

**Online Appendices for “Self Control and Commitment:  
Can Decreasing the Liquidity of a Savings Account Increase Deposits?”**

JOHN BESHEARS, HARVARD UNIVERSITY AND NBER

JAMES J. CHOI, YALE UNIVERSITY AND NBER

CHRISTOPHER HARRIS, UNIVERSITY OF CAMBRIDGE

DAVID LAIBSON, HARVARD UNIVERSITY AND NBER

BRIGITTE C. MADRIAN, HARVARD UNIVERSITY AND NBER

JUNG SAKONG, UNIVERSITY OF CHICAGO

August 29th 2013

This draft: July 23, 2015

A. INTRODUCTION

In this appendix we provide a complete analysis of the mechanism-design problem described in the main body of the paper.

B. PRELIMINARIES

**B.1. Functions of Bounded Variation.** We begin by discussing the concept of bounded variation. This concept will be used to formulate our assumptions on the distribution function  $F$  in the subsection immediately following this one, namely Section B.2. More importantly, it plays an essential role in our proof of sufficiency in a much later section, namely Section P.

The simplest definition of a function of bounded variation is probably that given in the main text: a function  $f : (0, \infty) \rightarrow \mathbb{R}$  is of bounded variation iff it is the difference of two bounded and non-decreasing functions  $f^+, f^- : (0, \infty) \rightarrow \mathbb{R}$ . This definition forms the starting point for the definition that we shall use. However, it needs to be developed into a form that is more convenient for the Lagrangean analysis below.

The first step is to collect the functions of bounded variation into equivalence classes. Intuitively speaking, two functions of bounded variation are equivalent iff they differ only at their points of discontinuity. This step is analogous to the first step in defining spaces of Lebesgue integrable functions. (In that case, one collects the Lebesgue integrable functions into equivalence classes. Two Lebesgue integrable functions are equivalent if they differ only on a set of measure 0.)

The second step is to place a norm on the resulting equivalence classes in such a way that the limit of a sequence of equivalence classes is again a suitable equivalence classes. (This step is analogous to the second step in defining spaces of Lebesgue integrable functions.) The main idea here is to note that, since  $f^+$  and  $f^-$  are non-decreasing, they are effectively the distribution functions of a pair of non-negative bounded measures  $df^+$  and  $df^-$ . Of course, neither  $df^+$  nor  $df^-$  is unique. But their difference  $df = df^+ - df^-$  is. The main component of the norm is therefore the total variation  $\|df\|_{TV}$  of  $df$ . The other idea is to note that, while  $\|df\|_{TV}$  effectively controls the derivative of  $f$ , it does not control the level of  $f$ . The remaining component of the norm can therefore be taken to be  $|f_R(1)|$ , where  $f_R(1)$  is the limit of  $f$  from the right at 1.

The best way of understanding how these ideas work is to note that we can easily reconstruct  $f$  from  $df$  and  $f_R(1)$ . For all  $\theta \in (0, 1)$ , we have

$$f_R(\theta) = f_R(1) - df((\theta, 1])$$

and

$$f_L(\theta) = f_R(1) - df([\theta, 1]),$$

where  $f_R(\theta)$  and  $f_L(\theta)$  are the limits of  $f$  from the right and left at  $\theta$ . And for all  $\theta \in (1, \infty)$ , we have

$$f_R(\theta) = f_R(1) + df((1, \theta])$$

and

$$f_L(\theta) = f_R(1) + df((1, \theta)).$$

We also need to work with the space  $\mathcal{BV}(\Theta, \mathbb{R})$  of functions of bounded variation on  $\Theta = [\underline{\theta}, \bar{\theta}]$ . By analogy with our discussion of the space  $\mathcal{BV}((0, \infty), \mathbb{R})$ , it should be clear that we can endow  $\mathcal{BV}(\Theta, \mathbb{R})$  with the norm

$$\|f\|_{\mathcal{BV}} = |f_R(\theta_0)| + \|df\|_{TV},$$

where  $\theta_0$  is a fixed element of  $(\underline{\theta}, \bar{\theta})$  and  $df$  is a bounded measure on  $\Theta$ . There is, however, one surprise: a function  $f \in \mathcal{BV}(\Theta, \mathbb{R})$  has a limit on the left at  $\underline{\theta}$  and a limit on the right at  $\bar{\theta}$ . Indeed, we have

$$f_L(\underline{\theta}) = f_R(\theta_0) - df([\underline{\theta}, \theta_0])$$

and

$$f_R(\bar{\theta}) = f_R(\theta_0) + df((\theta_0, \bar{\theta}]) .$$

To summarize, we denote the space of functions of bounded variation on  $(0, \infty)$  by  $\mathcal{BV}((0, \infty), \mathbb{R})$ , and we denote the space of functions of bounded variation on  $\Theta = [\underline{\theta}, \bar{\theta}]$  by  $\mathcal{BV}(\Theta, \mathbb{R})$ . Unless explicitly stated to the contrary, we shall always use the right-continuous representative of a function of bounded variation. We will usually denote this representative simply by  $f$ , but we will occasionally denote it by  $f_R$  for emphasis. We will denote the left-continuous representative of  $f$  by  $f_L$ .

**B.2. Assumptions on  $F$ .** We are now in a position to introduce our assumptions on the distribution function  $F$  of the taste shock  $\theta$ . They are:

**A1** Both  $F$  and  $F'$  are functions of bounded variation on  $(0, \infty)$ .

**A2** The support of  $F'$  is contained in  $[\underline{\theta}, \bar{\theta}]$ , where  $0 < \underline{\theta} < \bar{\theta} < \infty$ .

**A3** There exists  $\theta_M \in [\underline{\theta}, \bar{\theta}]$  such that: (i)  $G' \geq 0$  on  $(0, \theta_M)$ ; and (ii)  $G' \leq 0$  on  $(\theta_M, \infty)$ .

Here  $G$  is given by the formula  $G(\theta) = (1 - \beta)\theta F'(\theta) + F(\theta)$ . If A1 holds then  $G$ , like  $F$  and  $F'$ , is a function of bounded variation on  $(0, \infty)$ .

**B.3. The Support of  $F'$  is Connected.** Fourth, we note that either  $\beta = 1$ , in which case the analysis is trivial, or  $\beta < 1$ , in which case the support of  $F'$  is connected.<sup>1</sup> More precisely, we have:

**Proposition 1.** *Suppose that  $\beta < 1$  and that A1-A3 are satisfied. Then there exist  $\underline{\kappa}, \bar{\kappa} \in [\underline{\theta}, \bar{\theta}]$  such that: (i)  $\underline{\kappa} < \bar{\kappa}$ ; (ii)  $F' > 0$  on  $(\underline{\kappa}, \bar{\kappa})$ ; and (iii)  $F' = 0$  on  $(0, \infty) \setminus [\underline{\kappa}, \bar{\kappa}]$ .*

In what follows we shall therefore take it that  $\beta < 1$ , and that the support of  $F'$  is  $[\underline{\theta}, \bar{\theta}]$ .

---

<sup>1</sup>Notice that  $F$  is a distribution function, not a distribution. Also,  $F'$  has a dual interpretation. It can be regarded as: either (i) the non-negative finite measure with distribution function  $F$ ; or (ii) the density of that measure with respect to Lebesgue measure. By the same token, the support of  $F'$  has a dual interpretation. It can be regarded as: either (i) the support of the non-negative finite measure  $F'$ ; or (ii) the support of the non-negative function of bounded variation  $F'$ . It makes no difference which of these two interpretations is adopted.

**Proof.** Note first that there exists  $\kappa_1 \in (\underline{\theta}, \bar{\theta})$  such that  $F'(\kappa_1) > 0$ . Otherwise we would have  $F' = 0$  everywhere on  $(0, \infty)$ , by right-continuity of  $F'$ . Next, there exists  $\kappa_2 \in (\kappa_1, \bar{\theta})$  such that  $F' > 0$  on  $(\kappa_1, \kappa_2)$ , again by right-continuity of  $F'$ . Third, put  $\underline{\kappa} = \inf \{\theta \mid F'(\theta) > 0\}$  and  $\bar{\kappa} = \sup \{\theta \mid F'(\theta) > 0\}$ . Then certainly  $\underline{\theta} \leq \underline{\kappa} < \bar{\kappa} \leq \bar{\theta}$ . Fourth, put  $\alpha = \frac{2-\beta}{1-\beta}$ . Then  $G' \geq 0$  iff  $(\theta^\alpha F'(\theta))' \geq 0$  and  $G' \leq 0$  iff  $(\theta^\alpha F'(\theta))' \leq 0$ . There are therefore two possibilities. If  $G' \geq 0$  at  $\theta_M$  (i.e.  $\Delta G(\theta_M) \geq 0$ ), then we must have  $\theta^\alpha F'(\theta) > 0$  for all  $\theta \in (\underline{\kappa}, \theta_M]$  (because  $(\theta^\alpha F'(\theta))' \geq 0$  on this interval) and  $\theta^\alpha F'(\theta) > 0$  for all  $\theta \in (\theta_M, \bar{\kappa})$  (because  $(\theta^\alpha F'(\theta))' \leq 0$  on this interval); and if  $G' \leq 0$  at  $\theta_M$  (i.e.  $\Delta G(\theta_M) \leq 0$ ), then we must have  $\theta^\alpha F'(\theta) > 0$  for all  $\theta \in (\underline{\kappa}, \theta_M)$  and  $\theta^\alpha F'(\theta) > 0$  for all  $\theta \in [\theta_M, \bar{\kappa})$ .<sup>2</sup> Either way, we see that: (i)  $\theta^\alpha F'(\theta) > 0$ , and hence  $F'(\theta) > 0$ , for all  $\theta \in (\underline{\kappa}, \bar{\kappa})$ ; (ii)  $\theta_M \leq \bar{\kappa}$ , for otherwise we would have  $F' > 0$  on the non-empty interval  $(\bar{\kappa}, \theta_M)$ , and this contradicts the choice of  $\bar{\kappa}$ ; and (iii)  $\theta_M \geq \underline{\kappa}$ , for otherwise we would have  $F' > 0$  on the non-empty interval  $(\theta_M, \underline{\kappa})$ , and this contradicts the choice of  $\underline{\kappa}$ . ■

**B.4. Constraints on the Budget Set.** Fifth, recall that self 0 chooses a subset  $B$  of the ambient action set  $A$ , and that self 1’s choice of a consumption pair from  $B$  can therefore be described by a consumption curve  $(c_1, c_2) : [\underline{\theta}, \bar{\theta}] \rightarrow B$ . We consider three possible constraints on  $B$ , namely:

**Constraint 1.**  $B$  is a non-empty compact subset of  $A$ .

**Constraint 2.** The penalty for transferring consumption from period 2 to period 1 is no greater than  $\pi$ .<sup>3</sup>

**Constraint 3.** The penalty for transferring consumption from period 1 to period 2 is no greater than  $\pi$ .<sup>4</sup>

Constraint 1 involves no loss of generality. Indeed, it must be possible for all possible types  $\theta \in \Theta$  to find an optimum within  $B$ . This being the case, we can always take the closure of  $B$  without changing the outcome, since the utility function

---

<sup>2</sup>If  $\theta_M \leq \underline{\kappa}$  then we take the intervals  $(\underline{\kappa}, \theta_M)$  and  $(\underline{\kappa}, \theta_M]$  to be empty. Similarly, if  $\theta_M \geq \bar{\kappa}$ , then we take the intervals  $(\theta_M, \bar{\kappa})$  and  $[\theta_M, \bar{\kappa})$  to be empty.

<sup>3</sup>In other words, for any given  $(c_1, c_2) \in B$  and any  $\Delta c_1 \in [0, \frac{1}{1+\pi} c_2]$ , self 1 can increase her own consumption  $c_1$  by  $\Delta c_1$  at a cost of at most  $(1 + \pi) \Delta c_1$  in terms of the consumption  $c_2$  of self 2.

<sup>4</sup>In other words, for any given  $(c_1, c_2) \in B$  and any  $\Delta c_2 \in [0, \frac{1}{1+\pi} c_1]$ , self 1 can increase the consumption  $c_2$  of self 2 by  $\Delta c_2$  at a cost of at most  $(1 + \pi) \Delta c_2$  in terms of her own consumption  $c_1$ .

is continuous. Finally, since  $A$  itself is compact, so too is the closure of  $B$ . Constraint 2 is an essential part of the formulation of our problem. We wish to avoid extreme outcomes in which self 0 imposes an infinite penalty on self 1 for increasing her own consumption at the expense of self 2. Constraint 3 is simply the mirror image of Constraint 2. It eliminates extreme outcomes in which self 0 imposes an infinite penalty on self 1 for increasing the consumption of self 2 at her own expense.

**Remark 2.** *If we only impose Constraint 2, then the problem is one sided: Constraint 2 places a limit on the cost, in terms of  $c_2$ , of increasing  $c_1$ ; but there is no corresponding limit on the cost, in terms of  $c_1$ , of increasing  $c_2$ . By imposing Constraint 3, we eliminate this asymmetry.*

Now suppose that  $B$  must satisfy all three constraints. Then  $B$  must take one of two forms: either

1. it consists of the single point  $(0, 0)$ ; or
2. its frontier consists of a curve that begins at some  $(0, c_2)$  such that  $c_2 > 0$ , slopes downwards with slope between  $-(1 + \pi)$  and  $-(1 + \pi)^{-1}$ , and ends at some  $(c_1, 0)$  such that  $c_1 > 0$ .

Self 0 will never choose the first option, since the optimal pooling point on the frontier of the ambient budget set  $A$  is preferable. (By the same token, self 0 will never choose a  $B$ , the frontier of which is close to  $(0, 0)$ .) But, if she chooses the second option, then the resulting consumption curve  $(c_1, c_2)$  will be interior. That is, we will have  $c_1, c_2 > 0$  on  $\Theta$ .<sup>5</sup>

The ideal approach to our problem would therefore be to impose all three constraints on  $B$ , and to solve the optimization problem of self 0 subject to these constraints. One could then verify ex post that Constraint 3 was not binding.<sup>6</sup>

In practice, we shall take a shortcut. Rather than working explicitly with Constraint 3, we shall instead replace it by the weaker requirement that consumption curves are interior. Our analysis could, of course, be reworked in such a way as to incorporate Constraint 3 explicitly. But, in practice, this would simply involve an additional notational burden.

---

<sup>5</sup>This follows from our assumption that  $U'_t(0+) = \infty$ .

<sup>6</sup>It turns out that the slope of the optimal budget set is at most  $-1$ . So Constraint 3 certainly is not binding!

**Remark 3.** *The situation would be very different if  $\beta > 1$ . In that case, it would be Constraint 2 that would not bind. We would therefore replace Constraint 2 by the weaker requirement that consumption curves are interior.*

**B.5. Utility Curves.** Suppose accordingly that we are given a  $B$  satisfying Constraints 1 and 2, and that the associated consumption curve is interior. Define a utility curve

$$(u_1, u_2) : [\underline{\theta}, \bar{\theta}] \rightarrow (U_1(0), U_1(\infty)) \times (U_2(0), U_2(\infty))$$

by the formula  $(u_1, u_2)(\theta) = (U_1(c_1(\theta)), U_2(c_2(\theta)))$ . Then  $(u_1, u_2)$  must satisfy the following conditions:

**N1**  $C_1(u_1(\theta)) + C_2(u_2(\theta)) \leq y$  for all  $\theta \in [\underline{\theta}, \bar{\theta}]$ .

**N2**  $u_1$  is non-decreasing and  $u_2$  is non-increasing.

**N3**  $\theta du_1 + \beta du_2 = 0$ .

**N4**  $\beta(1 + \pi) U'_2(C_2(u_2(\theta))) \geq \theta U'_1(C_1(u_1(\theta)))$ .

Here:  $C_t = U_t^{-1}$ ; and  $du_1$  is a non-negative finite measure and  $du_2$  is a non-positive finite measure.

Conversely, suppose that a utility curve  $(u_1, u_2)$  is interior, in the sense that it satisfies  $u_1 > U_1(0)$  and  $u_2 > U_2(0)$  on  $\Theta$ , and that it satisfies Conditions N1-N4. Define  $(c_1, c_2)$  by the formula  $(c_1, c_2)(\theta) = (U_1^{-1}(u_1(\theta)), U_2^{-1}(u_2(\theta)))$ . Then there exists a  $B$  with slope at least  $-(1 + \pi)$  such that  $(c_1, c_2)$  is the consumption curve arising from  $B$ . Moreover  $(c_1, c_2)$  is interior.

**B.6. The CRRA Case.** Suppose now that  $U_1 = U_2 = U$  on  $(0, \infty)$ , and that  $U$  has constant relative risk aversion  $\rho > 0$ . Indeed, suppose for concreteness that  $U$  is given by the formula

$$U(c) = \begin{cases} \frac{c^{1-\rho}-1}{1-\rho} & \text{if } \rho \neq 1 \\ \log(c) & \text{if } \rho = 1 \end{cases}.$$

Then N4 is equivalent to

$$\mathbf{N4'} \quad u_2(\theta) \leq -\frac{1}{\rho} a\left(\frac{\theta}{(1+\pi)\beta}\right) + b\left(\frac{\theta}{(1+\pi)\beta}\right) u_1(\theta),$$

where  $a$  and  $b$  are given by the formulae

$$a(z) = \begin{cases} \frac{z^{1-\frac{1}{\rho}} - 1}{1-\frac{1}{\rho}} & \text{if } \rho \neq 1 \\ \log(z) & \text{if } \rho = 1 \end{cases} \quad (1)$$

and

$$b(z) = z^{1-\frac{1}{\rho}}. \quad (2)$$

**Remark 4.** *It is obvious that N4 becomes weaker as  $\pi$  increases. Since N4' is equivalent to N4, N4' likewise becomes weaker as  $\pi$  increases.*

### C. THE MAIN PROBLEM

Our strategy will be to study a relaxed version of the problem of self 0 in which we maximize self 0's expected utility  $\int (\theta u(\theta) + w(\theta)) dF(\theta)$  subject to N1, N3 and N4', but not N2. Following Luenberger (1969, Sections 8.3 and 8.4, pp. 216-221), we shall need:

1. A vector space<sup>7</sup>  $X$ , which we take to be  $\mathcal{C}(\Theta, \mathbb{R})^2$ . Here:  $\Theta = [\underline{\theta}, \bar{\theta}] \subset (0, \infty)$  is the space of types; and  $\mathcal{C}(\Theta, \mathbb{R})$  is the space of continuous functions from  $\Theta$  to  $\mathbb{R}$ .
2. A convex set<sup>8</sup>  $\Omega \subset X$ , which we take to consist of all

$$(u, w) \in \left( \mathcal{BV}(\Theta, \text{ran}(U)) \cap \mathcal{C}(\Theta, \text{ran}(U)) \right)^2$$

such that

$$\theta du + \beta dw = 0.$$

Here:  $\text{ran}(U)$  is the range of  $U$ ;<sup>9</sup>  $\mathcal{C}(\Theta, \text{ran}(U))$  is the space of continuous functions from  $\Theta$  to  $\text{ran}(U)$ ;  $\mathcal{BV}(\Theta, \text{ran}(U))$  is the space of all functions of bounded variation from  $\Theta$  to  $\text{ran}(U)$ ; and  $du$  and  $dw$  are in general elements of the space  $\mathcal{M}(\Theta, \mathbb{R})$  of finite Borel measures on  $\Theta$ .

---

<sup>7</sup>In our case  $X$  is actually a Banach space. For the Lagrangean analysis, we only need the fact that it is a vector space. When we later use calculus to find necessary and sufficient conditions for the maximization of the Lagrangean, we shall need the fact that it is a normed space. Cf. Luenberger (1969, Lemma 1, p. 227).

<sup>8</sup>In our case  $\Omega$  is actually a cone, the vertex of which is the constant mapping  $\frac{1}{\rho-1}$  when  $\rho \neq 1$  and the constant mapping 0 when  $\rho = 1$ .

<sup>9</sup>I.e.  $\text{ran}(U)$  is  $(\frac{1}{\rho-1}, \infty)$  when  $\rho < 1$ ,  $(-\infty, \infty)$  when  $\rho = 1$  and  $(-\infty, \frac{1}{\rho-1})$  when  $\rho > 1$ .

