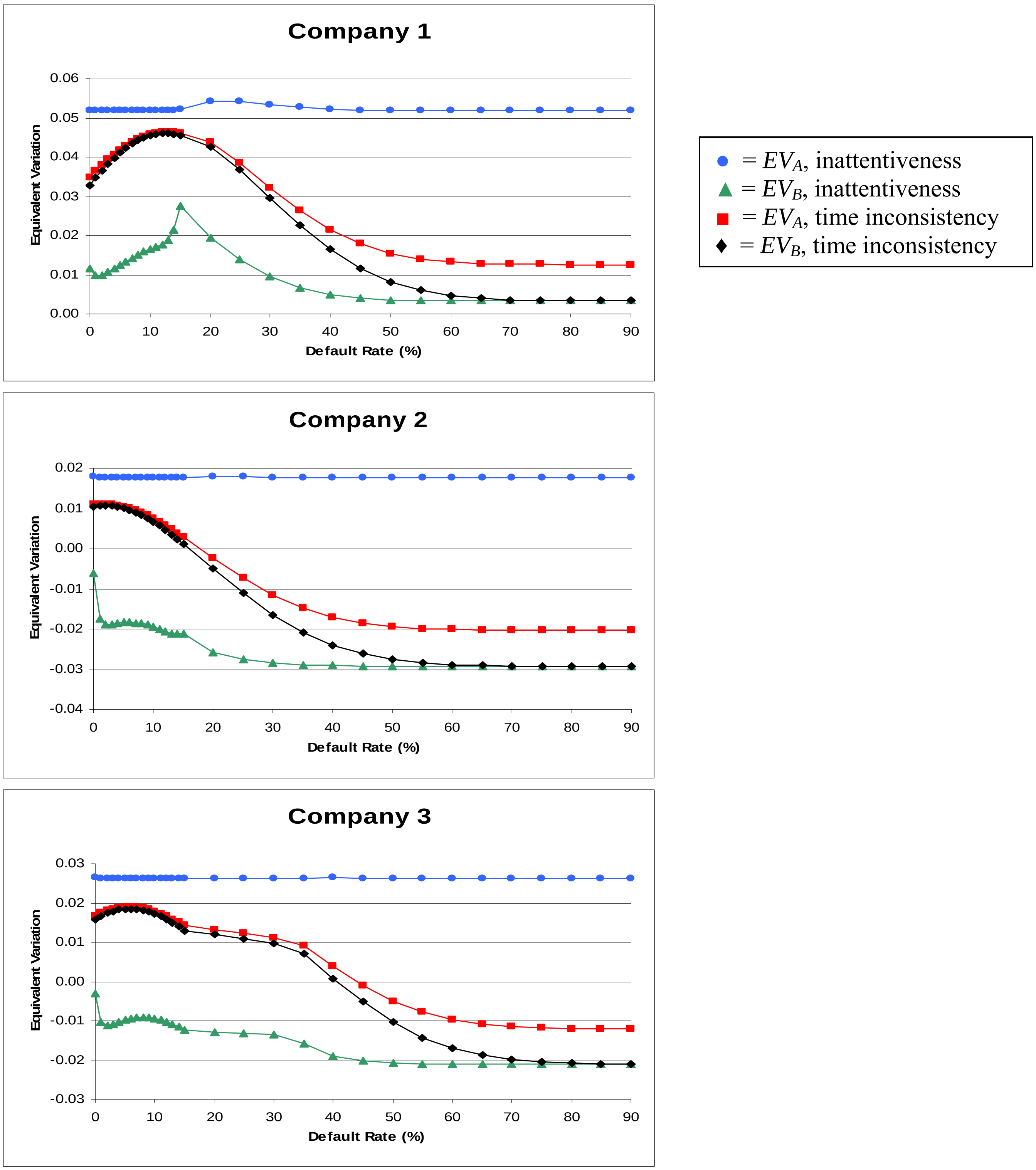

Figure A.7: Average equivalent variation, decisions made in the forward-looking frame for time inconsistency and in the fully attentive frame for inattentiveness, with no employer match