Supplier-induced demand as strategic framing

Kris De Jaegher
Tjalling C. Koopmans Research Institute
Utrecht University School of Economics, The Netherlands

Abstract
This paper develops a model of persuasive demand inducement, where a physician frames the healthcare decision faced by a patient with prospect-theoretic preferences, such as to make the patient see the decision from the perspective of a particular reference point. We derive the reference point that maximises the probability that the patient buys treatment under two scenarios, namely the decision to buy a safe preventive treatment rather than no treatment, and the decision to buy a risky curative treatment rather than no treatment. It is shown that in the curative scenario, the physician sets a high reference point, whereas in the preventive scenario, the physician sets a low reference point is the patient's loss aversion is small, and a high reference point when the patient's loss aversion is large. Furthermore, we derive how the patient's degree of risk aversion, how the manner in which this degree of risk aversion changes with wealth, and how the level of loss aversion, affect the reference point that is optimal from the perspective of the physician.

Key words: supplier-induced demand, prospect theory, framing

JEL-classification: D82, I11.

1. Introduction

One of the most popular themes in health economics is the *supplier-induced demand* (*SID*) *hypothesis*, stating that physicians whose income gets under pressure (e.g. because of physician entry) are able to create demand for their own services (Evans, 1974; for an overview, see e.g. Peacock and Richardson, 2007). From a theoretical point of view, the model underlying this hypothesis is simply a version of the Dorfman and Steiner (1954) model of advertising (Stano, 1987). By taking persuasive effort, the physician is able to shift the demand. A weakness of this model is that it does not explain how physicians can persuade patients, and that there is no limit on how much demand physicians can induce.

To fill the theoretical gap, part of the health economics literature takes a different approach to SID, and starts from the premise of a rational patient with less information than the physician (Dranove, 1988; Labelle et al., 1994; Calcott, 1999; De Jaegher and Jegers, 2000; De Jaegher and Jegers, 2001; Xie et al., 2006; De Jaegher, 2012). The physician may give wrong information, but the patient anticipates this, and if the physician overprescribes too much the patient refuses treatment or consults another physician. While an attractive feature of there models is that there now is a constraint on the extent to which physicians can induce demand, at the same time they are remote from the original concept of persuasive SID.

The current paper shows that behavioural economics (for health-economics applications see Frank, 2004; Barigozzi and Levaggi, 2008) can contribute to constructing a model of persuasive SID. In this model, the physician's information is always truthful, but the physician persuades the patient by *framing* the information. Our starting point is a modified version of an experiment due to Tversky and Kahneman (1981) known as the Asian Disease Problem. In a variant of this problem, assume that a patient is confronted with the following situation:

Your life expectancy was initially 80 years, but you have now been diagnosed with a rare disease. Without treatment, this rare disease reduces your life expectancy to 74 years. You can choose from two alternative treatments.

In the first scenario, the patient is given a choice between treatments A and B:

Scenario 1. *Without treatment, your life expectancy is 74 years. If you adopt treatment A, your life expectancy improves by 2 years. If you adopt treatment B, with probability 1/3 it improves by 6 years, and with probability 2/3 it does not improve."*

In the second scenario, the patient is given a choice between treatments C and D:

Scenario 2. *Without the rare disease, your life expectancy would have been 80 years. If you adopt treatment C, your life expectancy will be four years less. If you adopt treatment D, with probability 1/3 your life expectancy will not be less, and with probability 2/3 it will be six years less.*

Some reflection teaches us that the first and the second scenario are logically equivalent. The only difference between the first scenario and the second scenario is that the outcomes of the treatments have been *framed* in a different manner. In a meta-analysis of framing studies (Kühberger, 1998), framing by formulating the decision maker's reference point in different manners is found to have a systematic and significant effect. Concretely, in a gains frame such as Scenario 1, participants in experiments typically behave in a risk averse manner, and choose the safe treatment A. In a loss frame such as Scenario 2, however, participants typically behave in a risk loving manner. A physician who seeks to induce the safe rather than the risky treatment should therefore frame the patient's decision problem in terms of gains, presenting Scenario 1 rather than Scenario 2, while the opposite is true for a physician who seeks to induce the risky rather than the safe treatment. Applied to the question of how healthy behaviour can be induced, the simple conclusion that gains framing is more effective in inducing behaviour that avoids risk whereas loss framing is more effective in inducing behaviour that involves risk, can be found in Rothman and Salovey (1997) and Rothman et al. (2006), without any modelling effort.