3. A concave function<sup>10</sup>  $M : \Omega \rightarrow \mathbb{R}$  (the objective function), which we take to be given by the formula

$$M(u, w) = \int \left( \theta u(\theta) + w(\theta) \right) dF(\theta).$$

4. A normed space<sup>11</sup>  $Z$ , which we take to be  $\mathcal{C}(\Theta, \mathbb{R})$ .
5. A closed convex cone  $P$  in  $Z$  with vertex 0 and non-empty interior, which we take to be  $\mathcal{C}(\Theta, [0, \infty))$ . With this choice of  $P$ ,  $z_1 \geq z_2$  iff  $z_1(\theta) \geq z_2(\theta)$  for all  $\theta \in \Theta$  and  $z_1 > z_2$  iff  $z_1(\theta) > z_2(\theta)$  for all  $\theta \in \Theta$ . In other words,  $P$  is the positive cone of  $Z$ .
6. The space  $Z^*$  of continuous linear functionals on  $Z$ . Since  $Z = \mathcal{C}(\Theta, \mathbb{R})$ ,  $Z^*$  can be identified with  $\mathcal{M}(\Theta, \mathbb{R})$ .
7. The positive cone  $P^*$  of  $Z^*$ . Since  $P = \mathcal{C}(\Theta, [0, \infty))$ ,  $P^*$  can be identified with  $\mathcal{M}(\Theta, [0, \infty))$  (i.e. the space of non-negative finite Borel measures on  $\Theta$ ).
8. Concave mappings  $G_1, G_2 : \Omega \rightarrow Z$  (the constraint mappings),<sup>12</sup> which we take to be given by the formulae

$$(G_1(u, w))(\theta) = y - C(u(\theta)) - K(w(\theta))$$

and

$$(G_2(u, w))(\theta) = b \left( \frac{\theta}{(1+\pi)^\beta} \right) u(\theta) - \frac{1}{\rho} a \left( \frac{\theta}{(1+\pi)^\beta} \right) - w(\theta),$$

where  $C = K = U^{-1} : \text{ran}(U) \rightarrow (0, \infty)$ , and  $a$  and  $b$  are given by the formulae (1) and (2).

---

<sup>10</sup>In our case  $M$  is actually defined on the whole of  $X$  (and not just on  $\Omega$ ), and it is linear (and not just concave).

<sup>11</sup>In our case,  $Z$  is actually a Banach space, and not just a normed space.

<sup>12</sup>In our case  $G_1$  is actually defined on  $\Xi = \mathcal{C}(\Theta, \text{ran}(U))^2$  (and not just on  $\Omega$ ). This will be useful when we later want to do calculus. Furthermore  $G_2$  is defined on the whole of  $X$  (and not just on  $\Omega$ ), and it is linear (and not just concave).

Our problem is then to

$$\begin{aligned} & \text{maximize} && M(x) \\ & \text{subject to} && \left\{ \begin{array}{l} x \in \Omega \\ G_1(x) \geq 0 \\ G_2(x) \geq 0 \end{array} \right\}. \end{aligned} \quad (3)$$

#### D. CHARACTERIZING THE OPTIMUM

In our context, the Lagrangean is the mapping  $L : \Omega \times Z^* \times Z^* \rightarrow \mathbb{R}$  given by the formula

$$L(x, \lambda_1, \lambda_2) = M(x) + \langle G_1(x), \lambda_1 \rangle + \langle G_2(x), \lambda_2 \rangle,$$

where  $\langle G_i(x), \lambda_i \rangle$  denotes the real number obtained when the linear functional  $\lambda_i \in Z^*$  is evaluated at the point  $G_i(x) \in Z$ .

In view of our assumptions, the maximum is achieved at  $x_0 \in \Omega$  if and only if there exist  $\lambda_1, \lambda_2 \in Z^*$  such that:

$$1. \ L(x_0, \lambda_1, \lambda_2) \geq L(x, \lambda_1, \lambda_2) \quad \text{for all } x \in \Omega;$$

$$2. \ G_1(x) \geq 0, \lambda_1 \geq 0 \text{ and}$$

$$\langle G_1(x), \lambda_1 \rangle = 0; \quad (4)$$

$$3. \ G_2(x) \geq 0, \lambda_2 \geq 0 \text{ and}$$

$$\langle G_2(x), \lambda_2 \rangle = 0. \quad (5)$$

In other words, there exists multipliers  $\lambda_1$  and  $\lambda_2$  such that:  $x_0$  maximizes  $L(\cdot, \lambda_1, \lambda_2)$  over  $\Omega$ ; complementary slackness holds for the first constraint; and complementary slackness holds for the second constraint.

Since  $P^*$  can be identified with  $\mathcal{M}(\Theta, [0, \infty))$ , we have the following explicit representations of  $M(x)$ ,  $\langle G_1(x), \lambda_1 \rangle$  and  $\langle G_2(x), \lambda_2 \rangle$ :

$$M(x) = \int \left( \theta u(\theta) + w(\theta) \right) dF(\theta), \quad (6)$$

$$\langle G_1(x), \lambda_1 \rangle = \int \left( y - C(u(\theta)) - K(w(\theta)) \right) d\Lambda_1(\theta) \quad (7)$$

and

$$\langle G_2(x), \lambda_2 \rangle = \int \left( b\left(\frac{\theta}{(1+\pi)\beta}\right) u(\theta) - \frac{1}{\rho} a\left(\frac{\theta}{(1+\pi)\beta}\right) - w(\theta) \right) d\Lambda_2(\theta), \quad (8)$$

where  $\Lambda_1$  and  $\Lambda_2$  are the distribution functions of  $\lambda_1$  and  $\lambda_2$  respectively.

**Remark 5.** *In the interests of consistency, all integrals in this Appendix are Lebesgue-Stieltjes integrals, i.e. integrals with respect to functions of bounded variation.*

#### E. THE LAGRANGEAN IS FRÉCHET DIFFERENTIABLE

It is immediate from the formulae (6), (7) and (8) that  $L(x, \lambda_1, \lambda_2)$  is in fact well defined for all  $x \in \Xi = \mathcal{C}(\Theta, \text{ran}(U))^2$ . Let us consider accordingly any  $x_0 = (u_0, w_0) \in \Xi$  and any  $x_1 = (u_1, w_1) \in X$ . Because  $\Xi$  is open,  $x_0 + \varepsilon x_1 \in \Xi$  for all  $\varepsilon > 0$  sufficiently small. Furthermore, it can be verified that the directional derivative  $\nabla_{x_1} L(x_0, \lambda_1, \lambda_2)$  of  $L$  at  $x_0$  in the direction  $x_1$  takes the form

$$\int \left( \theta u_1 + w_1 \right) dF - \int \left( C'(u_0) u_1 + K'(w_0) w_1 \right) d\Lambda_1 + \int \left( b \left( \frac{\theta}{(1+\pi)^\beta} \right) u_1 - w_1 \right) d\Lambda_2. \quad (9)$$

This is easily seen to define a continuous linear functional

$$\nabla L(x_0, \lambda_1, \lambda_2) : x_1 \mapsto \nabla_{x_1} L(x_0, \lambda_1, \lambda_2)$$

on  $X$ . That is,  $L(\cdot, \lambda_1, \lambda_2)$  is Gâteaux differentiable at  $x_0$  with gradient  $\nabla L(x_0, \lambda_1, \lambda_2) \in X^*$ . Finally,  $\nabla L(\cdot, \lambda_1, \lambda_2) : \Xi \rightarrow X^*$  can be shown to be continuous. It follows that  $L(\cdot, \lambda_1, \lambda_2)$  is Fréchet differentiable on  $\Xi$ .

#### F. MAXIMIZING THE LAGRANGEAN

Since  $L(\cdot, \lambda_1, \lambda_2)$  is convex and Fréchet differentiable on  $\Xi$ , and since  $\Omega$  is convex, the maximum of  $L(\cdot, \lambda_1, \lambda_2)$  over  $\Omega$  is achieved at  $x_0 \in \Omega$  iff

$$\nabla_{x-x_0} L(x_0, \lambda_1, \lambda_2) \leq 0$$

for all  $x \in \Omega$ . In this section we shall identify the restrictions that this places on  $\lambda_1$  and  $\lambda_2$ .

To this end, put

$$Y = \left( \mathcal{BV}(\Theta, \mathbb{R}) \cap \mathcal{C}(\Theta, \mathbb{R}) \right) \times \mathbb{R};$$

and consider the affine transformation

$$S : Y \rightarrow \left( \mathcal{BV}(\Theta, \mathbb{R}) \cap \mathcal{C}(\Theta, \mathbb{R}) \right)^2$$

that maps  $y = (u, r)$  to  $x = (u_0 + u, w_0 + w)$ , where  $w$  is the unique solution of the equation  $\theta du + \beta dw = 0$  with boundary condition  $w(\bar{\theta}) = r$ .

For any  $y \in Y$ , we have

$$\begin{aligned}\nabla_{S(y)-x_0} L(x_0, \lambda_1, \lambda_2) &= \int (\theta u + w) dF - \int (C'(u_0) u + K'(w_0) w) d\Lambda_1 \\ &\quad + \int \left( b\left(\frac{\theta}{(1+\pi)\beta}\right) u - w \right) d\Lambda_2 \\ &= \int (\theta u + w) dF - \int \left( \frac{C'(u_0)}{K'(w_0)} u + w \right) d\tilde{\Lambda}_1 \\ &\quad + \int \left( b\left(\frac{\theta}{(1+\pi)\beta}\right) u - w \right) d\Lambda_2\end{aligned}$$

(where  $d\tilde{\Lambda}_1 = K'(w_0) d\Lambda_1$ ). Furthermore, integrating by parts, we have

$$\int w dF = [w F]_{\underline{\theta}-}^{\bar{\theta}} - \int F dw = w(\bar{\theta}) F(\bar{\theta}) + \int F \frac{\theta}{\beta} du$$

(because  $F(\underline{\theta}-) = 0$  and  $dw = -\frac{\theta}{\beta} du$ )

$$= r F(\bar{\theta}) + \int F \frac{\theta}{\beta} du.$$

Moreover

$$\begin{aligned}\int F \theta du &= \int F (\theta du + u d\theta) - \int F u d\theta \\ &= [\theta u F]_{\underline{\theta}-}^{\bar{\theta}} - \int \theta u dF - \int F u d\theta \\ &= \bar{\theta} u(\bar{\theta}) F(\bar{\theta}) - \int \theta u dF - \int F u d\theta\end{aligned}$$

(where we have again used the fact that  $F(\underline{\theta}-) = 0$ ). Hence

$$\int w dF = \left( \frac{\bar{\theta}}{\beta} u(\bar{\theta}) + r \right) F(\bar{\theta}) - \frac{1}{\beta} \int u (\theta dF + F d\theta).$$

Similarly,

$$\int w d\tilde{\Lambda}_1 = \left( \frac{\bar{\theta}}{\beta} u(\bar{\theta}) + r \right) \tilde{\Lambda}_1(\bar{\theta}) - \frac{1}{\beta} \int u (\theta d\tilde{\Lambda}_1 + \tilde{\Lambda}_1 d\theta)$$

and

$$\int w d\Lambda_2 = \left( \frac{\bar{\theta}}{\beta} u(\bar{\theta}) + r \right) \Lambda_2(\bar{\theta}) - \frac{1}{\beta} \int u(\theta) d\Lambda_2 + \Lambda_2 d\theta.$$

Overall, then,

$$\begin{aligned} \nabla_{S(y)-x_0} L(x_0, \lambda_1, \lambda_2) &= \left( \frac{\bar{\theta}}{\beta} u(\bar{\theta}) + r \right) \left( F(\bar{\theta}) - \tilde{\Lambda}_1(\bar{\theta}) - \Lambda_2(\bar{\theta}) \right) \\ &\quad - \frac{1}{\beta} \int u \left( (1 - \beta) \theta dF + F d\theta \right) \\ &\quad + \frac{1}{\beta} \int u \left( \left( \theta - \beta \frac{C'(u_0)}{K'(w_0)} \right) d\tilde{\Lambda}_1 + \tilde{\Lambda}_1 d\theta \right) \\ &\quad + \frac{1}{\beta} \int u \left( \left( \theta + \beta b \left( \frac{\theta}{(1+\pi)\beta} \right) \right) d\Lambda_2 + \Lambda_2 d\theta \right). \end{aligned}$$

Next, it is easy to see that the mapping

$$y \mapsto \nabla_{S(y)-x_0} L(x_0, \lambda_1, \lambda_2)$$

defines a continuous linear functional on  $Y$ . Since it does not depend on the derivatives of  $y$ , this functional extends uniquely to a continuous linear functional

$$y^* : \mathcal{C}(\Theta, \mathbb{R}) \times \mathbb{R} \rightarrow \mathbb{R}.$$

Indeed, we have

$$y^* = (u^*, r^*) \in \mathcal{M}(\Theta, \mathbb{R}) \times \mathbb{R} = (\mathcal{C}(\Theta, \mathbb{R}) \times \mathbb{R})^*,$$

where

$$\begin{aligned} u^* &= -\frac{1}{\beta} \left( (1 - \beta) \theta dF + F d\theta \right) + \frac{1}{\beta} \left( \left( \theta - \beta \frac{C'(u_0)}{K'(w_0)} \right) d\tilde{\Lambda}_1 + \tilde{\Lambda}_1 d\theta \right) \\ &\quad + \frac{1}{\beta} \left( \left( \theta + \beta b \left( \frac{\theta}{(1+\pi)\beta} \right) \right) d\Lambda_2 + \Lambda_2 d\theta \right) + \frac{\bar{\theta}}{\beta} \left( F(\bar{\theta}) - \tilde{\Lambda}_1(\bar{\theta}) - \Lambda_2(\bar{\theta}) \right) dI, \end{aligned}$$

$$r^* = F(\bar{\theta}) - \tilde{\Lambda}_1(\bar{\theta}) - \Lambda_2(\bar{\theta})$$

and  $I$  is the distribution function of the unit mass at  $\bar{\theta}$ .

Finally, it is easy to see that there exists  $\varepsilon > 0$  such that  $S(y) \in \Omega$  for all  $y \in Y \cap B_\varepsilon(0)$ , where  $B_\varepsilon(0)$  is the open ball in  $\mathcal{C}(\Theta, \mathbb{R}) \times \mathbb{R}$  with radius  $\varepsilon$  and centre

0. It follows that  $\langle y, y^* \rangle \leq 0$  for all  $y \in Y \cap B_\varepsilon(0)$ . But  $Y \cap B_\varepsilon(0)$  is dense in  $B_\varepsilon(0)$ . Hence  $\langle y, y^* \rangle \leq 0$  for all  $y \in B_\varepsilon(0)$ . Hence  $y^* = 0$ . In other words, we have

$$0 = -\frac{1}{\beta} \left( (1 - \beta) \theta dF + F d\theta \right) + \frac{1}{\beta} \left( \left( \theta - \beta \frac{C'(u_0)}{K'(w_0)} \right) d\tilde{\Lambda}_1 + \tilde{\Lambda}_1 d\theta \right) \\ + \frac{1}{\beta} \left( \left( \theta + \beta b \left( \frac{\theta}{(1+\pi)\beta} \right) \right) d\Lambda_2 + \Lambda_2 d\theta \right) + \frac{\bar{\theta}}{\beta} \left( F(\bar{\theta}) - \tilde{\Lambda}_1(\bar{\theta}) - \Lambda_2(\bar{\theta}) \right) dI \quad (10)$$

and

$$0 = F(\bar{\theta}) - \tilde{\Lambda}_1(\bar{\theta}) - \Lambda_2(\bar{\theta}). \quad (11)$$

Taking advantage of (11), (10) simplifies to

$$0 = -G d\theta + \left( \left( \theta - \beta \frac{C'(u_0)}{K'(w_0)} \right) d\tilde{\Lambda}_1 + \tilde{\Lambda}_1 d\theta \right) \\ + \left( \left( \theta + \beta b \left( \frac{\theta}{(1+\pi)\beta} \right) \right) d\Lambda_2 + \Lambda_2 d\theta \right), \quad (12)$$

where  $G$  is given by the formula  $G(\theta) = (1 - \beta) \theta F'(\theta) + F(\theta)$ .

#### G. A ONE-PARAMETER FAMILY OF UTILITY CURVES

We shall consider a family of utility curves depending on the single parameter  $\theta_1 \in (0, \bar{\theta}]$ . For each  $\theta_1$ , we begin by finding the point  $(c^*(\theta_1), k^*(\theta_1))$  that would be chosen by a self 1 of type  $\theta_1$  from the ambient budget set  $A$ . The utility curve corresponding to  $\theta_1$  is then the utility curve associated with a two-part budget set with slopes of  $-1$  and  $-(1 + \pi)$  to the left and right of a kink at  $(c^*(\theta_1), k^*(\theta_1))$ . Notice that we specifically allow for the possibility that  $\theta_1 < \underline{\theta}$ .

Put  $\theta_2 = (1 + \pi) \theta_1$ . Knife-edge cases apart, there are then five possible cases arising from the relative positions of the two non-trivial intervals  $[\underline{\theta}, \bar{\theta}]$  and  $[\theta_1, \theta_2]$ :

**Case 1**  $[\underline{\theta}, \bar{\theta}]$  contains  $[\theta_1, \theta_2]$ ;

**Case 2**  $[\theta_1, \theta_2]$  contains  $[\underline{\theta}, \bar{\theta}]$ ;

**Case 3** the two intervals overlap, with  $[\theta_1, \theta_2]$  lying to the left and  $[\underline{\theta}, \bar{\theta}]$  to the right;

**Case 4** the two intervals overlap, with  $[\underline{\theta}, \bar{\theta}]$  lying to the left and  $[\theta_1, \theta_2]$  to the right;

**Case 5**  $[\theta_1, \theta_2]$  lies entirely to the left of  $[\underline{\theta}, \bar{\theta}]$ .

(The case in which  $[\underline{\theta}, \bar{\theta}]$  lies entirely to the left of  $[\theta_1, \theta_2]$  cannot occur, because we are confining  $\theta_1$  to the interval  $(0, \bar{\theta}]$ .)

**G.1. Case 1.** If the utility curve corresponding to  $\theta_1$  is to be an optimum, then the associated multipliers  $\lambda_1$  and  $\lambda_2$  must satisfy the three necessary conditions (4), (5) and (12) (i.e. complementary slackness for the first constraint, complementary slackness for the second constraint and the measure equation). In this section, we show that these three necessary conditions tie down  $\lambda_1$  and  $\lambda_2$  uniquely. The fourth necessary condition, namely the boundary condition (11), is not needed at this stage. (It will be used below to tie down  $\theta_1$ .)

By construction, the consumption-penalty constraint is strictly slack on  $[\underline{\theta}, \theta_2)$  and the budget constraint is strictly slack on  $(\theta_2, \bar{\theta}]$ . Hence  $d\Lambda_2 = 0$  on the former interval and  $d\tilde{\Lambda}_1 = 0$  on the latter. Furthermore (12) implies that

$$(1 - \beta) \theta_2 \Delta F(\theta_2) = (\theta_2 - \theta_1) \Delta \tilde{\Lambda}_1(\theta_2) + \left( \theta_2 + \beta b \left( \frac{\theta_2}{(1+\pi)\beta} \right) \right) \Delta \Lambda_2(\theta_2),$$

where  $\Delta F(\theta_2)$ ,  $\Delta \tilde{\Lambda}_1(\theta_2)$  and  $\Delta \Lambda_2(\theta_2)$  are the atoms of  $dF$ ,  $d\tilde{\Lambda}_1$  and  $d\Lambda_2$  at  $\theta_2$ . But Assumption A1 implies that  $\Delta F(\theta_2) = 0$ . Since all the terms on the right-hand side of the equation are non-negative, it follows that  $\Delta \tilde{\Lambda}_1(\theta_2) = \Delta \Lambda_2(\theta_2) = 0$ . Hence  $d\Lambda_2 = 0$  on  $\{\theta_2\}$  (and therefore on  $[\underline{\theta}, \theta_2]$ ) and  $d\tilde{\Lambda}_1 = 0$  on  $\{\theta_2\}$  (and therefore on  $[\theta_2, \bar{\theta}]$ ).

Now let us consider the three intervals  $[\underline{\theta}, \theta_1]$ ,  $[\theta_1, \theta_2]$  and  $[\theta_2, \bar{\theta}]$  in turn. On  $[\underline{\theta}, \theta_1]$ , we have  $\frac{C'(u_0)}{K'(w_0)} = \frac{\theta}{\beta}$ ,  $\Lambda_2 = 0$  and  $d\Lambda_2 = 0$ . Hence (12) becomes

$$0 = -G d\theta + \tilde{\Lambda}_1 d\theta.$$

It follows that  $\tilde{\Lambda}_1 = G$  almost everywhere with respect to Lebesgue measure  $d\theta$ . Since both  $\tilde{\Lambda}_1$  and  $G$  are functions of bounded variation, it then follows (bearing in mind the convention that functions of bounded variation are right continuous) that  $\tilde{\Lambda}_1 = G$  everywhere on  $[\underline{\theta}, \theta_1)$ , and hence that  $\tilde{\Lambda}_1(\theta_1-) = G(\theta_1-)$ .

On  $[\theta_1, \theta_2]$ , we have  $\frac{C'(u_0)}{K'(w_0)} = \frac{\theta_1}{\beta}$ ,  $\Lambda_2 = 0$  and  $d\Lambda_2 = 0$ . Hence (12) becomes

$$0 = -G d\theta + (\theta - \theta_1) d\tilde{\Lambda}_1 + \tilde{\Lambda}_1 d\theta.$$

This implies that  $\tilde{\Lambda}_1$  takes the form

$$\tilde{\Lambda}_1(\theta) = \frac{1}{\theta - \theta_1} \int_{\theta_1}^{\theta} G(t) dt$$

for all  $\theta \in (\theta_1, \theta_2)$ , that  $\tilde{\Lambda}_1(\theta_1) = G(\theta_1)$  (since  $\tilde{\Lambda}_1$  and  $G$  are right continuous) and that  $\tilde{\Lambda}_1(\theta_2) = \tilde{\Lambda}_1(\theta_2-) = \frac{1}{\theta_2 - \theta_1} \int_{\theta_1}^{\theta_2} G(\theta) d\theta$  (since  $\tilde{\Lambda}_1$  cannot have a jump at  $\theta_2$ ).

On  $[\theta_2, \bar{\theta}]$ , we have  $d\tilde{\Lambda}_1 = 0$ . Hence (12) becomes

$$0 = -G d\theta + \tilde{\Lambda}_1 d\theta + \left( \theta + \beta b\left(\frac{\theta}{(1+\pi)\beta}\right) \right) d\Lambda_2 + \Lambda_2 d\theta.$$

Furthermore, we have the boundary condition  $\Lambda_2(\theta_2) = 0$  (since  $\Lambda_2$  cannot have a jump at  $\theta_2$ ). Putting  $\tilde{\Lambda}_2 = \Lambda_2 + \tilde{\Lambda}_1(\theta_2)$ , this equation simplifies slightly to

$$0 = -G d\theta + \left( \theta + \beta b\left(\frac{\theta}{(1+\pi)\beta}\right) \right) d\tilde{\Lambda}_2 + \tilde{\Lambda}_2 d\theta,$$

with boundary condition  $\tilde{\Lambda}_2(\theta_2) = \tilde{\Lambda}_1(\theta_2)$ .

**G.2. Cases 2 to 5.** In order to handle the remaining cases, we need a unified construction. (This construction includes Case 1 too.) It is more convenient to work in terms of the distribution function  $\Psi = \Psi(\cdot; \theta_1)$  of the total multiplier  $d\tilde{\Lambda}_1 + d\Lambda_2$ , and to view this as a function on  $[0, \bar{\theta}]$ . The construction is then very simple. For all  $\theta_1 \in (0, \bar{\theta}]$ :

1. put  $\Psi = G$  on  $[0, \theta_1]$ ;
2. if  $\theta_1 < \bar{\theta}$  (so that  $\Psi$  is not yet defined on the whole of  $[0, \bar{\theta}]$ ), then let  $\Psi$  be the unique bounded solution of the o.d.e.

$$0 = -G + (\theta - \theta_1) \Psi' + \Psi$$

on  $(\theta_1, \theta_2 \wedge \bar{\theta}]$ , i.e. put

$$\Psi(\theta) = \frac{1}{\theta - \theta_1} \int_{\theta_1}^{\theta} G(t) dt;$$

3. if  $\theta_2 < \bar{\theta}$  (so that  $\Psi$  is still not defined on the whole of  $[0, \bar{\theta}]$ ), then let  $\Psi$  be

the unique solution of the o.d.e.

$$0 = -G + \left( \theta + \beta b\left(\frac{\theta}{(1+\pi)\beta}\right) \right) \Psi' + \Psi$$

on  $(\theta_2, \bar{\theta}]$  with boundary condition

$$\Psi(\theta_2) = \frac{1}{\theta_2 - \theta_1} \int_{\theta_1}^{\theta_2} G(t) dt.$$

Then, using arguments similar to those of the preceding section, it is easy to verify that the three necessary conditions (4), (5) and (12) together imply that, for all  $\theta_1 \in (0, \bar{\theta}]$ ,  $\Psi$  must take the given form.

**Remark 6.** *We can easily extend the construction of  $\Psi$  to include the case  $\theta_1 = 0$ . Indeed, in line with the construction above, we can let  $\Psi(\cdot; 0)$  be the unique bounded solution of the o.d.e.*

$$0 = -G + \left( \theta + \beta b\left(\frac{\theta}{(1+\pi)\beta}\right) \right) \Psi' + \Psi$$

on  $(0, \bar{\theta}]$ .

**Remark 7.** *With this definition of  $\Psi(\cdot; 0)$ ,  $\Psi(\cdot; \theta_1)$  is independent of  $\theta_1$  for  $\theta_1 \in [0, \frac{1}{1+\pi} \underline{\theta}]$ .*

## H. EXISTENCE OF AN OPTIMUM

For all  $\theta_1 \in (0, \bar{\theta}]$ , we have shown that there exists a unique  $\Psi = \Psi(\cdot; \theta_1)$  satisfying the two complementary slackness conditions (4) and (5) and the measure equation (12). This  $\Psi(\cdot; \theta_1)$  does not in general satisfy the boundary condition (11). The purpose of the current section is to establish that there is at least one choice of  $\theta_1$  for which (11) is satisfied.

For all  $\theta_1 \in (0, \bar{\theta}]$ , put  $\psi(\theta_1) = \Psi(\bar{\theta}; \theta_1)$ . Now consider  $G$ . We certainly have  $G = 0$  on  $(0, \underline{\theta})$  and  $G = F(\bar{\theta})$  on  $[\bar{\theta}, \infty)$ . Furthermore Assumption A3 tells us that  $G$  is first increasing (on  $(0, \theta_M)$ ) and then decreasing (on  $(\theta_M, \infty)$ ). It is therefore obvious that there exists  $\underline{\theta}_F \in [\underline{\theta}, \bar{\theta}]$  such that  $G < F(\bar{\theta})$  on  $(0, \underline{\theta}_F)$  and  $G \geq F(\bar{\theta})$  on  $[\underline{\theta}_F, \bar{\theta}]$ . Our first lemma sharpens this observation.

**Lemma 8.** *There exists  $\bar{\theta}_F \in [\underline{\theta}, \bar{\theta})$  such that  $G \leq F(\bar{\theta})$  on  $(0, \bar{\theta}_F)$  and  $G > F(\bar{\theta})$  on  $(\bar{\theta}_F, \bar{\theta})$ .*

**Proof.** Suppose first that  $\theta_M < \bar{\theta}$ . Suppose further that there exists  $\xi_0 \in (\theta_M, \bar{\theta})$  such that  $G(\xi_0) = G(\bar{\theta})$ . Since  $G' \leq 0$  on  $(\theta_M, \infty) \supset (\xi_0, \bar{\theta}]$ , it follows that  $G' = 0$  on  $(\xi_0, \bar{\theta}]$ . We also know that  $G = F(\bar{\theta})$  on  $[\bar{\theta}, \infty)$ , and therefore that  $G' = 0$  on  $(\bar{\theta}, \infty)$ . Overall, then,  $G' = 0$  on  $(\xi_0, \infty)$ . Hence  $\theta^\alpha F'(\theta)$  is constant on  $(\xi_0, \infty)$ , where  $\alpha = \frac{2-\beta}{1-\beta}$ . (Cf. the proof of Proposition 1.) Hence  $F' = 0$  on  $(\xi_0, \infty)$ . But this contradicts the fact that  $\bar{\theta}$  is the maximum of the support of  $F$ . We may therefore conclude that  $G > G(\bar{\theta}) = F(\bar{\theta})$  on  $(\theta_M, \bar{\theta})$ .

Suppose second that  $\theta_M = \bar{\theta}$ . Then  $G' \geq 0$  on  $(0, \bar{\theta})$ . Hence  $\theta^\alpha F'(\theta)$  is non-decreasing on  $(0, \bar{\theta})$ . In particular, if we fix  $\xi_1 \in (\underline{\theta}, \bar{\theta})$ , then we will have  $\theta^\alpha F'(\theta) \geq \xi_1^\alpha F'(\xi_1)$  for all  $\theta \in (\xi_1, \bar{\theta})$ . Letting  $\theta \uparrow \bar{\theta}$ , it follows that  $\bar{\theta}^\alpha F'(\bar{\theta}-) \geq \xi_1^\alpha F'(\xi_1)$ . But  $F'(\xi_1) > 0$ , since  $\xi_1$  lies in the interior of the support of  $F$ . Hence  $F'(\bar{\theta}-) > 0$ . Hence  $G(\bar{\theta}-) = (1-\beta)\bar{\theta}F'(\bar{\theta}-) + F(\bar{\theta}) > F(\bar{\theta})$ . Hence there exists  $\varepsilon > 0$  such that  $G > F(\bar{\theta})$  on  $(\bar{\theta}-\varepsilon, \bar{\theta})$ .

Overall, then, we have the following picture:  $G = 0$  on  $(0, \underline{\theta})$ ; there exists  $\xi_2 \in [\underline{\theta}, \bar{\theta})$  such that  $G > F(\bar{\theta})$  on  $(\xi_2, \bar{\theta})$ ; and  $G = F(\bar{\theta})$  on  $[\bar{\theta}, \infty)$ . Furthermore Assumption A3 ensures that  $\{\theta \mid G(\theta) > F(\bar{\theta})\}$  is an interval. We may therefore put  $\bar{\theta}_F = \inf \{\theta \mid \theta \in [\underline{\theta}, \bar{\theta}), G(\theta) > F(\bar{\theta})\}$ . ■

Now, it follows from the construction of  $\Psi$  given in Section G.2 that  $\Psi(\bar{\theta}; \theta_1)$  is a convex combination of the values of  $G$  on the interval  $(\theta_1, \bar{\theta})$ . Combining this observation with Lemma 8, we obtain:

**Lemma 9.**  $\psi > F(\bar{\theta})$  on  $[\bar{\theta}_F, \bar{\theta})$ . ■

We also have:

**Lemma 10.**  $\psi < F(\bar{\theta})$  on  $(0, \frac{1}{1+\pi}\underline{\theta}]$ .

**Proof.** We begin by noting that  $\Psi = \Psi(\cdot; \theta_1)$  is independent of  $\theta_1$  for  $\theta_1 \in [0, \frac{1}{1+\pi}\underline{\theta}]$ . It is therefore the unique bounded solution of the o.d.e.

$$0 = -G + \left( \theta + \beta b\left(\frac{\theta}{(1+\pi)\beta}\right) \right) \Psi' + \Psi \quad (13)$$

on  $(0, \bar{\theta}]$ . We compare  $\Psi$  with the function  $\Phi$  which is the unique bounded solution of the o.d.e.

$$0 = -G + \theta \Phi' + \Phi \quad (14)$$

on  $(0, \bar{\theta}]$ . Now

$$\Phi(\theta) = \frac{1}{\theta} \int_0^\theta G(t) dt = (1 - \beta) F(\theta) + \beta \frac{1}{\theta} \int_0^\theta F(t) dt.$$

Hence:  $\Phi = 0$  on  $(0, \bar{\theta}]$ ; and  $0 < \Phi < F$  on  $(\underline{\theta}, \bar{\theta}]$ . Hence

$$\Phi' = \frac{G - \Phi}{\theta} \geq \frac{F - \Phi}{\theta} \geq 0$$

on  $(0, \bar{\theta}]$ , with strict inequality on  $(\underline{\theta}, \bar{\theta}]$ . Furthermore,  $\Phi$  is a supersolution of the equation for  $\Psi$ . Indeed, we have

$$\begin{aligned} & -G + \left( \beta b \left( \frac{\theta}{(1+\pi)\beta} \right) + \theta \right) \Phi' + \Phi \\ = & -G + \beta b \left( \frac{\theta}{(1+\pi)\beta} \right) \Phi' + \theta \Phi' + \Phi \end{aligned}$$

(on rearranging)

$$= \beta b \left( \frac{\theta}{(1+\pi)\beta} \right) \Phi'$$

(since  $\Phi$  satisfies (14))

$$\geq 0$$

on  $(0, \bar{\theta}]$ , with strict inequality on  $(\underline{\theta}, \bar{\theta}]$ . That is,  $\Phi$  is a supersolution of the equation for  $\Psi$ , and a strict supersolution on  $(\underline{\theta}, \bar{\theta}]$ . Hence  $\Phi > \Psi$  on  $(\underline{\theta}, \bar{\theta}]$ . In particular,  $\psi(0) = \Psi(\bar{\theta}) < \Phi(\bar{\theta}) < F(\bar{\theta})$ . The general case now follows on noting that  $\Psi(\cdot; \theta_1) = \Psi(\cdot; 0)$  for all  $\theta_1 \in (0, \frac{1}{1+\pi}\underline{\theta}]$ . Cf. the remark at the end of Section G.2. ■

Since  $\psi$  is continuous, we can combine Lemmas 9 and 10 to obtain:

**Proposition 11.** *There exists  $\theta_1 \in (\frac{1}{1+\pi}\underline{\theta}, \bar{\theta}_F)$  such that  $\psi(\theta_1) = F(\bar{\theta})$ . ■*

That is, there exists  $\theta_1 \in (\frac{1}{1+\pi}\underline{\theta}, \bar{\theta}_F)$  such that equation (11) is satisfied. However, we still need to verify that the multipliers associated with any such  $\theta_1$  are non-negative.

## I. NON-NEGATIVITY OF THE MULTIPLIER

In this section we show that, if  $\theta_1 \in (0, \bar{\theta}_F)$  is such that  $\psi(\theta_1) \leq F(\bar{\theta})$ , then  $\Psi = \Psi(\cdot; \theta_1)$  is non-decreasing on  $[0, \bar{\theta}]$ . We treat the intervals  $[0, \theta_1]$ ,  $(\theta_1, \bar{\theta}_F)$  and  $[\bar{\theta}_F, \bar{\theta}]$  in turn. We begin with a simple lemma.

**Lemma 12.**  $\bar{\theta}_F \leq \theta_M$ .

**Proof.** In the light of Lemma 8,  $\sup G > F(\bar{\theta})$ . Moreover it follows from the definition of  $\theta_M$  that  $\sup G = \max \{G_L(\theta_M), G(\theta_M)\}$ . There are therefore two possibilities. Either  $G_L(\theta_M) > F(\bar{\theta})$ , in which case there is a left neighbourhood of  $\theta_M$  on which  $G > F(\bar{\theta})$ , and therefore  $\bar{\theta}_F < \theta_M$ . Or  $G(\theta_M) > F(\bar{\theta})$ , in which case necessarily  $\theta_M \geq \bar{\theta}_F$ . ■

**Lemma 13.**  $G' \geq 0$  on  $[0, \bar{\theta}_F]$ .

**Proof.** From Lemma 12 we know that  $[0, \bar{\theta}_F) \subset [0, \theta_M)$ . Hence  $G' \geq 0$  on  $[0, \bar{\theta}_F)$ . However,  $G \leq F(\bar{\theta})$  on  $[0, \bar{\theta}_F)$  and  $G > F(\bar{\theta})$  on  $(\bar{\theta}_F, \bar{\theta})$ . Hence  $\Delta G(\bar{\theta}_F) \geq 0$ . Hence  $G' \geq 0$  on  $[0, \bar{\theta}_F]$ . ■

**Proposition 14.** Suppose that  $\psi(\theta_1) \leq F(\bar{\theta})$ . Then  $\Psi' \geq 0$  on  $[0, \theta_1]$ .

**Proof.** Since  $\psi(\theta_1) \leq F(\bar{\theta})$ , Lemma 9 implies that  $\theta_1 < \bar{\theta}_F$ . Hence  $[0, \theta_1] \subset [0, \bar{\theta}_F]$ . But  $\Psi = G$  on  $[0, \theta_1]$  by construction of  $\Psi$ , and Lemma 13 tells us that  $G' \geq 0$  on  $[0, \bar{\theta}_F]$ . It follows that  $\Psi' \geq 0$  on  $[0, \theta_1]$ . ■

In order to show that  $\Psi' \geq 0$  on  $(\theta_1, \bar{\theta}]$ , we use the fact that  $\Psi$  solves

$$0 = -G + (\theta - \theta_1) \Psi' + \Psi \quad (15)$$

on  $(\theta_1, \theta_2 \wedge \bar{\theta}]$  and

$$0 = -G + \left( \theta + \beta b \left( \frac{\theta}{(1+\pi)\beta} \right) \right) \Psi' + \Psi \quad (16)$$

on  $(\theta_2 \wedge \bar{\theta}, \bar{\theta}]$ . We also make use of the corresponding o.d.e. for  $\theta$ , namely

$$\dot{\theta} = -(\theta - \theta_1) \quad (17)$$

on  $(\theta_1, \theta_2 \wedge \bar{\theta}]$  and

$$\dot{\theta} = - \left( \theta + \beta b \left( \frac{\theta}{(1+\pi)\beta} \right) \right) \quad (18)$$

on  $(\theta_2 \wedge \bar{\theta}, \bar{\theta}]$ . Specifically, for all  $g, h \in (\theta_1, \bar{\theta}]$  such that  $h < g$ , let  $T(h, g)$  denote the time at which the solution of the o.d.e. (17-18) for  $\theta$  starting from  $g$  hits  $h$ , and put  $S(h, g) = \exp(-T(h, g))$ . Notice that  $S(\cdot, g)$  increases from 0 at  $\theta_1$  to 1 at  $g$ .

**Lemma 15.** Suppose that  $\psi(\theta_1) \leq F(\bar{\theta})$ . Then  $\Psi \leq G$  on  $(\theta_1, \theta_M)$ .

**Proof.** Since  $\psi(\theta_1) \leq F(\bar{\theta})$ , Lemma 9 implies that  $\theta_1 < \bar{\theta}_F$ . Furthermore Lemma 12 tells us that  $\bar{\theta}_F \leq \theta_M$ . For all  $g \in (\theta_1, \theta_M)$ , we therefore have

$$\Psi(g) = \int_{\theta_1}^g \frac{\partial S}{\partial h}(h, g) G(h) dh \leq \int_{\theta_1}^g \frac{\partial S}{\partial h}(h, g) G(g-) dh$$

(with equality iff  $G(g-) = G(\theta_1)$ )

$$= G(g-).$$

Taking limits from the right (and using the continuity of  $\Psi$  and the right continuity of  $G$ ) then yields  $\Psi \leq G$ . ■

**Lemma 16.** *The sign of  $\Psi'$  coincides with that of  $G - \Psi$  on  $(\theta_1, \bar{\theta}]$ .*

**Proof.** We have

$$\Psi' = \frac{G - \Psi}{\theta - \theta_1}$$

on  $(\theta_1, \theta_2 \wedge \bar{\theta}]$  (from equation (15)) and

$$\Psi' = \frac{G - \Psi}{\theta + \beta b\left(\frac{\theta}{(1+\pi)\beta}\right)}$$

on  $(\theta_2 \wedge \bar{\theta}, \bar{\theta}]$  (from equation (16)). Bearing in mind that we have  $\theta - \theta_1 > 0$  on  $(\theta_1, \theta_2 \wedge \bar{\theta}]$ , it follows that the sign of  $\Psi'$  coincides with that of  $G - \Psi$  on  $(\theta_1, \theta_2 \wedge \bar{\theta}] \cup (\theta_2 \wedge \bar{\theta}, \bar{\theta}] = (\theta_1, \bar{\theta}]$ . ■

**Proposition 17.** *Suppose that  $\psi(\theta_1) \leq F(\bar{\theta})$ . Then  $\Psi' \geq 0$  on  $(\theta_1, \bar{\theta}_F)$ .*

**Proof.** From Lemma 15 we know that  $\Psi \leq G$  on  $(\theta_1, \theta_M)$  and therefore on  $(\theta_1, \bar{\theta}_F) \subset (\theta_1, \theta_M)$ . Lemma 16 then implies that  $\Psi' \geq 0$  there. ■

**Proposition 18.** *Suppose that  $\psi(\theta_1) \leq F(\bar{\theta})$ . Then  $\Psi' > 0$  on  $[\bar{\theta}_F, \bar{\theta})$ .*

**Proof.** For all  $\theta \in [\bar{\theta}_F, \bar{\theta})$ , we have

$$\Psi(\bar{\theta}) = S(\theta, \bar{\theta}) \Psi(\theta) + \int_{\theta}^{\bar{\theta}} \frac{\partial S}{\partial h}(h, \bar{\theta}) G(h) dh. \quad (19)$$

Since

$$\int_{\theta}^{\bar{\theta}} \frac{\partial S}{\partial h}(h, \bar{\theta}) dh = 1 - S(\theta, \bar{\theta}),$$

this means that  $\Psi(\bar{\theta})$  is a convex combination of  $\Psi(\theta)$  and the values of  $G$  on  $(\theta, \bar{\theta}]$ . But  $\Psi(\bar{\theta}) = \psi(\theta_1) \leq F(\bar{\theta})$  and  $G > F(\bar{\theta})$  on  $(\theta, \bar{\theta})$ . So we must have  $\Psi(\theta) < F(\bar{\theta})$ . On the other hand,  $G(\theta) \geq F(\bar{\theta})$  since  $\theta \geq \bar{\theta}_F$ . Lemma 16 therefore implies that  $\Psi'(\theta) > 0$ . ■

## J. UNIQUENESS OF THE OPTIMUM

At this point we have established that the utility curve associated with  $\theta_1$  solves the main maximization problem (3) iff  $\psi(\theta_1) = F(\bar{\theta})$ . Furthermore  $\psi < F(\bar{\theta})$  on  $[0, \frac{1}{1+\pi}\underline{\theta}]$  and  $\psi > F(\bar{\theta})$  on  $[\bar{\theta}_F, \bar{\theta})$ . Hence there exists  $\theta_1 \in (\frac{1}{1+\pi}\underline{\theta}, \bar{\theta}_F)$  such that  $\psi(\theta_1) = F(\bar{\theta})$ . In the current section, we show that the set of  $\theta_1$  for which  $\psi(\theta_1) = F(\bar{\theta})$  is an interval. Furthermore, if we strengthen Assumption A3 by requiring that  $G$  is strictly increasing to the left of its peak, then  $\psi' > 0$  over the relevant range. It then follows that there is a unique  $\theta_1$  for which  $\psi(\theta_1) = F(\bar{\theta})$ . This result is limited: it shows that – under the strengthened version of A3 – there is exactly one solution to problem (3) within our one-parameter family of utility curves; but it does not show that there is exactly one solution to problem (3) in  $\Omega$ . It is, however, very suggestive.

The main idea of the proof is to find an explicit formula for  $\psi'$ , and then use this formula to determine the sign of  $\psi'$ . Of course, the formula depends on whether  $\theta_1 < \frac{1}{1+\pi}\bar{\theta}$  or  $\theta_1 > \frac{1}{1+\pi}\bar{\theta}$ . In the former case:  $\theta_2 = (1 + \pi)\theta_1 < \bar{\theta}$ ; the consumption-penalty constraint is strictly binding; and the types in the range  $[\theta_2, \bar{\theta}]$  will choose to incur the consumption penalty. In the latter case:  $\theta_2 = (1 + \pi)\theta_1 > \bar{\theta}$ ; the consumption-penalty constraint is strictly slack; and no type will choose to incur the consumption penalty.

In order to state the formula for  $\psi'$ , it will be helpful to introduce the functions  $\phi$ ,  $\chi$ ,  $\zeta$  and  $\eta$  given by the formulae

$$\phi(\theta_1) = \frac{1}{\theta_2 - \theta_1} \int_{\theta_1}^{\theta_2} G(\theta) d\theta \tag{20}$$

for all  $\theta_1 \in (0, \infty)$  (where we have suppressed the dependence of  $\theta_2$  on  $\theta_1$ ),

$$\chi(\theta_1) = \frac{1}{\bar{\theta} - \theta_1} \int_{\theta_1}^{\bar{\theta}} G(\theta) d\theta \quad (21)$$

for all  $\theta_1 \in (0, \bar{\theta})$ ,

$$\zeta(\theta_1) = \frac{\frac{\beta}{1+\pi} b\left(\frac{\theta_1}{\beta}\right)}{\theta_1 \left(\theta_1 + \frac{\beta}{1+\pi} b\left(\frac{\theta_1}{\beta}\right)\right)} (G(\theta_2) - \phi(\theta_1)) + \frac{1}{\pi \theta_1} (G(\theta_2) - G(\theta_1)) \quad (22)$$

for all  $\theta_1 \in (0, \infty)$  (where we have suppressed the dependence of  $\theta_2$  on  $\theta_1$ ) and

$$\eta(\theta_1) = \frac{\chi(\theta_1) - G(\theta_1)}{\bar{\theta} - \theta_1} \quad (23)$$

for all  $\theta_1 \in (0, \bar{\theta})$ .

**Lemma 19.**  $\psi'(\theta_1) = S(\theta_2, \bar{\theta}) \zeta(\theta_1)$  for  $\theta_1 \in (0, \frac{1}{1+\pi} \bar{\theta})$ .

**Proof.** Equation (19) gives

$$\psi(\theta_1) = \int_{\theta_2}^{\bar{\theta}} \frac{\partial S}{\partial h}(h, \bar{\theta}) G(h) dh + S(\theta_2, \bar{\theta}) \phi(\theta_1),$$

where we have used the fact that  $\Psi(\bar{\theta}; \theta_1) = \psi(\theta_1)$  and  $\Psi(\theta_2; \theta_1) = \phi(\theta_1)$ . Hence

$$\begin{aligned} \psi' &= -\frac{\partial S}{\partial h}(\theta_2, \bar{\theta}) G(\theta_2) \theta_2' + \frac{\partial S}{\partial h}(\theta_2, \bar{\theta}) \theta_2' \phi + S(\theta_2, \bar{\theta}) \phi' \\ &= \exp(-T(\theta_2, \bar{\theta})) \left( \frac{\partial T}{\partial h}(\theta_2, \bar{\theta}) \theta_2' (G(\theta_2) - \phi) + \phi' \right), \end{aligned}$$

where we have suppressed the dependence of  $\phi$  and  $\psi$  on  $\theta_1$ , and where  $\theta_2'$  and  $\phi'$  denote the derivatives of  $\theta_2$  and  $\phi$  with respect to  $\theta_1$ . Furthermore

$$\phi = \frac{1}{\theta_2 - \theta_1} \int_{\theta_1}^{\theta_2} G(t) dt$$

and

$$T(\theta_2, \bar{\theta}) = \int_{\theta_2}^{\bar{\theta}} \frac{dt}{t + \beta b\left(\frac{t}{(1+\pi)\beta}\right)}.$$

Hence

$$\begin{aligned}\phi' &= -\frac{\theta_2' - 1}{(\theta_2 - \theta_1)^2} \int_{\theta_1}^{\theta_2} G(t) dt + \frac{1}{\theta_2 - \theta_1} (G(\theta_2) \theta_2' - G(\theta_1)) \\ &= \frac{1}{\theta_2 - \theta_1} (-\pi \phi + ((1 + \pi) G(\theta_2) - G(\theta_1))) \\ &= \frac{1}{\theta_1} (G(\theta_2) - \phi) + \frac{1}{\pi \theta_1} (G(\theta_2) - G(\theta_1))\end{aligned}$$

and

$$\frac{\partial T}{\partial h}(\theta_2, \bar{\theta}) = -\frac{1}{\theta_2 + \beta b\left(\frac{\theta_2}{(1+\pi)\beta}\right)}. \quad (24)$$

Overall, then,

$$\begin{aligned}\exp(T(\theta_2, \bar{\theta})) \psi' &= \frac{\partial T}{\partial h}(\theta_2, \bar{\theta}) \theta_2' (G(\theta_2) - \phi) + \phi' \\ &= -\frac{1 + \pi}{\theta_2 + \beta b\left(\frac{\theta_2}{(1+\pi)\beta}\right)} (G(\theta_2) - \phi) + \frac{1}{\theta_1} (G(\theta_2) - \phi) + \frac{1}{\pi \theta_1} (G(\theta_2) - G(\theta_1)) \\ &= \left( \frac{1}{\theta_1} - \frac{1}{\theta_1 + \frac{\beta}{1+\pi} b\left(\frac{\theta_1}{\beta}\right)} \right) (G(\theta_2) - \phi) + \frac{1}{\pi \theta_1} (G(\theta_2) - G(\theta_1))\end{aligned}$$

(collecting terms in  $(G(\theta_2) - \phi)$  and  $(G(\theta_2) - G(\theta_1))$ , and using the fact that  $\theta_2 = (1 + \pi) \theta_1$ )

$$= \frac{\frac{\beta}{1+\pi} b\left(\frac{\theta_1}{\beta}\right)}{\theta_1 \left( \theta_1 + \frac{\beta}{1+\pi} b\left(\frac{\theta_1}{\beta}\right) \right)} (G(\theta_2) - \phi) + \frac{1}{\pi \theta_1} (G(\theta_2) - G(\theta_1)).$$

Making  $\psi'$  the subject of this equation, we obtain the required result. ■

The second of the two formulae for  $\psi'$  is given by the following lemma.

**Lemma 20.**  $\psi' = \eta$  on  $\left[\frac{1}{1+\pi} \bar{\theta}, \bar{\theta}\right)$ .

**Proof.** We have  $\psi = \chi$  on  $\left[\frac{1}{1+\pi} \bar{\theta}, \bar{\theta}\right)$ . Moreover it is easy to check that

$$\chi'(\theta_1) = \frac{\chi(\theta_1) - G(\theta_1)}{\bar{\theta} - \theta_1}$$

on  $(0, \bar{\theta})$ . ■

There are now two main cases to consider. The more general of the two main cases occurs when  $\frac{1}{1+\pi} \bar{\theta} < \bar{\theta}_F$ . In this case, there are three main subcases to consider:

**Subcase 1**  $\theta_1 \in (\frac{1}{1+\pi} \underline{\theta}, \frac{1}{1+\pi} \bar{\theta}_F]$ . In this subcase, the consumption-penalty constraint is strictly binding in the sense that  $\theta_2 < \bar{\theta}$ . I.e. all the types in the non-trivial range of  $[\theta_2, \bar{\theta}]$  choose to make an early withdrawal from the penalty account.

**Subcase 2**  $\theta_1 \in [\frac{1}{1+\pi} \bar{\theta}_F, \frac{1}{1+\pi} \bar{\theta}]$ . In this subcase, the consumption-penalty constraint is weakly binding in the sense that  $\theta_2 \leq \bar{\theta}$ .

**Subcase 3**  $\theta_1 \in [\frac{1}{1+\pi} \bar{\theta}, \bar{\theta}_F)$ . In this subcase, the consumption-penalty constraint is weakly slack in the sense that  $\theta_2 \geq \bar{\theta}$ .

The less general of the two main cases occurs when  $\frac{1}{1+\pi} \bar{\theta} \geq \bar{\theta}_F$ . In this case, the third subcase does not arise.

We deal with both of the two main cases simultaneously. The first subcase is settled by the following lemma.

**Lemma 21.** *Suppose that  $\theta_1 \in (0, \frac{1}{1+\pi} \bar{\theta}_F]$ . Then  $\psi'(\theta_1) \geq 0$ .*

**Proof.** The proof relies on the formula  $\psi'(\theta_1) = S(\theta_2, \bar{\theta}) \zeta(\theta_1)$  given in Lemma 19. This formula is valid for  $\theta_1 \in (0, \frac{1}{1+\pi} \bar{\theta}) \supset (0, \frac{1}{1+\pi} \bar{\theta}_F]$ .

We have  $[\theta_1, \theta_2] \subset (0, \bar{\theta}_F]$  and therefore  $G' \geq 0$  on  $[\theta_1, \theta_2]$ . Hence  $G(\theta_2) \geq \phi(\theta_1)$  (with equality iff  $G_L(\theta_2) = G(\theta_1)$ ) and  $G(\theta_2) \geq G(\theta_1)$ . It then follows from formula (22) that  $\zeta(\theta_1) \geq 0$  (with equality iff  $G(\theta_2) = G(\theta_1)$ ), and thence that  $\psi'(\theta_1) \geq 0$  (with equality iff  $G(\theta_2) = G(\theta_1)$ ). ■

We now turn to the second subcase.

**Lemma 22.** *Suppose that  $\theta_1 \in [\frac{1}{1+\pi} \bar{\theta}_F, \frac{1}{1+\pi} \bar{\theta}]$  and that  $\psi(\theta_1) \leq F(\bar{\theta})$ . Then  $G(\theta_2) - \phi(\theta_1) \geq 0$ .*

**Proof.** We have  $\theta_2 \in [\bar{\theta}_F, \bar{\theta}]$  and therefore

$$G(\theta_2) \geq F(\bar{\theta})$$

(with strict inequality if  $\theta_1 \in (\frac{1}{1+\pi} \bar{\theta}_F, \frac{1}{1+\pi} \bar{\theta})$ , because then  $\theta_2 \in (\bar{\theta}_F, \bar{\theta})$ )

$$\geq \psi(\theta_1) = \Psi(\bar{\theta}; \theta_1)$$

(by assumption and by definition of  $\psi$  respectively)

$$\geq \Psi(\theta_2; \theta_1)$$

(with strict inequality if  $\theta_1 \in [\frac{1}{1+\pi}\bar{\theta}_F, \frac{1}{1+\pi}\bar{\theta}]$ , because then  $\theta_2 < \bar{\theta}$ )

$$= \phi(\theta_1)$$

(by definition of  $\phi$ ). ■

**Lemma 23.** *Suppose that  $\theta_1 \in [\frac{1}{1+\pi}\bar{\theta}_F, \frac{1}{1+\pi}\bar{\theta}]$  and that  $\psi(\theta_1) \leq F(\bar{\theta})$ . Then  $G(\theta_2) - G(\theta_1) > 0$ .*

**Proof.** We have

$$G(\theta_2) \geq F(\bar{\theta})$$

(with strict inequality if  $\theta_1 \in (\frac{1}{1+\pi}\bar{\theta}_F, \frac{1}{1+\pi}\bar{\theta})$ )

$$\geq \psi(\theta_1) = \Psi(\bar{\theta}; \theta_1) > \Psi(\theta_1; \theta_1) = G(\theta_1)$$

(by assumption, by definition of  $\psi$ , because  $\theta_1 < \bar{\theta}$  and by construction of  $\Psi$  respectively). ■

Combining Lemmas 22 and 23, we obtain the following result about the right-hand derivative of  $\psi$ .

**Lemma 24.** *Suppose that  $\theta_1 \in [\frac{1}{1+\pi}\bar{\theta}_F, \frac{1}{1+\pi}\bar{\theta})$  and that  $\psi(\theta_1) \leq F(\bar{\theta})$ . Then  $\psi'(\theta_1) > 0$ .*

**Proof.** The proof again relies on the formula  $\psi'(\theta_1) = S(\theta_2, \bar{\theta})\zeta(\theta_1)$  given in Lemma 19. In view of this formula,  $\psi'(\theta_1) > 0$  if  $G(\theta_2) \geq \phi(\theta_1)$  and  $G(\theta_2) > G(\theta_1)$  with at least one strict inequality. But Lemmas 22 and 23 show that  $G(\theta_2) \geq \phi(\theta_1)$  and  $G(\theta_2) > G(\theta_1)$  respectively. ■

We also need the corresponding result about the left-hand derivative of  $\psi$ .

**Lemma 25.** *Suppose that  $\theta_1 \in (\frac{1}{1+\pi}\bar{\theta}_F, \frac{1}{1+\pi}\bar{\theta}]$  and that  $\psi(\theta_1) \leq F(\bar{\theta})$ . Then  $\psi'_L(\theta_1) > 0$ .*

**Proof.** The proof parallels that of Lemma 24, with minor changes. First of all, bearing in mind that  $\phi$  is continuous, we have

$$\zeta_L(\theta_1) = \frac{\frac{\beta}{1+\pi} b\left(\frac{\theta_1}{\beta}\right)}{\theta_1 \left(\theta_1 + \frac{\beta}{1+\pi} b\left(\frac{\theta_1}{\beta}\right)\right)} (G_L(\theta_2) - \phi(\theta_1)) + \frac{1}{\pi \theta_1} (G_L(\theta_2) - G_L(\theta_1)) \quad (25)$$

for all  $\theta_1 \in (0, \infty)$ . As in Lemma 19, we then have  $\psi'_L(\theta_1) = S(\theta_2, \bar{\theta}) \zeta_L(\theta_1)$  for  $\theta_1 \in (0, \frac{1}{1+\pi} \bar{\theta}]$ . Next, just as the single peakedness of  $G$  implies that  $G > F(\bar{\theta})$  on  $(\bar{\theta}_F, \bar{\theta})$ , so it also implies that  $G_L > F(\bar{\theta})$  on  $(\bar{\theta}_F, \bar{\theta})$ . Arguing as in Lemma 22, we can therefore show that  $G_L(\theta_2) - \phi(\theta_1) \geq 0$  for  $\theta_1 \in (\frac{1}{1+\pi} \bar{\theta}_F, \frac{1}{1+\pi} \bar{\theta}]$ , with strict inequality if  $\theta_1 < \frac{1}{1+\pi} \bar{\theta}$ . (We cannot however extend this to  $\theta_1 \in [\frac{1}{1+\pi} \bar{\theta}_F, \frac{1}{1+\pi} \bar{\theta}]$ , since we cannot deduce from the fact that  $G_L > F(\bar{\theta})$  on  $(\bar{\theta}_F, \bar{\theta})$  that  $G_L \geq F(\bar{\theta})$  at  $\bar{\theta}_F$ .) Next, as in Lemma 23, we have  $G_L(\theta_2) - G_L(\theta_1) > 0$  on  $(\frac{1}{1+\pi} \bar{\theta}_F, \frac{1}{1+\pi} \bar{\theta}]$ . Indeed, as in the proof of that lemma, we can show that

$$G_L(\theta_2) \geq F(\bar{\theta})$$

(with strict inequality if  $\theta_1 \in (\frac{1}{1+\pi} \bar{\theta}_F, \frac{1}{1+\pi} \bar{\theta})$ )

$$\geq \psi(\theta_1) = \Psi(\bar{\theta}; \theta_1) > \Psi(\theta_1; \theta_1) = G(\theta_1).$$

In particular,  $G(\theta_1) < F(\bar{\theta})$ ; and therefore  $\theta_1 < \bar{\theta}_F \leq \theta_M$ ; and therefore  $G_L(\theta_1) \leq G(\theta_1)$ . Finally, applying (25) yields the required result. ■

Next, we prove a lemma that will be needed for the third subcase.

**Lemma 26.**  $\chi > G$  on  $[0, \bar{\theta}_F)$ .

**Proof.** For all  $\theta_1 \in [0, \bar{\theta}_F)$ , we have

$$\chi(\theta_1) = \frac{1}{\bar{\theta} - \theta_1} \int_{\theta_1}^{\bar{\theta}} G(\theta) d\theta = \frac{1}{\bar{\theta} - \theta_1} \int_{\theta_1}^{\bar{\theta}_F} G(\theta) d\theta + \frac{1}{\bar{\theta} - \theta_1} \int_{\bar{\theta}_F}^{\bar{\theta}} G(\theta) d\theta.$$

Moreover:

$$\int_{\theta_1}^{\bar{\theta}_F} G(\theta) d\theta \geq (\bar{\theta}_F - \theta_1) G(\theta_1),$$

since  $G' \geq 0$  on  $[0, \bar{\theta}_F]$  by Lemma 13; and

$$\int_{\bar{\theta}_F}^{\bar{\theta}} G(\theta) d\theta > (\bar{\theta} - \bar{\theta}_F) F(\bar{\theta}) \geq (\bar{\theta} - \bar{\theta}_F) G(\theta_1),$$

since  $G > F(\bar{\theta})$  on  $(\bar{\theta}_F, \bar{\theta})$  and (by Lemma 8)  $G \leq F(\bar{\theta})$  on  $(0, \bar{\theta}_F)$ . Hence

$$\chi(\theta_1) > \frac{\bar{\theta}_F - \theta_1}{\bar{\theta} - \theta_1} G(\theta_1) + \frac{\bar{\theta} - \bar{\theta}_F}{\bar{\theta} - \theta_1} G(\theta_1) = G(\theta_1),$$

as required. ■

We can now deal with the third subcase, which arises only in the first scenario.

**Lemma 27.** *Suppose that  $\frac{1}{1+\pi} \bar{\theta} < \bar{\theta}_F$  – i.e. that we are in the first scenario – and that  $\theta_1 \in [\frac{1}{1+\pi} \bar{\theta}, \bar{\theta}_F)$ . Then  $\psi'(\theta_1) > 0$ .*

**Proof.** Since  $\theta_1 \geq \frac{1}{1+\pi} \bar{\theta}$ , we may apply Lemma 20 to obtain

$$\psi'(\theta_1) = \frac{\chi(\theta_1) - G(\theta_1)}{\bar{\theta} - \theta_1}.$$

Since  $\theta_1 < \bar{\theta}_F$ , we may apply Lemma 26 to obtain  $\chi(\theta_1) - G(\theta_1) > 0$ . The result follows. ■

Combining Lemmas 21, 24, 25 and 27, we obtain:

**Proposition 28.** *The set of  $\theta_1 \in (0, \bar{\theta}_F)$  such that  $\psi(\theta_1) = F(\bar{\theta})$  is a closed interval. ■*

The idea behind the proof of the proposition is straightforward. We know from Proposition 11 that all solutions to the equation  $\psi = F(\bar{\theta})$  lie in  $(\frac{1}{1+\pi} \underline{\theta}, \bar{\theta}_F)$ . Hence, to prove the proposition, we need only show that  $\psi' \geq 0$  on this interval. Furthermore this is what Lemma 21 (for the interval  $(\frac{1}{1+\pi} \underline{\theta}, \frac{1}{1+\pi} \bar{\theta}_F]$ ), Lemmas 24 and 25 (for the interval  $(\frac{1}{1+\pi} \bar{\theta}_F, \frac{1}{1+\pi} \bar{\theta})$ ) and Lemma 27 (for the interval  $[\frac{1}{1+\pi} \bar{\theta}, \bar{\theta}_F)$ ) seem to tell us. The only complication is that Lemmas 24 and 25 both require the side condition  $\psi \leq F(\bar{\theta})$ . However, they make up for this by providing strict rather than weak inequalities. The proof of the Proposition does therefore go through.

Indeed, we actually have  $\psi' > 0$  on the interval  $(\frac{1}{1+\pi} \bar{\theta}_F, \bar{\theta}_F)$ . Hence, the only way in which non-uniqueness can occur at all is if there is a non-trivial interval, contained

in  $(\frac{1}{1+\pi}\underline{\theta}, \frac{1}{1+\pi}\bar{\theta}_F]$ , on which  $\psi = F(\bar{\theta})$ . Unfortunately, it is possible to construct an example in which precisely this form of non-uniqueness occurs. The spirit of the example is that there exist  $\theta_3, \theta_4 \in [\underline{\theta}, \bar{\theta}_F)$  such that: (i)  $\theta_4 > (1 + \pi)\theta_3$  (i.e. it is possible that the entire interval of types associated with the kink lies within  $[\theta_3, \theta_4)$ ); and (ii)  $G_L(\theta_4) = G(\theta_3)$  (i.e.  $G$  is constant on  $[\theta_3, \theta_4)$ ). It then follows that, if there exists  $\theta_1 \in [\theta_3, \frac{1}{1+\pi}\theta_4]$  such that  $\psi(\theta_1) = F(\bar{\theta})$ , then  $\psi(\theta_1) = F(\bar{\theta})$  for all  $\theta_1 \in [\theta_3, \frac{1}{1+\pi}\theta_4]$ .

There are two ways to eliminate this possibility. The first way is to ensure that  $G$  cannot have a “flat” of the type envisaged. The following assumption is more than sufficient to ensure this:

**A4**  $G$  is strictly increasing on  $[\underline{\theta}, \theta_M)$ .<sup>13</sup>

We then have:

**Proposition 29.** *Suppose that Assumptions A1-A4 hold. Then there is a unique  $\theta_1 \in (0, \bar{\theta}_F)$  such that  $\psi(\theta_1) = F(\bar{\theta})$ . ■*

Working with Assumption A4 certainly simplifies our comparative statics. See Sections K-N below. However, we can still obtain satisfactory comparative-statics results without it. See Section O below.

The second way to eliminate the possibility of non-uniqueness is ensure that  $G$  cannot have a long enough flat:

**Proposition 30.** *Suppose that Assumptions A1-A3 hold, and that  $\pi > \frac{\bar{\theta}-\underline{\theta}}{\underline{\theta}}$ . Then there is a unique  $\theta_1 \in (0, \bar{\theta}_F)$  such that  $\psi(\theta_1) = F(\bar{\theta})$ . ■*

In particular, if  $\pi = \infty$ , then we there is a unique optimum within our one-parameter family of candidate optima.

## K. COMPARATIVE STATICS WITH A4

The analysis of Sections H-J shows that, for all  $\pi \in [0, \infty)$ , the set of solutions of the equation

$$\psi(\theta_1, \pi) = F(\bar{\theta}) \tag{26}$$

---

<sup>13</sup>Notice that  $G$  is identically 0 on  $(0, \underline{\theta})$ . It does not therefore make sense to require that  $G$  is strictly increasing on  $(0, \theta_M)$ .

is a non-empty interval. We denote this interval by  $\tau(\pi) = [\underline{\tau}(\pi), \bar{\tau}(\pi)]$ . The purpose of the current section is to investigate the dependence of  $\tau$  on  $\pi$ .

In order to simplify the exposition, it will be helpful to assume for the time being that A4 holds. This ensures that the interval  $\tau(\pi)$  collapses to a single point, which we shall denote by  $\tau_1(\pi)$ . It also ensures that  $\frac{\partial \psi}{\partial \theta_1}(\tau_1(\pi), \pi) > 0$ .

If we assume further that all the functions involved are sufficiently smooth, then we can apply the implicit function to the equation

$$\psi(\tau_1(\pi), \pi) = F(\bar{\theta})$$

to conclude that

$$\tau'_1 = - \frac{\frac{\partial \psi}{\partial \pi}}{\frac{\partial \psi}{\partial \theta_1}}. \quad (27)$$

In particular:  $\tau_1$  will be increasing (decreasing) in  $\pi$  iff  $\frac{\partial \psi}{\partial \pi} < 0$  ( $\frac{\partial \psi}{\partial \pi} > 0$ ); and the allocation to the illiquid account will be increasing (decreasing) in  $\pi$  iff  $\frac{\partial \psi}{\partial \pi} > 0$  ( $\frac{\partial \psi}{\partial \pi} < 0$ ).

Motivated by these observations, we look first at the case in which the maximum-penalty constraint is strictly binding. More precisely, we put  $\tau_2(\pi) = (1 + \pi)\tau_1(\pi)$ , and we consider the case in which  $\tau_2(\pi) < \bar{\theta}$ . In other words, there is a non-trivial interval of types  $(\tau_2(\pi), \bar{\theta})$  who choose to consume out of the illiquid account and therefore pay the penalty for doing so. In this case we begin by finding explicit formulae for  $\frac{\partial \psi}{\partial \pi}$  and  $\frac{\partial \psi}{\partial \theta_1}$ . We then go on to find conditions under which  $\frac{\partial \psi}{\partial \pi} > 0$  and  $\frac{\partial \psi}{\partial \theta_1} > 0$ , thereby ensuring that  $\tau'_1 < 0$  (and hence that the allocation to the illiquid account will be strictly increasing in  $\pi$ ).

We look second at the case in which the maximum-penalty constraint is strictly slack. More precisely, we consider the case in which  $\tau_2(\pi) > \bar{\theta}$ . In other words, even the highest type is not tempted to consume out of the illiquid account. In this case we again begin by finding explicit formulae for  $\frac{\partial \psi}{\partial \pi}$  and  $\frac{\partial \psi}{\partial \theta_1}$ . It turns out that  $\frac{\partial \psi}{\partial \pi} = 0$ . We therefore concentrate on finding conditions under which  $\frac{\partial \psi}{\partial \theta_1} > 0$ , thereby ensuring that  $\tau'_1 = 0$  (and hence that the allocation to the illiquid account does not change with  $\pi$ ).

We look third at the intermediate case in which  $\tau_2(\pi) = \bar{\theta}$ . This case is important because it is  $\tau_2(\pi)$  that determines whether we are in the strictly binding case  $\tau_2(\pi) < \bar{\theta}$  or the strictly slack case  $\tau_2(\pi) > \bar{\theta}$ . Our analysis of the comparative statics of our problem is not therefore complete until we have understood how the transition

between these two cases occurs.

### L. THE STRICTLY BINDING CASE

In this section we focus on the set  $V$  of  $(\theta_1, \pi)$  such that

1.  $\theta_1 \in (0, \bar{\theta})$ ,
2.  $\pi \in (0, \infty)$  and
3.  $\theta_2 = (1 + \pi)\theta_1 < \bar{\theta}$ .

In other words, we do not impose the requirement that  $\theta_1 = \tau_1(\pi)$  (i.e. that  $\theta_1$  be optimal for the given  $\pi$ ), but we do require that the maximum-penalty constraint is binding (in the sense that types in the non-empty interval  $(\theta_2, \bar{\theta})$  are choosing to pay the penalty).

**L.1. The formula for  $\frac{\partial \psi}{\partial \pi}$ .** Consider the o.d.e.

$$\dot{\theta} = - \left( \theta + \beta b \left( \frac{\theta}{(1+\pi)\beta} \right) \right) \quad (28)$$

on  $[\theta_2, \bar{\theta}]$ , with initial condition  $\theta(0) = \bar{\theta}$ . Let  $T(h; \pi)$  denote the first hitting time of  $h \in [\theta_2, \bar{\theta}]$ , and put  $S(h; \pi) = \exp(-T(h; \pi))$ . Then the formula for  $\frac{\partial \psi}{\partial \pi}$  is given by the following proposition.

**Proposition 31.** *Suppose that  $\theta_2 < \bar{\theta}$ . Then*

$$\begin{aligned} \frac{\partial \psi}{\partial \pi}(\theta_1, \pi) &= \left( \frac{1}{\pi} S(\theta_2, \pi) - \theta_1 \frac{\partial S}{\partial h}(\theta_2, \pi) - \frac{\partial S}{\partial \pi}(\theta_2, \pi) \right) (G(\theta_2) - \phi(\theta_1, \pi)) \\ &\quad - \int_{(\theta_2, \bar{\theta}]} \frac{\partial S}{\partial \pi}(h, \pi) dG(h). \end{aligned}$$

**Proof.** Equation (19) can be written

$$\psi(\theta_1, \pi) = \int_{[\theta_2, \bar{\theta}]} \frac{\partial S}{\partial h}(h, \pi) G(h) dh + S(\theta_2, \pi) \phi(\theta_1, \pi).$$

Hence

$$\begin{aligned} \frac{\partial \psi}{\partial \pi} &= \int_{[\theta_2, \bar{\theta}]} \frac{\partial^2 S}{\partial h \partial \pi}(h, \pi) G(h) dh - \frac{\partial S}{\partial h}(\theta_2, \pi) G(\theta_2) \frac{\partial \theta_2}{\partial \pi} \\ &\quad + \left( \frac{\partial S}{\partial h}(\theta_2, \pi) \frac{\partial \theta_2}{\partial \pi} + \frac{\partial S}{\partial \pi}(\theta_2, \pi) \right) \phi + S(\theta_2, \pi) \frac{\partial \phi}{\partial \pi}, \end{aligned} \quad (29)$$

where we have suppressed the dependence of  $\psi$  and  $\phi$  on  $\theta_1$  and  $\pi$ . Now:

$$\begin{aligned}
 \int_{[\theta_2, \bar{\theta}]} \frac{\partial^2 S}{\partial h \partial \pi}(h, \pi) G(h) dh &= \int_{[\theta_2, \bar{\theta}]} \frac{\partial^2 S}{\partial \pi \partial h}(h, \pi) G(h) dh \\
 &= \left[ \frac{\partial S}{\partial \pi}(h, \pi) G(h) \right]_{\theta_2-}^{\bar{\theta}} - \int_{[\theta_2, \bar{\theta}]} \frac{\partial S}{\partial \pi}(h, \pi) dG(h) \\
 &= -\frac{\partial S}{\partial \pi}(\theta_2, \pi) G(\theta_2-) - \int_{[\theta_2, \bar{\theta}]} \frac{\partial S}{\partial \pi}(h, \pi) dG(h) \\
 &= -\frac{\partial S}{\partial \pi}(\theta_2, \pi) G(\theta_2) - \int_{(\theta_2, \bar{\theta}]} \frac{\partial S}{\partial \pi}(h, \pi) dG(h),
 \end{aligned}$$

where we have used the fact that  $\frac{\partial S}{\partial \pi}(\bar{\theta}) = 0$ ;

$$\frac{\partial \phi}{\partial \pi} = \frac{G(\theta_2) - \phi}{\theta_2 - \theta_1} \frac{\partial \theta_2}{\partial \pi} = \frac{G(\theta_2) - \phi}{\pi};$$

and

$$\frac{\partial \theta_2}{\partial \pi} = \theta_1.$$

Substituting into (29), we therefore obtain

$$\begin{aligned}
 \frac{\partial \psi}{\partial \pi} &= -\frac{\partial S}{\partial \pi}(\theta_2, \pi) G(\theta_2) - \int_{(\theta_2, \bar{\theta}]} \frac{\partial S}{\partial \pi}(h, \pi) dG(h) - \frac{\partial S}{\partial h}(\theta_2, \pi) G(\theta_2) \theta_1 \\
 &\quad + \left( \frac{\partial S}{\partial h}(\theta_2, \pi) \theta_1 + \frac{\partial S}{\partial \pi}(\theta_2, \pi) \right) \phi + S(\theta_2, \pi) \frac{G(\theta_2) - \phi}{\pi} \\
 &= -\frac{\partial S}{\partial \pi}(\theta_2, \pi) (G(\theta_2) - \phi(\theta_1, \pi)) - \int_{(\theta_2, \bar{\theta}]} \frac{\partial S}{\partial \pi}(h, \pi) dG(h) \\
 &\quad + \left( \frac{1}{\pi} S(\theta_2, \pi) - \theta_1 \frac{\partial S}{\partial h}(\theta_2, \pi) \right) (G(\theta_2) - \phi(\theta_1, \pi)).
 \end{aligned}$$

The required formula now follows on rearranging. ■

In view of Proposition 31, it is clear that there are three main contributions to  $\frac{\partial \psi}{\partial \pi}$ , namely:

1.  $\frac{1}{\pi} S(\theta_2, \pi) - \theta_1 \frac{\partial S}{\partial h}(\theta_2, \pi) - \frac{\partial S}{\partial \pi}(\theta_2, \pi);$
2.  $G(\theta_2) - \phi(\theta_1, \pi);$
3.  $-\int_{(\theta_2, \bar{\theta}]} \frac{\partial S}{\partial \pi}(h, \pi) dG(h).$

We discuss these contributions in turn.

The first contribution can be signed quite generally:

**Proposition 32.** *Suppose that  $\theta_2 \leq \bar{\theta}$ . Then*

$$\frac{1}{\pi} S(\theta_2, \pi) - \theta_1 \frac{\partial S}{\partial h}(\theta_2, \pi) - \frac{\partial S}{\partial \pi}(\theta_2, \pi) > 0.$$

*In other words, Contribution 1 is strictly positive.*

**Proof.** Explicit calculation shows that

$$\frac{1}{\pi} S(\theta_2, \pi) - \theta_1 \frac{\partial S}{\partial h}(\theta_2, \pi) - \frac{\partial S}{\partial \pi}(\theta_2, \pi) = \frac{N}{D},$$

where

$$\begin{aligned} N = 1 + (1 + \pi) & \left( \frac{\theta_2}{\beta(1 + \pi)} \right)^{1/\rho} + \left( \frac{\bar{\theta}}{\beta(1 + \pi)} \right)^{1/\rho} \\ & + (1 + \pi) \left( \frac{\theta_2}{\beta(1 + \pi)} \right)^{1/\rho} \left( \frac{\bar{\theta}}{\beta(1 + \pi)} \right)^{1/\rho} \\ & + \rho \pi \left( \left( \frac{\bar{\theta}}{\beta(1 + \pi)} \right)^{1/\rho} - \left( \frac{\theta_2}{\beta(1 + \pi)} \right)^{1/\rho} \right). \end{aligned}$$

and

$$D = \pi \left( 1 + (1 + \pi) \left( \frac{\theta_2}{\beta(1 + \pi)} \right)^{1/\rho} \right)^{1-\rho} \left( 1 + (1 + \pi) \left( \frac{\bar{\theta}}{\beta(1 + \pi)} \right)^{1/\rho} \right)^{1+\rho}.$$

Now the last term in the formula for  $N$  is non-negative, since  $\theta_2 \leq \bar{\theta}$ . Hence  $N > 0$ . Finally, it is clear that  $D > 0$ . ■

The second contribution can only be signed when  $\theta_1 = \tau_1(\pi)$  (or, more generally, when  $\theta_2 \in (\underline{\theta}, \bar{\theta})$  and  $\psi(\theta_1) \leq F(\bar{\theta})$ ). This, however, is enough for the purpose of our comparative statics.

**Proposition 33.** *Suppose that:*

1.  $\theta_2 \in (\underline{\theta}, \bar{\theta})$ ;
2.  $\psi(\theta_1) \leq F(\bar{\theta})$ ;

3. *Assumption A4 holds.*

Then  $G(\theta_2) - \phi(\theta_1, \pi) > 0$ . In other words, *Contribution 2 is strictly positive.*

**Proof.** We break the proof down into the cases  $\theta_2 \in (\underline{\theta}, \bar{\theta}_F)$  and  $\theta_2 \in [\bar{\theta}_F, \bar{\theta})$ . In the first case, the proof parallels that of Lemma 21. We have  $[\theta_1, \theta_2] \subset (\frac{1}{1+\pi}\underline{\theta}, \bar{\theta}_F) \subset (0, \theta_M)$  and  $\theta_2 > \underline{\theta}$ . Assumption A4 therefore implies that  $G(\theta_2) > \phi(\theta_1, \pi)$ . In the second case, it follows from the proof of Lemma 22 that  $G(\theta_2) - \phi(\theta_1, \pi) > 0$ . ■

The third contribution cannot be signed under our primary assumptions. It is, however, worth drawing attention to three special cases in which it can be signed. In all three cases, the comparative statics end up going the same way:  $\frac{\partial \psi}{\partial \pi} > 0$ , and therefore the allocation to the illiquid account will be increasing in  $\pi$ . We state these three cases as separate propositions, corresponding to the cases  $\rho < 1$ ,  $\rho = 1$  and  $\rho > 1$ .

**Proposition 34.** *Suppose that:*

1.  $\rho < 1$ ;
2.  $G' \leq 0$  on  $(\underline{\theta}, \infty)$  (i.e.  $\theta_M = \underline{\theta}$ );
3.  $\theta_2 \in (\underline{\theta}, \bar{\theta})$ .

Then  $-\int_{(\theta_2, \bar{\theta}]} \frac{\partial S}{\partial \pi}(h, \pi) dG(h) \geq 0$ . In other words, *Contribution 3 is non-negative.*

**Proof.** It is easy to show that we have

$$\frac{\partial S}{\partial \pi}(h, \pi) \begin{cases} > 0 & \text{if } \rho < 1 \\ = 0 & \text{if } \rho = 1 \\ < 0 & \text{if } \rho > 1 \end{cases}$$

for all  $h \in [\theta_2, \bar{\theta})$ . Furthermore, we have

$$\frac{\partial S}{\partial \pi}(\bar{\theta}, \pi) = 0 \quad \text{for all } \rho,$$

because  $S(\bar{\theta}, \pi) = 1$ . We can therefore proceed as follows.

First, we know that  $\theta_2 \in (\underline{\theta}, \bar{\theta})$ . Hence  $G' \leq 0$  on  $(\theta_2, \bar{\theta}] \subset (\underline{\theta}, \infty)$ . Second,  $\rho < 1$ . Hence  $\frac{\partial S}{\partial \pi}(\cdot, \pi) \geq 0$  on  $(\theta_2, \bar{\theta}] \subset [\theta_2, \bar{\theta}]$ . Putting these two observations together gives us the required conclusion. ■

**Remark 35.** If  $G' \leq 0$  on  $(\underline{\theta}, \infty)$  then necessarily  $\Delta G(\underline{\theta}) > 0$ . Hence it is essential for the proof of Proposition 34 that we restrict attention to  $\theta_2 > \underline{\theta}$ .

**Proposition 36.** Suppose that:

1.  $\rho = 1$ ;
2.  $\theta_2 \in (0, \bar{\theta})$ .

Then  $-\int_{(\theta_2, \bar{\theta}]} \frac{\partial S}{\partial \pi}(h, \pi) dG(h) = 0$ . In other words, Contribution 3 is zero.

**Proof.** This follows at once from the fact that  $\frac{\partial S}{\partial \pi}(\cdot, \pi) = 0$  on  $[\theta_2, \bar{\theta}]$ . ■

**Proposition 37.** Suppose that:

1.  $\rho > 1$ ;
2.  $G' \geq 0$  on  $(0, \bar{\theta})$  (i.e.  $\theta_M = \bar{\theta}$ );
3.  $\theta_2 \in (0, \bar{\theta})$ .

Then  $-\int_{(\theta_2, \bar{\theta}]} \frac{\partial S}{\partial \pi}(h, \pi) dG(h) \geq 0$ . In other words, Contribution 3 is non-negative.

**Proof.** Note first that  $G' \geq 0$  on  $(\theta_2, \bar{\theta}) \subset (0, \bar{\theta})$ . Second,  $\rho > 1$ . Hence  $\frac{\partial S}{\partial \pi}(\cdot, \pi) < 0$  on  $(\theta_2, \bar{\theta}) \subset [\theta_2, \bar{\theta}]$ . (Cf. the proof of Proposition .) Third,  $\frac{\partial S}{\partial \pi}(\bar{\theta}, \pi) = 0$ . Putting these three observations together, we obtain

$$\begin{aligned} \int_{(\theta_2, \bar{\theta}]} \frac{\partial S}{\partial \pi}(h, \pi) dG(h) &= \int_{(\theta_2, \bar{\theta})} \frac{\partial S}{\partial \pi}(h, \pi) dG(h) + \int_{[\bar{\theta}, \bar{\theta}]} \frac{\partial S}{\partial \pi}(h, \pi) dG(h) \\ &= \int_{(\theta_2, \bar{\theta})} \frac{\partial S}{\partial \pi}(h, \pi) dG(h) \leq 0, \end{aligned}$$

as required. ■

**Remark 38.** If  $G' \geq 0$  on  $(0, \bar{\theta})$  then necessarily  $\Delta G(\bar{\theta}) < 0$ . The fact that  $\frac{\partial S}{\partial \pi}(\bar{\theta}, \pi) = 0$  therefore plays an essential role in the proof of Proposition 37.

**L.2. The formula for  $\frac{\partial \psi}{\partial \theta_1}$ .** Let  $T(h; \pi)$  and  $S(h; \pi) = \exp(-T(h; \pi))$  as in the preceding section. Then the formula for  $\frac{\partial \psi}{\partial \theta_1}$  is given by the following proposition.

**Proposition 39.** *Suppose that  $\theta_2 < \bar{\theta}$ . Then*

$$\begin{aligned} \frac{\partial \psi}{\partial \theta_1}(\theta_1, \pi) = & \left( \frac{\beta b\left(\frac{\theta_1}{\beta}\right)}{\theta_1 \left(\theta_2 + \beta b\left(\frac{\theta_1}{\beta}\right)\right)} (G(\theta_2) - \phi(\theta_1, \pi)) \right. \\ & \left. + \frac{1}{\pi \theta_1} (G(\theta_2) - G(\theta_1)) \right) S(\theta_2, \bar{\theta}). \end{aligned}$$

**Proof.** This is simply a restatement of Lemma 19. ■

In view of Proposition 39, there are two main contributions to  $\frac{\partial \psi}{\partial \theta_1}$ , namely

1.  $G(\theta_2) - \phi(\theta_1, \pi)$ ;
2.  $G(\theta_2) - G(\theta_1)$ .

We have already given conditions under which the first is strictly positive (in Proposition 33). The second is strictly positive under the same conditions:

**Proposition 40.** *Suppose that:*

1.  $\theta_2 \in (\underline{\theta}, \bar{\theta})$ ;
2.  $\psi(\theta_1) \leq F(\bar{\theta})$ ;
3. Assumption A4 holds.

*Then  $G(\theta_2) - G(\theta_1) > 0$ .*

**Proof.** We break the proof down into the cases  $\theta_1 \in (\frac{1}{1+\pi}\underline{\theta}, \frac{1}{1+\pi}\bar{\theta}_F)$  and  $\theta_1 \in [\frac{1}{1+\pi}\bar{\theta}_F, \frac{1}{1+\pi}\bar{\theta})$ . In the first case, the proof parallels that of Lemma 21. We have  $[\theta_1, \theta_2] \subset (\frac{1}{1+\pi}\bar{\theta}_F, \bar{\theta}_F) \subset (0, \theta_M)$ , and moreover  $\theta_2 > \underline{\theta}$ . Assumption A4 therefore implies that  $G(\theta_2) > G(\theta_1)$ . In the second case, Lemma 23 implies directly that that  $G(\theta_2) > G(\theta_1)$ . ■

M. THE STRICTLY SLACK CASE

In this section we focus on the set  $W$  of  $(\theta_1, \pi)$  such that

1.  $\theta_1 \in (0, \bar{\theta})$ ,
2.  $\pi \in (0, \infty)$  and
3.  $\theta_2 = (1 + \pi)\theta_1 > \bar{\theta}$ .

In other words, we do not impose the requirement that  $\theta_1 = \tau_1(\pi)$  (i.e. that  $\theta_1$  be optimal for the given  $\pi$ ), but we do require that the maximum-penalty constraint is slack in the sense that no types are choosing to pay the penalty.

**M.1. The formula for  $\frac{\partial \psi}{\partial \pi}$ .** The formula for  $\frac{\partial \psi}{\partial \pi}$  is given by the following proposition.

**Proposition 41.** *Suppose that  $\theta_2 > \bar{\theta}$ . Then*

$$\frac{\partial \psi}{\partial \pi}(\theta_1, \pi) = 0.$$

**Proof.** As in the proof of Lemma 20, we have  $\psi(\theta_1, \pi) = \chi(\theta_1)$  for  $\theta_1 \in (\frac{1}{1+\pi}\bar{\theta}, \bar{\theta})$ , where

$$\chi(\theta_1) = \frac{1}{\bar{\theta} - \theta_1} \int_{\theta_1}^{\bar{\theta}} G(\theta) d\theta.$$

Hence  $\psi$  is independent of  $\pi$  for such  $\theta_1$ . ■

**M.2. The formula for  $\frac{\partial \psi}{\partial \theta_1}$ .** The formula for  $\frac{\partial \psi}{\partial \theta_1}$  is given by the following proposition.

**Proposition 42.** *Suppose that  $\theta_2 > \bar{\theta}$ . Then*

$$\frac{\partial \psi}{\partial \theta_1}(\theta_1, \pi) = \frac{\chi(\theta_1) - G(\theta_1)}{\bar{\theta} - \theta_1}.$$

**Proof.** As already noted, we have  $\psi(\theta_1, \pi) = \chi(\theta_1)$  for  $\theta_1 \in (\frac{1}{1+\pi}\bar{\theta}, \bar{\theta})$ . Moreover

$$\chi'(\theta_1) = \frac{\chi(\theta_1) - G(\theta_1)}{\bar{\theta} - \theta_1},$$

as in the proof of Lemma 20. ■

In view of Proposition 42, there is really only one contribution to  $\frac{\partial \psi}{\partial \theta_1}$ , namely  $\chi(\theta_1) - G(\theta_1)$ . It is not possible to sign  $\chi(\theta_1) - G(\theta_1)$  for all  $\theta_1$ , but it is possible to sign it when  $\theta_1 = \tau_1(\pi)$ , and indeed much more generally when  $\theta_1 \in (0, \bar{\theta}_F)$ . As before, this is enough for the purpose of our comparative statics.

**Proposition 43.** *Suppose that  $\theta_1 \in (0, \bar{\theta}_F)$ . Then  $\chi(\theta_1) - G(\theta_1) > 0$ .*

**Proof.** This is simply a special case of Lemma 26. ■

## N. THE INTERMEDIATE CASE

Up to now we have focussed on the comparative statics of  $\tau_1$ . For example, we have shown that if A4 is satisfied and  $\rho = 1$  then: (i)  $\tau_1' < 0$  when  $\tau_2(\pi) < \bar{\theta}$ ; and (ii)  $\tau_1' = 0$  when  $\tau_2(\pi) > \bar{\theta}$ . However, this leaves open the question of what happens at the transition between the two cases. For example, does  $\tau_1$  jump up when  $\tau_2(\pi) = \bar{\theta}$ ? Does it jump down? Or is there more than one value of  $\pi$  for which  $\tau_2(\pi) = \bar{\theta}$ ?

In order to address these questions, we need to understand the comparative statics of  $\tau_2$ . These comparative statics are quite complex in the binding case. However, they simplify as the borderline between the two cases is approached. Moreover they are simpler still in the slack case.

**N.1. Comparative Statics of  $\tau_2$  in the Weakly Binding Case.** We begin this section by looking at the comparative statics of  $\tau_2$  when the maximum-penalty constraint is strictly binding (in the sense that  $\tau_2(\pi) < \bar{\theta}$ ). More precisely, we show that  $\tau_2'(\pi)$  satisfies a simple linear equation. We then go on to check whether this equation remains valid when the maximum-penalty constraint is only weakly binding (in the sense that  $\tau_2(\pi) \uparrow \bar{\theta}$ ).

**Proposition 44.** *Suppose that  $\tau_2(\pi) < \bar{\theta}$ . Then*

$$D(\tau_1(\pi), \pi) \tau_2'(\pi) = N(\tau_1(\pi), \pi), \quad (30)$$

where

$$D(\theta_1, \pi) = (1 + \pi) \frac{\theta_1 + \beta b \left( \frac{\theta_1}{\beta} \right)}{\theta_2 + \beta b \left( \frac{\theta_1}{\beta} \right)} (G(\theta_2) - \phi(\theta_1, \pi)) + (\phi(\theta_1, \pi) - G(\theta_1))$$

and

$$N(\theta_1, \pi) = \left( \frac{\pi(1+\pi) \frac{\partial S}{\partial \pi}(\theta_2, \pi)}{S(\theta_2, \pi)} (G(\theta_2) - \phi(\theta_1, \pi)) \right. \\ \left. + (\phi(\theta_1, \pi) - G(\theta_1)) + \frac{\pi(1+\pi)}{S(\theta_2, \pi)} \int_{(\theta_2, \bar{\theta}]} \frac{\partial S}{\partial \pi}(h; \pi) dG(h) \right) \theta_1.$$

**Proof.** We have

$$\tau_2(\pi) = (1 + \pi) \tau_1(\pi)$$

and therefore

$$\tau_2'(\pi) = \tau_1(\pi) + (1 + \pi) \tau_1'(\pi). \quad (31)$$

Now,

$$\frac{\partial \psi}{\partial \theta_1}(\tau_1(\pi), \pi) \tau_1'(\pi) + \frac{\partial \psi}{\partial \pi}(\tau_1(\pi), \pi) = 0.$$

Hence, multiplying (31) through by  $\frac{\partial \psi}{\partial \theta_1}(\tau_1(\pi), \pi)$ , we obtain

$$\begin{aligned} \frac{\partial \psi}{\partial \theta_1} \tau_2' &= \frac{\partial \psi}{\partial \theta_1} \tau_1 + (1 + \pi) \frac{\partial \psi}{\partial \theta_1} \tau_1' \\ &= \frac{\partial \psi}{\partial \theta_1} \tau_1 - (1 + \pi) \frac{\partial \psi}{\partial \pi}, \end{aligned}$$

where we have suppressed the dependence of  $\frac{\partial \psi}{\partial \theta_1}$  and  $\frac{\partial \psi}{\partial \pi}$  on  $\tau_1(\pi)$  and  $\pi$ , and the dependence of  $\tau_1$  and  $\tau_2$  on  $\pi$ . We may therefore put

$$D(\theta_1, \pi) = \frac{\pi \theta_1}{S(\theta_2, \bar{\theta})} \frac{\partial \psi}{\partial \theta_1}(\theta_1, \pi)$$

and

$$N(\theta_1, \pi) = \frac{\pi \theta_1}{S(\theta_2, \bar{\theta})} \left( \frac{\partial \psi}{\partial \theta_1}(\theta_1, \pi) \theta_1 - (1 + \pi) \frac{\partial \psi}{\partial \pi}(\theta_1, \pi) \right).$$

Equation (30) now follows on applying the formulae for  $\frac{\partial \psi}{\partial \pi}(\theta_1, \pi)$  and  $\frac{\partial \psi}{\partial \theta_1}(\theta_1, \pi)$  given in Propositions 31 and 39. ■

Equation (30) can be solved for  $\tau_2'(\pi)$  under the conditions of Proposition 33, namely that: (i)  $\tau_2(\pi) \in (\underline{\theta}, \bar{\theta})$ ; (ii)  $\psi(\tau_1(\pi)) \leq F(\bar{\theta})$ ; and (iii) Assumption A4 holds. This is not, however, enough for our current purposes: we need to make sure that it can still be solved for  $\tau_2'(\pi)$  in the limiting case  $\tau_2(\pi) \uparrow \bar{\theta}$ . To this end, recall

that

$$V = \{(\theta_1, \pi) \mid \theta_1 \in (0, \bar{\theta}), \pi \in (0, \infty), \theta_2 < \bar{\theta}\},$$

and put

$$\partial V = \{(\theta_1, \pi) \mid \theta_1 \in (0, \bar{\theta}), \pi \in (0, \infty), \theta_2 = \bar{\theta}\}.$$

Furthermore, for all  $(\theta_1, \pi) \in V \cup \partial V$ , put

$$\begin{aligned} \underline{D}(\theta_1, \pi) &= (1 + \pi) \frac{\theta_1 + \beta b \left( \frac{\theta_1}{\beta} \right)}{\theta_2 + \beta b \left( \frac{\theta_1}{\beta} \right)} (G_L(\theta_2) - \phi(\theta_1, \pi)) \\ &\quad + (\phi(\theta_1, \pi) - \max \{G(\theta_1), G_L(\theta_1)\}) \end{aligned}$$

and

$$\begin{aligned} \overline{D}(\theta_1, \pi) &= (1 + \pi) \frac{\theta_1 + \beta b \left( \frac{\theta_1}{\beta} \right)}{\theta_2 + \beta b \left( \frac{\theta_1}{\beta} \right)} (G_L(\theta_2) - \phi(\theta_1, \pi)) \\ &\quad + (\phi(\theta_1, \pi) - \min \{G(\theta_1), G_L(\theta_1)\}). \end{aligned}$$

Then we have:

**Lemma 45.** *Suppose that  $(\tilde{\theta}_1, \tilde{\pi}) \in V \rightarrow (\theta_1, \pi) \in \partial V$ . Then*

$$\underline{D}(\theta_1, \pi) \leq \liminf D(\tilde{\theta}_1, \tilde{\pi}) \leq \limsup D(\tilde{\theta}_1, \tilde{\pi}) \leq \overline{D}(\theta_1, \pi).$$

**Proof.** Note first that  $b$  and  $\phi$  are both continuous. Hence  $b \left( \frac{\tilde{\theta}_1}{\beta} \right) \rightarrow b \left( \frac{\theta_1}{\beta} \right)$  and  $\phi(\tilde{\theta}_1, \tilde{\pi}) \rightarrow \phi(\theta_1, \pi)$ . Next, put  $\tilde{\theta}_2 = (1 + \pi) \tilde{\theta}_1$  and  $\theta_2 = (1 + \pi) \theta_1$ . Then  $\tilde{\theta}_2 \uparrow \theta_2$ , and therefore  $G(\tilde{\theta}_2) \rightarrow G_L(\theta_2)$ . Finally,

$$\begin{aligned} \min \{G(\theta_1), G_L(\theta_1)\} &\leq \liminf G(\tilde{\theta}_1) \\ &\leq \limsup G(\tilde{\theta}_1) \\ &\leq \max \{G(\theta_1), G_L(\theta_1)\}. \end{aligned}$$

The result follows. ■

The next step is to sign  $\underline{D}$ . This cannot be done everywhere on  $V \cup \partial V$ . But it can be done when  $\theta_2 = \bar{\theta}$  and  $\theta_1 = \tau_1(\pi)$ . Indeed, it is enough to require that  $\theta_2 \in (\bar{\theta}_F, \bar{\theta}]$  (i.e. we do not actually have to be on the boundary) and that  $\psi(\theta_1, \pi) \leq F(\bar{\theta})$  (i.e. we do not actually have to be at an optimum). We begin with a lemma.

**Lemma 46.** *Suppose that:*

1.  $\theta_2 \in (\bar{\theta}_F, \bar{\theta}]$ ;
2.  $\psi(\theta_1) \leq F(\bar{\theta})$ .

Then  $G_L(\theta_2) > G(\theta_1) \geq G_L(\theta_1)$ .

**Proof.** The proof is similar to that of Lemma 23. Note first that

$$G_L(\theta_2) \geq F(\bar{\theta})$$

(with strict inequality if  $\theta_2 < \bar{\theta}$ )

$$\geq \psi(\theta_1) = \Psi(\bar{\theta}; \theta_1) > \Psi(\theta_1; \theta_1) = G(\theta_1)$$

(by assumption, by definition of  $\psi$ , because  $\theta_1 < \theta_2 \leq \bar{\theta}$  and by construction of  $\Psi$  respectively). Second, Lemma 9 tells us that  $\psi > F(\bar{\theta})$  on  $[\bar{\theta}_F, \bar{\theta})$ . But we have  $\psi(\theta_1) \leq F(\bar{\theta})$ . Hence  $\theta_1 < \bar{\theta}_F$  and therefore  $G' \geq 0$  at  $\theta_1 \in (0, \bar{\theta}_F) \subset (0, \theta_M)$ . That is,  $G(\theta_1) - G_L(\theta_1) = \Delta G(\theta_1) \geq 0$ . ■

We can now sign  $\underline{D}$ .

**Proposition 47.** *Suppose that:*

1.  $\theta_2 \in (\bar{\theta}_F, \bar{\theta}]$ ;
2.  $\psi(\theta_1) \leq F(\bar{\theta})$ .

Then  $\underline{D}(\theta_1, \pi) > 0$ .

**Proof.** Two things follow from Lemma 46. First,  $G(\theta_1) \geq G_L(\theta_1)$ . Hence the formula for  $\underline{D}(\theta_1, \pi)$  simplifies to

$$\underline{D}(\theta_1, \pi) = (1 + \pi) \frac{\theta_1 + \beta b\left(\frac{\theta_1}{\beta}\right)}{\bar{\theta} + \beta b\left(\frac{\theta_1}{\beta}\right)} (G_L(\theta_2) - \phi(\theta_1, \pi)) + (\phi(\theta_1, \pi) - G(\theta_1)).$$

In particular,  $\underline{D}(\theta_1, \pi)$  is a strictly positive linear combination of the two terms  $G_L(\theta_2) - \phi(\theta_1, \pi)$  and  $\phi(\theta_1, \pi) - G(\theta_1)$ . Second,  $G_L(\theta_2) - G(\theta_1) > 0$ . Hence the sum of the two terms  $G_L(\theta_2) - \phi(\theta_1, \pi)$  and  $\phi(\theta_1, \pi) - G(\theta_1)$  is strictly positive. It therefore suffices to show that each of these two terms is non-negative. We have

$$G_L(\theta_2) \geq F(\bar{\theta}) \geq \psi(\theta_1) = \Psi(\bar{\theta}; \theta_1)$$

(as in the proof of Lemma 46)

$$\geq \Psi(\theta_2; \theta_1) > \Psi(\theta_1; \theta_1)$$

(since  $\Psi' \geq 0$  on  $(\theta_1, \bar{\theta}_F)$  (by Proposition 17) and  $\Psi' > 0$  on  $[\bar{\theta}_F, \bar{\theta})$  (by Proposition 18))

$$= G(\theta_1)$$

(again as in the proof of Lemma 46). In particular, since  $\Psi(\theta_2; \theta_1) = \phi(\theta_1, \pi)$ , we have  $G_L(\theta_2) \geq \phi(\theta_1, \pi)$  and  $\phi(\theta_1, \pi) > G(\theta_1)$ . ■

Since  $\underline{D} > 0$ , finding the sign of  $N$  and finding the sign of  $\tau'_2(\pi)$  amount to the same thing. Note first that

$$\tau_2(\pi) = (1 + \pi) \tau_1(\pi)$$

and hence

$$\tau'_2(\pi) = (1 + \pi) \tau'_1(\pi) + \tau_1(\pi).$$

We therefore face a tension. On the one hand, we are mainly interested in the case in which  $\tau'_1(\pi) < 0$ . For our purposes, then, the first contribution to  $\tau'_2(\pi)$  (namely  $(1 + \pi) \tau'_1(\pi)$ ) is negative. However, the second contribution (namely  $\tau_1(\pi)$ ) is necessarily positive. The net effect is therefore ambiguous. Worse still, what we really need to show for the purposes of comparative statics is that  $\tau'_2(\pi) > 0$  (so that the curve  $(\tau_1(\pi), \pi)$  crosses the boundary  $\theta_2 = \bar{\theta}$  in a simple way). This is directly at odds with our interest in the case in which  $\tau'_1(\pi) < 0$ .

Fortunately, the problem of signing  $\tau'_2(\pi)$  at the boundary is much simpler than the problem of signing  $\tau'_2(\pi)$  in  $V$ . With this end in mind, for all  $(\theta_1, \pi) \in \partial V$ , put

$$\underline{N}(\theta_1, \pi) = (\phi(\theta_1, \pi) - \max\{G(\theta_1), G_L(\theta_1)\}) \theta_1$$

and

$$\overline{N}(\theta_1, \pi) = (\phi(\theta_1, \pi) - \min \{G(\theta_1), G_L(\theta_1)\}) \theta_1.$$

Then we have the following lemma.

**Lemma 48.** *Suppose that  $(\tilde{\theta}_1, \tilde{\pi}) \in V \rightarrow (\theta_1, \pi) \in \partial V$ . Then*

$$\underline{N}(\theta_1, \pi) \leq \liminf N(\tilde{\theta}_1, \tilde{\pi}) \leq \limsup N(\tilde{\theta}_1, \tilde{\pi}) \leq \overline{N}(\theta_1, \pi).$$

**Proof.** The proof is similar to that of Lemma 45. Put  $\tilde{\theta}_2 = (1 + \pi)\tilde{\theta}_1$  and  $\theta_2 = (1 + \pi)\theta_1$ . Then  $\tilde{\theta}_2 \uparrow \theta_2 = \bar{\theta}$ , and therefore  $\frac{\partial S}{\partial \pi}(\tilde{\theta}_2, \pi) \rightarrow 0$  and  $\int_{(\theta_2, \bar{\theta}]} \frac{\partial S}{\partial \pi}(h; \pi) dG(h) \rightarrow 0$ . Furthermore  $\phi(\tilde{\theta}_1, \tilde{\pi}) \rightarrow \phi(\theta_1, \pi)$  and

$$\begin{aligned} \min \{G(\theta_1), G_L(\theta_1)\} &\leq \liminf G(\tilde{\theta}_1) \\ &\leq \limsup G(\tilde{\theta}_1) \\ &\leq \max \{G(\theta_1), G_L(\theta_1)\}. \end{aligned}$$

Passing to the limit in the formula given for  $N$  in the statement of Proposition 44, we therefore obtain the required result. ■

Combining Lemma 48 with the earlier Lemma 46, we obtain:

**Proposition 49.** *Suppose that:*

1.  $\theta_2 = \bar{\theta}$ ;
2.  $\psi(\theta_1) \leq F(\bar{\theta})$ .

Then  $\underline{N}(\theta_1, \pi) > 0$ .

**Proof.** The proof is similar to that of Proposition 47. First, because  $\psi(\theta_1) \leq F(\bar{\theta})$  and therefore  $\theta_1 < \bar{\theta}_F$ , the formula for  $\underline{N}(\theta_1, \pi)$  simplifies to

$$\underline{N}(\theta_1, \pi) = (\phi(\theta_1, \pi) - G(\theta_1)) \theta_1.$$

Second, we have

$$\Psi(\bar{\theta}; \theta_1) > \Psi(\theta_1; \theta_1) = G(\theta_1).$$

It remains only to note that, because  $\theta_2 = \bar{\theta}$ , we have  $\phi(\theta_1, \pi) = \Psi(\bar{\theta}; \theta_1)$ . ■

Combining Propositions 47 and 49, we see that  $\tau'_2(\pi) > 0$  on  $\partial V$ . In other words, whatever the behaviour of the curve  $(\tau_1(\pi), \pi)$  in  $V$ , it points out of  $V$  at  $\partial V$ . I.e. it can exit, but not enter,  $V$  at  $\partial V$ . In particular, there exists  $\pi_1 \in (0, \infty)$  such that  $\tau_2(\pi) < \bar{\theta}$  iff  $\pi \in [0, \pi_1)$ .

**N.2. Comparative Statics of  $\tau_2$  in the Weakly Slack Case.** We begin this section by looking at the comparative statics of  $\tau_2$  when the maximum-penalty constraint is strictly slack (in the sense that  $\tau_2(\pi) > \bar{\theta}$ ). More precisely, we show that  $\tau'_2(\pi)$  satisfies a simple linear equation. We then go on to check whether this equation remains valid when the maximum-penalty constraint is only weakly slack (in the sense that  $\tau_2(\pi) \downarrow \bar{\theta}$ ). Our first proposition is analogous to Proposition 44.

**Proposition 50.** *Suppose that  $\tau_2(\pi) > \bar{\theta}$ . Then*

$$D(\tau_1(\pi), \pi) \tau'_2(\pi) = N(\tau_1(\pi), \pi), \quad (32)$$

where

$$D(\theta_1, \pi) = \frac{\chi(\theta_1) - G(\theta_1)}{\bar{\theta} - \theta_1}$$

and

$$N(\theta_1, \pi) = \frac{\chi(\theta_1) - G(\theta_1)}{\bar{\theta} - \theta_1} \theta_1.$$

Notice that, if  $\chi(\theta_1) - G(\theta_1) > 0$ , then we can divide through by  $D(\tau_1(\pi), \pi)$  to conclude that  $\tau'_2(\pi) = \theta_1$ . Furthermore  $\chi(\theta_1) - G(\theta_1) > 0$  if  $\theta_1 = \tau_1(\pi)$ , and indeed much more generally if  $\theta_1 \in (0, \bar{\theta}_F)$ . Cf. Proposition 43. But it does not hold for all  $(\theta_1, \pi) \in W$ .

**Proof.** As in the proof of Proposition 44, we have

$$\frac{\partial \psi}{\partial \theta_1} \tau'_2 = \frac{\partial \psi}{\partial \theta_1} \tau_1 - (1 + \pi) \frac{\partial \psi}{\partial \pi}.$$

We may therefore put

$$D(\theta_1, \pi) = \frac{\partial \psi}{\partial \theta_1}(\theta_1, \pi)$$

and

$$N(\theta_1, \pi) = \frac{\partial \psi}{\partial \theta_1}(\theta_1, \pi) \theta_1 - (1 + \pi) \frac{\partial \psi}{\partial \pi}(\theta_1, \pi).$$

Equation (32) now follows on applying the formulae for  $\frac{\partial \psi}{\partial \pi}(\theta_1, \pi)$  and  $\frac{\partial \psi}{\partial \theta_1}(\theta_1, \pi)$  given in Propositions 41 and 42. ■

The next step is to ensure that equation (32) can still be solved for  $\tau_2'(\pi)$  in the limiting case  $\tau_2(\pi) \downarrow \bar{\theta}$ . To this end, recall that

$$W = \{(\theta_1, \pi) \mid \theta_1 \in (0, \bar{\theta}), \pi \in (0, \infty), \theta_2 > \bar{\theta}\},$$

and put

$$\partial W = \{(\theta_1, \pi) \mid \theta_1 \in (0, \bar{\theta}), \pi \in (0, \infty), \theta_2 = \bar{\theta}\}.$$

Furthermore, for all  $(\theta_1, \pi) \in W \cup \partial W$ , put

$$\begin{aligned} \underline{D}(\theta_1, \pi) &= \frac{\chi(\theta_1) - \max\{G(\theta_1), G_L(\theta_1)\}}{\bar{\theta} - \theta_1}, \\ \overline{D}(\theta_1, \pi) &= \frac{\chi(\theta_1) - \min\{G(\theta_1), G_L(\theta_1)\}}{\bar{\theta} - \theta_1} \end{aligned}$$

and

$$\begin{aligned} \underline{N}(\theta_1, \pi) &= \frac{\chi(\theta_1) - \max\{G(\theta_1), G_L(\theta_1)\}}{\bar{\theta} - \theta_1} \theta_1, \\ \overline{N}(\theta_1, \pi) &= \frac{\chi(\theta_1) - \min\{G(\theta_1), G_L(\theta_1)\}}{\bar{\theta} - \theta_1} \theta_1. \end{aligned}$$

Then we have:

**Lemma 51.** *Suppose that  $(\tilde{\theta}_1, \tilde{\pi}) \in W \rightarrow (\theta_1, \pi) \in \partial W$ . Then*

$$\underline{D}(\theta_1, \pi) \leq \liminf D(\tilde{\theta}_1, \tilde{\pi}) \leq \limsup D(\tilde{\theta}_1, \tilde{\pi}) \leq \overline{D}(\theta_1, \pi)$$

and

$$\underline{N}(\theta_1, \pi) \leq \liminf N(\tilde{\theta}_1, \tilde{\pi}) \leq \limsup N(\tilde{\theta}_1, \tilde{\pi}) \leq \overline{N}(\theta_1, \pi).$$

**Proof.** Note first that  $\chi$  is continuous. Hence  $\chi(\tilde{\theta}_1) \rightarrow \chi(\theta_1)$ . On the other hand, as in the proof of Lemma 45,

$$\begin{aligned} \min\{G(\theta_1), G_L(\theta_1)\} &\leq \liminf G(\tilde{\theta}_1) \\ &\leq \limsup G(\tilde{\theta}_1) \\ &\leq \max\{G(\theta_1), G_L(\theta_1)\}. \end{aligned}$$

The result follows. ■

The next step is to sign  $\underline{D}$ . This cannot be done everywhere on  $W \cup \partial W$ . But it can be done when  $\theta_2 = \bar{\theta}$  and  $\theta_1 = \tau_1(\pi)$ . Indeed, it is enough to require that  $\theta_2 \in [\bar{\theta}, \infty)$  (i.e. we do not actually have to be on the boundary) and that  $\psi(\theta_1, \pi) \leq F(\bar{\theta})$  (i.e. we do not actually have to be at an optimum). We begin with a lemma.

**Lemma 52.** *Suppose that:*

1.  $\theta_2 \in [\bar{\theta}, \infty)$ ;
2.  $\psi(\theta_1) \leq F(\bar{\theta})$ .

*Then  $G(\theta_1) \geq G_L(\theta_1)$ .*

**Proof.** The proof is identical to the relevant part of that of Lemma 46. Since  $\psi(\theta_1) \leq F(\bar{\theta})$ , we must have  $\theta_1 < \bar{\theta}_F$ . Hence  $G' \geq 0$  at  $\theta_1 \in (0, \bar{\theta}_F) \subset (0, \theta_M)$ . ■

**Proposition 53.** *Suppose that:*

1.  $\theta_2 \in [\bar{\theta}, \infty)$ ;
2.  $\psi(\theta_1) \leq F(\bar{\theta})$ .

*Then  $\underline{D}(\theta_1, \pi), \underline{N}(\theta_1, \pi) > 0$ .*

**Proof.** Note first that, in view of Lemma 52, we have

$$\underline{D}(\theta_1, \pi) = \frac{\chi(\theta_1) - G(\theta_1)}{\bar{\theta} - \theta_1}$$

and

$$\underline{N}(\theta_1, \pi) = \frac{\chi(\theta_1) - G(\theta_1)}{\bar{\theta} - \theta_1} \theta_1.$$

Second, since  $\psi(\theta_1) \leq F(\bar{\theta})$ , we have  $\theta_1 < \bar{\theta}_F$ . Finally, Proposition 43 tells us that  $\chi(\theta_1) - G(\theta_1) > 0$  for  $\theta_1 \in (0, \bar{\theta}_F)$ . ■

It follows from Proposition 53 that  $\tau'_2(\pi) > 0$  on  $\partial W$ . In other words, whatever the behaviour of the curve  $(\tau_1(\pi), \pi)$  in  $W$ , it points into  $W$  at  $\partial W$ . I.e. it can enter, but not exit,  $W$  at  $\partial W$ . In particular, there exists  $\pi_2 \in (0, \infty)$  such that  $\tau_2(\pi) > \bar{\theta}$  iff  $\pi \in (\pi_2, \infty)$ .

**N.3. Comparative Statics of  $\tau_2$  in the Remaining Case.** At this point we have established that there exist  $0 < \pi_1 \leq \pi_2 < \infty$  such that  $\tau_2(\pi) < \bar{\theta}$  iff  $\pi \in [0, \pi_1)$  and  $\tau_2(\pi) > \bar{\theta}$  iff  $\pi \in (\pi_2, \infty)$ .<sup>14</sup> The remaining question is therefore whether it is possible that  $\pi_1 < \pi_2$ , in other words that there is a non-trivial interval  $(\pi_1, \pi_2)$  over which  $\tau_2(\pi) = \bar{\theta}$ .

Suppose for a contradiction that there is such an interval. Then, over this interval, we must have both

$$\psi(\tau_1(\pi), \pi) = F(\bar{\theta}) \quad (33)$$

(because  $\tau_1(\pi)$  is the optimal  $\theta_1$ ) and

$$\tau_2(\pi) = \bar{\theta}. \quad (34)$$

Hence

$$F(\bar{\theta}) = \psi(\tau_1(\pi), \pi) = \Psi(\bar{\theta}; \tau_1(\pi), \pi) = \Psi(\tau_2(\pi); \tau_1(\pi), \pi)$$

(by equation (33), by definition of  $\psi$  and by equation (34))

$$= \phi(\tau_1(\pi); \pi) = \frac{1}{\tau_2(\pi) - \tau_1(\pi)} \int_{\tau_1(\pi)}^{\tau_2(\pi)} G(\theta) d\theta = \frac{1}{\bar{\theta} - \tau_1(\pi)} \int_{\tau_1(\pi)}^{\bar{\theta}} G(\theta) d\theta$$

(by construction of  $\Psi$ , by definition of  $\phi$ , by equation (34) again). Multiplying through by  $\bar{\theta} - \tau_1(\pi)$ , we therefore obtain

$$\int_{\tau_1(\pi)}^{\bar{\theta}} G(\theta) d\theta = (\bar{\theta} - \tau_1(\pi)) F(\bar{\theta}).$$

Differentiating with respect to  $\pi$ , we then obtain

$$-G(\tau_1(\pi)) \tau_1'(\pi) = -\tau_1'(\pi) F(\bar{\theta})$$

or

$$(G(\tau_1(\pi)) - F(\bar{\theta})) \tau_1'(\pi) = 0.$$

---

<sup>14</sup>That  $\pi_1 \leq \pi_2$  follows at once from the fact that we cannot have  $\tau_2(\pi) < \bar{\theta}$  and  $\tau_2(\pi) > \bar{\theta}$  simultaneously.

But equation (34) implies that  $(1 + \pi) \tau_1(\pi) = \bar{\theta}$  and therefore

$$\tau_1'(\pi) = -\frac{\theta_1}{1 + \pi} \neq 0.$$

We conclude that  $G(\tau_1(\pi)) - F(\bar{\theta}) = 0$ . This, however, is impossible. For we have

$$G(\tau_1(\pi)) = \Psi(\tau_1(\pi); \tau_1(\pi), \pi) \leq \Psi(\bar{\theta}_F; \tau_1(\pi), \pi) < \Psi(\bar{\theta}; \tau_1(\pi), \pi)$$

(by construction of  $\Psi$ , by Proposition 17 and by Proposition 18)

$$= \psi(\tau_1(\pi), \pi) = F(\bar{\theta})$$

(as above). The only possible conclusion is therefore that  $\pi_1 = \pi_2$ .

#### O. COMPARATIVE STATICS WITHOUT A4

We divide our discussion into the same three cases that we considered in Section L.1, namely:

1.  $\rho < 1$  and  $G' \leq 0$  on  $(\underline{\theta}, \infty)$ ;
2.  $\rho = 1$ ;
3.  $\rho > 1$  and  $G' \geq 0$  on  $(0, \bar{\theta})$ .

Of these, the first is by far the simplest, because if  $G' \leq 0$  on  $(\underline{\theta}, \infty)$  then A4 holds.

**Proposition 54.** *Suppose that  $\rho < 1$  and  $G' \leq 0$  on  $(\underline{\theta}, \infty)$  (i.e.  $\theta_M = \underline{\theta}$ ). Then  $\underline{\tau} = \bar{\tau}$  for all  $\pi \in (0, \infty)$ . Furthermore there exists  $\pi_1 \in (0, \infty)$  such that: the maximum-penalty constraint is strictly binding for all  $\pi \in (0, \pi_1)$ ; and the maximum-penalty constraint is strictly slack for all  $\pi \in (\pi_1, \infty)$ . Finally:*

1.  $\underline{\tau} = \bar{\tau}$  is strictly decreasing on  $(0, \pi_1)$ ; and
2.  $\underline{\tau} = \bar{\tau}$  is constant on  $(\pi_1, \infty)$ . ■

In other words, for all values of the maximum penalty  $\pi \in [0, \infty)$ , there is a unique optimum within our one-parameter family. Furthermore there exists a critical level  $\pi_1$  of  $\pi$ . Below  $\pi_1$ , the maximum-penalty constraint is strictly binding and the optimal savings target is strictly increasing in  $\pi$ . Above  $\pi_1$ , the maximum-penalty constraint is strictly slack and the optimal savings target is independent of  $\pi$ .

**Proposition 55.** *Suppose that  $\rho = 1$ . Then there exists  $\pi_0 \in [0, \infty)$  such that:  $\underline{\tau} < \bar{\tau}$  for all  $\pi \in (0, \pi_0)$ ; and  $\underline{\tau} = \bar{\tau}$  for all  $\pi \in (\pi_0, \infty)$ . Furthermore there exists  $\pi_1 \in (\pi_0, \infty)$  such that: the maximum-penalty constraint is strictly binding for all  $\pi \in (0, \pi_1)$ ; and the maximum-penalty constraint is strictly slack for all  $\pi \in (\pi_1, \infty)$ . Finally:*

1.  $\underline{\tau}$  is constant, and  $\bar{\tau}$  is strictly decreasing, on  $(0, \pi_0)$ ;
2.  $\underline{\tau} = \bar{\tau}$  is strictly decreasing on  $(\pi_0, \pi_1)$ ; and
3.  $\underline{\tau} = \bar{\tau}$  is constant on  $(\pi_1, \infty)$ . ■

In other words, there are two critical levels of  $\pi$ , namely  $\pi_0$  and  $\pi_1$ . Below  $\pi_0$ , there is a continuum of optima from within our one-parameter family; and, above  $\pi_0$ , there is a unique optimum from within our one-parameter family. Below  $\pi_1$ , the maximum-penalty constraint is strictly binding; and, above  $\pi_1$ , the maximum-penalty constraint is strictly slack. Furthermore, below  $\pi_0$ : the smallest of the possible optimal savings targets is strictly increasing in  $\pi$ ; and the largest of the possible optimal savings targets is independent of  $\pi$ . Between  $\pi_0$  and  $\pi_1$ : there is only one optimal savings target, and this is strictly increasing in  $\pi$ . And, above  $\pi_1$ : there is again only one optimal savings target, and this is independent of  $\pi$ .

**Proposition 56.** *Suppose that  $\rho > 1$  and  $G' \geq 0$  on  $(0, \bar{\theta})$  (i.e.  $\theta_M = \bar{\theta}$ ). Then there there exists  $\pi_1 \in (0, \infty)$  such that: the maximum-penalty constraint is strictly binding for all  $\pi \in (0, \pi_1)$ ; and the maximum-penalty constraint is strictly slack for all  $\pi \in (\pi_1, \infty)$ . Furthermore:*

1.  $\underline{\tau}$  and  $\bar{\tau}$  are both strictly decreasing on  $(0, \pi_1)$ ;
2.  $\underline{\tau} = \bar{\tau}$  is constant on  $(\pi_1, \infty)$ . ■

In other words, there exists a critical level  $\pi_1$  of  $\pi$ . Below  $\pi_1$ : the maximum-penalty constraint is strictly binding; and the set of optimal savings targets is strictly increasing in  $\pi$ . Above  $\pi_1$ : the maximum-penalty constraint is strictly slack; and the optimal savings target is independent of  $\pi$ .

**Remark 57.** *It is interesting to compare the levels of uniqueness obtained in Propositions 54, 55 and 56. When  $\rho < 1$ , we have uniqueness for all  $\pi \in (0, \infty)$ . When*

$\rho = 1$ , a limited form of non-uniqueness can develop: there exists  $\pi_0 \in [0, \pi_1)$  such that there is non-uniqueness on  $(0, \pi_0)$  and uniqueness on  $(\pi_0, \infty)$ . And, when  $\rho > 1$ , non-uniqueness takes the form that one might expect in a convex optimization problem. However, we do at least get strict monotonicity on the whole of  $(0, \pi_1)$ .

#### P. EXISTENCE OF A FULL OPTIMUM

Suppose that self 0 is required to pick a  $B$  satisfying Constraints 1 and 2. Then the utility curve  $(u, w)$  that results will satisfy the following three conditions:

**I**  $(u, w)$  is interior, in the sense that  $u, w > U(0)$  on  $\Theta$ .

**M**  $(u, w)$  is monotonic, in the sense that  $u$  is non-decreasing and  $w$  is non-increasing.

**DE**  $(u, w)$  satisfies the differential equation  $\theta du + \beta dw = 0$ .

If  $B$  is also convex, then  $(u, w)$  will also satisfy:

**C**  $(u, w)$  is continuous.

Now, the set  $\Omega$  with which we have worked so far consists of utility curves  $(u, w)$  that satisfy I, BV, DE and C, where BV is the condition:

**BV**  $(u, w)$  is of bounded variation.

Since BV is weaker than M, this means that  $\Omega$  contains all the utility curves that can result from convex  $B$ , and more besides. We have therefore solved a relaxed version of the convex- $B$  problem. Since the solution of this relaxed problem is feasible in the convex- $B$  problem, we have therefore also solved the convex- $B$  problem. The purpose of the present section is to solve the general problem in which  $B$  is not required to be convex.

Suppose accordingly that  $\Omega$  consists of all  $(u, w) \in \mathcal{BV}(\Theta, \text{ran}(U))^2$  such that  $\theta du + \beta dw = 0$ . In other words, let  $\Omega$  consist of utility curves  $(u, w)$  that satisfy I, BV, DE but not C. Put  $X = \mathcal{BV}(\Theta, \mathbb{R})^2$ ,  $\Xi = \mathcal{BV}(\Theta, \text{ran}(U))^2$  and  $Z = \mathcal{BV}(\Theta, \mathbb{R})$ . Then the objective function  $M$  and the constraint mappings  $G_1$  and  $G_2$  continue to make sense. The analysis of Luenberger (1969) therefore shows that  $x_0 \in \Omega$  solves the problem

$$\begin{aligned} & \text{maximize} && M(x) \\ & \text{subject to} && \left\{ \begin{array}{l} x \in \Omega \\ G_1(x) \geq 0 \\ G_2(x) \geq 0 \end{array} \right\} \end{aligned}$$

iff there exist  $\lambda_1, \lambda_2 \in Z^*$  such that:

1.  $L(x_0, \lambda_1, \lambda_2) \geq L(x, \lambda_1, \lambda_2)$  for all  $x \in \Omega$ , where

$$L(x, \lambda_1, \lambda_2) = M(x) + \langle G_1(x), \lambda_1 \rangle + \langle G_2(x), \lambda_2 \rangle;$$

2.  $G_1(x) \geq 0$ ,  $\lambda_1 \geq 0$  and  $\langle G_1(x), \lambda_1 \rangle = 0$ ;
3.  $G_2(x) \geq 0$ ,  $\lambda_2 \geq 0$  and  $\langle G_2(x), \lambda_2 \rangle = 0$ .

In other words, there exists multipliers  $\lambda_1$  and  $\lambda_2$  such that: (1)  $x_0$  maximizes  $L(\cdot, \lambda_1, \lambda_2)$  over  $\Omega$ ; (2) complementary slackness holds for the first constraint; and (3) complementary slackness holds for the second constraint.

At this point, however, we encounter an obstacle. While the dual space  $\mathcal{C}(\Theta, \mathbb{R})^*$  of  $\mathcal{C}(\Theta, \mathbb{R})$  has a convenient representation as the space  $\mathcal{M}(\Theta, \mathbb{R})$ , the dual space  $\mathcal{BV}(\Theta, \mathbb{R})^*$  of  $\mathcal{BV}(\Theta, \mathbb{R})$  does not have a similarly convenient representation. This makes it difficult to use the necessity part of the Lagrangean characterization of the optimum. We can, however, still hope to use the sufficiency part.

The idea here is to note that the elements of  $\mathcal{M}(\Theta, \mathbb{R})$  can be used to induce continuous linear functionals on  $\mathcal{BV}(\Theta, \mathbb{R})$ . For example,  $\mu \in \mathcal{M}(\Theta, \mathbb{R})$  induces  $\mu_R \in \mathcal{BV}(\Theta, \mathbb{R})^*$  via the formula

$$\langle z, \mu_R \rangle = \int z_R d\mu,$$

where  $z_R$  denotes the right-continuous version of  $z$ . However, in pursuing this idea, it is important to note that  $\mu$  also induces  $\mu_L \in \mathcal{BV}(\Theta, \mathbb{R})^*$  via the formula

$$\langle z, \mu_L \rangle = \int z_L d\mu,$$

where  $z_L$  denotes the left-continuous version of  $z$ . In other words, there is no canonical association between elements of  $\mathcal{M}(\Theta, \mathbb{R})$  and continuous linear functionals on  $\mathcal{BV}(\Theta, \mathbb{R})$ .

Our plan is therefore to start from a  $\theta_1$  such that  $\Psi(\bar{\theta}; \theta_1) = F(\bar{\theta})$ , in the hope that  $\Psi(\cdot; \theta_1)$  can be used to generate multipliers that can be used in the sufficiency part of the Lagrangean characterization of an optimum. Indeed, suppose that we are given such a  $\theta_1$ . Then, bearing in mind that  $\Delta\Psi(\theta_2; \theta_1) = 0$ , we may put  $d\tilde{\Lambda}_1 = d\Psi(\cdot; \theta_1)$  on  $[\underline{\theta}, \theta_2]$ ,  $d\Lambda_2 = d\Psi(\cdot; \theta_1)$  on  $[\theta_2, \bar{\theta}]$  and  $d\Lambda_1 = \frac{1}{K'(w_0)} d\tilde{\Lambda}_1$ . Furthermore,

if we let  $\lambda_1$  and  $\lambda_2$  be the continuous linear functionals induced on  $\mathcal{BV}(\Theta, \mathbb{R})$  by  $d\Lambda_1$  and  $d\Lambda_2$  using integration with respect to the right-continuous versions of functions, then we have

$$\begin{aligned} L(x, \lambda_1, \lambda_2) &= \int \left( \theta u(\theta) + w(\theta) \right) dF(\theta) \\ &\quad + \int \left( y - C(u(\theta)) - K(w(\theta)) \right) d\Lambda_1(\theta) \\ &\quad + \int \left( b\left(\frac{\theta}{(1+\pi)\beta}\right) u(\theta) - \frac{1}{\rho} a\left(\frac{\theta}{(1+\pi)\beta}\right) - w(\theta) \right) d\Lambda_2(\theta) \end{aligned}$$

for all  $x \in X$ . Our objective is then to show that the utility curve  $x_0 = (u_0, w_0)$  associated with  $\theta_1$  maximizes  $L(\cdot, \lambda_1, \lambda_2)$ .

It suffices to show that, for all  $x_1 \in \Omega$ , the directional derivative  $\nabla_x L(x_0, \lambda_1, \lambda_2)$  of  $L$  at  $x_0$  in the direction  $x = x_1 - x_0$  is non-positive. As in Section F, we have

$$\begin{aligned} \nabla_x L(x_0, \lambda_1, \lambda_2) &= \int \left( \theta u + w \right) dF - \int \left( C'(u_0) u + K'(w_0) w \right) d\Lambda_1 \\ &\quad + \int \left( b\left(\frac{\theta}{(1+\pi)\beta}\right) u - w \right) d\Lambda_2 \\ &= \int \left( \theta u + w \right) dF - \int \left( \frac{C'(u_0)}{K'(w_0)} u + w \right) d\tilde{\Lambda}_1 \\ &\quad + \int \left( b\left(\frac{\theta}{(1+\pi)\beta}\right) u - w \right) d\Lambda_2. \end{aligned}$$

Furthermore, notwithstanding the fact that we are now working in a more general context, we can eliminate the terms  $\int w dF$ ,  $\int w d\tilde{\Lambda}_1$  and  $\int w d\Lambda_2$  using integration by parts.

Indeed, the general formula for integration by parts tells us that

$$\int_{[\underline{\theta}, \bar{\theta}]} w(\theta) dF(\theta) = [w F]_{\underline{\theta}-}^{\bar{\theta}} - \int_{[\underline{\theta}, \bar{\theta}]} F(\theta) dw(\theta) + \sum_{\theta \in [\underline{\theta}, \bar{\theta}]} \Delta w(\theta) \Delta F(\theta),$$

where

$$[w F]_{\underline{\theta}-}^{\bar{\theta}} = w(\bar{\theta}) F(\bar{\theta}) - w(\underline{\theta}-) F(\underline{\theta}-).$$

We therefore have

$$\int w dF = [w F]_{\underline{\theta}-}^{\bar{\theta}} - \int F dw + \sum \Delta w \Delta F$$

(where we have suppressed the dependence on  $\theta$  and where the domains of all integrals and sums are understood to be the whole of  $[\underline{\theta}, \bar{\theta}]$ )

$$= w(\bar{\theta}) F(\bar{\theta}) + \int F \frac{\theta}{\beta} du - \sum \frac{\theta}{\beta} \Delta u \Delta F$$

(because  $F(\underline{\theta}-) = 0$  and  $dw = -\frac{\theta}{\beta} du$ )

$$= w(\bar{\theta}) F(\bar{\theta}) + \frac{1}{\beta} \int F \theta du - \frac{1}{\beta} \sum \theta \Delta u \Delta F.$$

Moreover

$$\int F \theta du = [(F \theta) u]_{\underline{\theta}-}^{\bar{\theta}} - \int u d(F \theta) + \sum \Delta(F \theta) \Delta u$$

(applying the general formula for integration by parts to  $\int F \theta du$ )

$$= \bar{\theta} u(\bar{\theta}) F(\bar{\theta}) - \int u (\theta dF + F d\theta) + \sum \theta \Delta F \Delta u$$

(since  $F(\underline{\theta}-) = 0$ ,  $d(F \theta) = \theta dF + F d\theta$  and  $\Delta(F \theta) = \theta \Delta F$ ). Overall, then,

$$\int w dF = \left( \frac{\bar{\theta}}{\beta} u(\bar{\theta}) + w(\bar{\theta}) \right) F(\bar{\theta}) - \frac{1}{\beta} \int u (\theta dF + F d\theta).$$

By the same token, and bearing in mind that we did not use the fact that  $\Delta F = 0$  in the derivation of the previous paragraph, we have

$$\int w d\tilde{\Lambda}_1 = \left( \frac{\bar{\theta}}{\beta} u(\bar{\theta}) + w(\bar{\theta}) \right) \tilde{\Lambda}_1(\bar{\theta}) - \frac{1}{\beta} \int u (\theta d\tilde{\Lambda}_1 + \tilde{\Lambda}_1 d\theta)$$

and

$$\int w d\Lambda_2 = \left( \frac{\bar{\theta}}{\beta} u(\bar{\theta}) + w(\bar{\theta}) \right) \Lambda_2(\bar{\theta}) - \frac{1}{\beta} \int u (\theta d\Lambda_2 + \Lambda_2 d\theta).$$

We therefore have

$$\nabla_x L(x_0, \lambda_1, \lambda_2) = \int u du^* + w(\bar{\theta}) r^*$$

where, as in section F above,

$$\begin{aligned} du^* &= -\frac{1}{\beta} \left( (1 - \beta) \theta dF + F d\theta \right) + \frac{1}{\beta} \left( \left( \theta - \beta \frac{C'(u_0)}{K'(w_0)} \right) d\tilde{\Lambda}_1 + \tilde{\Lambda}_1 d\theta \right) \\ &\quad + \frac{1}{\beta} \left( \left( \theta + \beta b \left( \frac{\theta}{(1+\pi)\beta} \right) \right) d\Lambda_2 + \Lambda_2 d\theta \right) + \frac{\bar{\theta}}{\beta} \left( F(\bar{\theta}) - \tilde{\Lambda}_1(\bar{\theta}) - \Lambda_2(\bar{\theta}) \right) dI, \\ r^* &= F(\bar{\theta}) - \tilde{\Lambda}_1(\bar{\theta}) - \Lambda_2(\bar{\theta}) \end{aligned}$$

and  $I$  is the distribution function of the unit mass at  $\bar{\theta}$ . Finally, by construction of  $\tilde{\Lambda}_1$  and  $\Lambda_2$ , we have  $u^* = 0$  and  $r^* = 0$ . So in fact  $\nabla_x L(x_0, \lambda_1, \lambda_2) = 0$ . In particular, the utility curve  $x_0 = (u_0, w_0)$  associated with  $\theta_1$  does indeed maximize  $L(\cdot, \lambda_1, \lambda_2)$ .

**Remark 58.** Great care is needed in choosing the space  $Z$ . One possible choice is  $\mathcal{C}(\Theta, \mathbb{R})$ , the space of all continuous functions on  $\Theta$  endowed with the sup norm. This choice has the advantage that there is a convenient representation for  $Z^*$ . However, it also requires that  $\Omega \subset \mathcal{C}(\Theta, \mathbb{R})$ , and this is not an economically reasonable restriction. Another possible choice is  $\mathcal{B}(\Theta, \mathbb{R})$ , the space of all bounded functions on  $\Theta$  endowed with the sup norm. This choice has the advantage that it includes all economically relevant utility curves. Unfortunately, it leads to a different problem: the measures  $d\Lambda_1$  and  $d\Lambda_2$  associated with  $\Psi(\cdot; \theta_1)$  do not induce continuous linear functionals on  $\mathcal{B}(\Theta, \mathbb{R})$ , since functions in  $\mathcal{B}(\Theta, \mathbb{R})$  are not in general measurable. The results of Luenberger (1969) do not therefore apply. Our solution to this double problem is to use  $\mathcal{BV}(\Theta, \mathbb{R})$ . This space is big enough to include all economically relevant utility curves, but small enough that  $d\Lambda_1$  and  $d\Lambda_2$  can be used to induce continuous linear functionals on it.

## Q. DISTRIBUTIONS

**Q.1. Beta Distribution.** The density of the generalization of the Beta that we consider is proportional to

$$(x - a)^{\zeta-1} (b - x)^{\eta-1}$$

on the interval  $(a, b)$ , where  $0 < a < b$  and  $\zeta, \eta > 0$ . It is unbounded at  $a$  if  $\zeta < 1$ , in which case we require that  $\underline{\theta} \in (a, b)$  in order to ensure that A1 is satisfied, and unbounded at  $b$  if  $\eta < 1$ , in which case we require that  $\bar{\theta} \in (a, b)$  in order to ensure that A1 is satisfied.

There are then four main cases. Three of the cases are easy to describe:

**Case 1** if  $\zeta > 1$  and  $\eta \geq 1$  then A3 is satisfied for all choices of  $\underline{\theta}, \bar{\theta} \in (a, b)$ ;

**Case 2** if  $\zeta > 1$  and  $\eta < 1$  then A3 is again satisfied for all choices of  $\underline{\theta}, \bar{\theta} \in (a, b)$ , albeit for somewhat different reasons;

**Case 4** if  $\zeta < 1$  and  $\eta < 1$ , then A3 is violated for some choices of  $\underline{\theta}, \bar{\theta} \in (a, b)$ .

Case 3 is more involved. If  $\zeta < 1$  and  $\eta \geq 1$ , then A3 is satisfied for all choices of  $\underline{\theta}, \bar{\theta} \in (a, b)$  iff

$$\frac{(\sqrt{1-\zeta} + \sqrt{\eta-1}\sqrt{\frac{a}{b}})^2}{1 - \frac{a}{b}} \geq 1 + \frac{1}{1-\beta}. \quad (35)$$

As this inequality makes clear, A3 is more likely to be satisfied if: either (i)  $\zeta$  is close to 0 (i.e. the spike at  $a$  is very pronounced); or (ii)  $\eta$  is large (i.e. the density decays very quickly towards  $b$ ); or (iii)  $\frac{a}{b}$  is close to 1 (i.e. the density is concentrated in a narrow band).<sup>15</sup> It is also worth noting that, as  $\frac{a}{b} \rightarrow 0$ , the left-hand side of (35) converges to  $1 - \zeta < 1$ . Hence A3 is violated for some choices of  $\underline{\theta}, \bar{\theta} \in (a, b)$  when  $\zeta < 1$  and  $\frac{a}{b}$  is small. This is in striking contrast with the standard case studied in both Rice and Hogg et al. In that case A3 is satisfied for all  $\underline{\theta}, \bar{\theta} \in (a, b)$  when  $\zeta < 1$  and  $\frac{a}{b} = 0$ .

Note finally that the right-hand side of (35) is strictly increasing in  $\beta$ . Hence, if we fix a distribution for which  $\zeta < 1$  and  $\eta \geq 1$ , then the conclusion is that A3 will be satisfied provided that  $\beta$  is *far enough below* 1. I.e. A3 is more likely to be satisfied when the decision maker is *more* time-inconsistent.

**Q.2. Cauchy Distribution.** The density of the general form of the Cauchy distribution is proportional to

$$\left(1 + \left(\frac{x - \mu}{\sigma}\right)^2\right)^{-1}$$

on  $\mathbb{R}$ , where  $\mu \in \mathbb{R}$  is a location parameter and  $\sigma > 0$  is a scale parameter. This distribution satisfies A3 for all  $\underline{\theta}, \bar{\theta} \in (0, \infty)$  iff

$$\frac{\mu}{\sigma} \leq \sqrt{\frac{1 - (1 - \beta)^2}{(1 - \beta)^2}}. \quad (36)$$

In other words, taking  $\beta$  as given, A3 is satisfied iff the distribution is not located too far to the right. If (36) does not hold then, for some choices of  $\underline{\theta}, \bar{\theta} \in (0, \infty)$ ,  $G$  is

---

<sup>15</sup>For the purposes of the present discussion,  $\zeta \in (0, 1)$ ,  $\eta \in [1, \infty)$  and  $\frac{a}{b} \in (0, 1)$ .

first increasing (at  $\underline{\theta}$ ), then decreasing, then increasing again, then finally decreasing again (at  $\bar{\theta}$ ).

We can also make  $1 - \beta$  the subject of the inequality (36). Doing so, we find that A3 is satisfied for all  $\underline{\theta}, \bar{\theta} \in (0, \infty)$  iff

$$1 - \beta \leq \left(1 + \frac{\mu^2}{\sigma^2}\right)^{-\frac{1}{2}}.$$

In other words, taking the parameters  $\mu$  and  $\sigma$  of the Cauchy distribution as given, A3 is satisfied iff  $\beta$  is sufficiently close to 1. I.e. A3 is more likely to be satisfied when the decision maker is less time-inconsistent.

**Q.3. Log-Gamma Distribution.** The density of the Log-Gamma distribution is proportional to

$$x^{-\frac{\eta+1}{\eta}} (\log(x))^{\zeta-1}$$

on  $(1, \infty)$ , where  $\zeta, \eta > 0$ . It is unbounded at 1 if  $\zeta < 1$ , in which case we require that  $\underline{\theta} > 1$  in order to ensure that A1 is satisfied. It violates A3 for some choices of  $\underline{\theta}, \bar{\theta} \in (1, \infty)$  iff  $\zeta < 1$  and  $\eta > 1 - \beta$ . In other words, taking  $\beta$  as given, A3 is violated iff there is a singularity at 1 and the rate of decay at  $\infty$  is sufficiently slow.

Note finally that, if we fix a distribution for which  $\zeta < 1$ , then the conclusion is that A3 will be satisfied provided that  $\beta$  is *far enough below* 1. I.e. A3 is more likely to be satisfied when the decision maker is *more* time-inconsistent.

**Q.4. Pareto Distribution.** The density of the Pareto type II distribution is proportional to

$$\left(1 + \frac{x - \mu}{\sigma}\right)^{-\zeta-1}$$

on  $(\mu, \infty)$ , where  $\mu \in \mathbb{R}$  is a location parameter,  $\sigma > 0$  is a scale parameter and  $\zeta > 0$  is a shape parameter. It violates A3 for some choices of  $\underline{\theta}, \bar{\theta} \in (\mu, \infty)$  iff

$$\zeta < \frac{1}{1 - \beta} \tag{37}$$

and

$$\frac{\mu}{\sigma} > \frac{1}{\zeta + 1} \left(1 + \frac{1}{1 - \beta}\right). \tag{38}$$

ONLINE APPENDICES FOR “SELF CONTROL AND COMMITMENT:  
CAN DECREASING THE LIQUIDITY OF A SAVINGS ACCOUNT INCREASE DEPOSITS?”

In other words, it violates A3 iff its right-hand tail is sufficiently fat and, taking the fatness of the tail as given, it is located sufficiently far to the right. In particular, if  $\frac{\mu}{\sigma} \leq 1$ , then the Pareto type II distribution satisfies A3 for all  $\underline{\theta}, \bar{\theta} \in (\mu, \infty)$ . For in that case: either (i)  $\zeta \geq \frac{1}{1-\beta}$  and therefore (37) is violated; or (ii)  $\zeta < \frac{1}{1-\beta}$ , in which case  $\frac{1}{\zeta+1} \left(1 + \frac{1}{1-\beta}\right) > \frac{1}{\zeta+1} (1 + \zeta) = 1 \geq \frac{\mu}{\sigma}$  and therefore (38) is violated.

We can also make  $1 - \beta$  the subject of these inequalities. Doing so, we find that A3 is violated for some  $\underline{\theta}, \bar{\theta} \in (\mu, \infty)$  iff

$$\left(\frac{\mu}{\sigma}(\zeta + 1) - 1\right)^{-1} < 1 - \beta < \zeta^{-1}.$$

In particular, if  $\frac{\mu}{\sigma} > 1$  and  $\zeta > 1$  (so that  $\left(\frac{\mu}{\sigma}(\zeta + 1) - 1\right)^{-1} < \zeta^{-1} < 1$ ), then A3 is satisfied iff  $\beta$  is either close enough to 1 or *far enough below* 1.