Yet, the physician need not limit herself to either scenarios with pure gains or with pure losses, as illustrated by a third logically equivalent scenario:

Scenario 3. *If you adopt treatment E, your life expectancy is 76 years. If you adopt treatment F, with probability 1/3 you do four years better, and with probability 2/3 you do two years worse.*

In Scenario 3, the risky treatment *F* is framed as a mix of gains and losses. It is well-known that in lotteries involving mixed gains and losses, losses have more impact than gains, a phenomenon referred to as *loss aversion*. Because of loss aversion the patient may be risk averse over mixed gains and losses, and possibly more risk averse than in the pure gains region. A physician who wants to induce the safe rather than the risky treatment may then be better off when framing Scenario 3 rather than Scenario 1. More generally, many more logically equivalent scenarios can be envisaged: e.g., if average life expectancy is larger than 80 years, this may be framed as the reference point, where all outcomes are then seen as losses. As the level of risk aversion need not be constant for pure gains and pure losses, an open question is then what is precisely the optimal reference point of a physician who wants to induce a safe action, and one who wants to induce a risky action.

The paper is structured as follows. Section 2 contains the model, where the patient is assumed to have prospect-theoretic preferences. In Section 3, we find the physician's optimal frame depending on the precise form of the patient's preferences. We end with a discussion in Section 4.

2. The model

The physician-patient game proceeds as follows. The uninformed patient (he) visits his informed physician (she). At stage 1, the physician observes the patient's health status, and by assumption truthfully communicates to the patient all available objective information, meaning the actions available to the patient, and for each action the corresponding objective density function over outcomes. Only a risky action and a safe action are available. The safe action always yields outcome *M*. The risky action yields outcome *H* with objective probability *p*, and outcome *L* with objective probability $(1-p)$, where $H > M > L$. While the physician by assumption reveals the aforementioned objective information to the patient, she can still strategically set the patient's reference point *R*. This *R* can be chosen from a set $[R_{\min}, R_{\max}]$, where $R_{\min} < L$ and $R_{\max} > H$. The patient sees outcome $X = L, M, H$ as a gain $(X - R)$ if $X \geq R$, and as a loss $(R - X)$ if $X < R$. At stage 2, following the principle of informed consent, after hearing the objective information provided by the physician and given the reference point *R* set by the physician, the patient calculates his expected psychic valuation (defined below, henceforth in short valuation) for each action, and picks the action with the highest expected valuation. Crucially, by assumption, the patient does not consider expected valuations from the perspective of reference points other than the reference point set by the physician. Our patient is thus not aware that the physician strategically sets his reference point, and that if he would consider the decision problem from the perspective of another reference point, he could come to a different decision. This makes our physician-patient game a behavioural game rather than a standard game, as the patient does not act as a standard rational player.

We now look at the physician's and the patient's payoffs. The physician always prefers the patient to get treatment rather than not to get treatment. This translates into two possible scenarios for the game. In the *preventive scenario*, the safe action is interpreted as getting a costly treatment that prevents the patient from ever getting ill, and the risky action is interpreted as not getting any treatment. If the patient does not get ill, not getting any treatment leads to a high outcome (*H*), as the cost of treatment is then also avoided. If the patient does get ill, not getting treatment leads to illness and a low outcome (*L*). Getting the treatment leads to a safe, intermediate outcome (*M*).

In the *curative scenario*, the risky action is interpreted as getting treatment, and the safe action is interpreted as not getting any treatment. Not getting the treatment means that the patient is never cured, leading to a safe intermediate outcome (*M*). The costly treatment either cures the patient, leading to a high outcome (*H*), or does not cure the patient, leading to a low outcome (*L*), as the patient is then not cured but still incurs the cost of treatment.

We now specify the patient's valuation function, in accordance with prospect theory. The patient does not directly value the absolute outcomes *L*, *M* and *H*, but instead values these outcomes as gains or losses with respect to the reference point *R* set by the physician, where the patient thinks differently about gains and losses. For given *R*, the patient's valuation function *f*(*X*) of any outcome *X* takes the following form (Tversky and Kahneman, 1992):

$$X \geq R: f(X) = v(X - R)$$
$$X < R: f(X) = -\lambda v(R - X)$$
with $v' > 0, v'' < 0, v(0) = 0$ and $\lambda > 1$. (1)

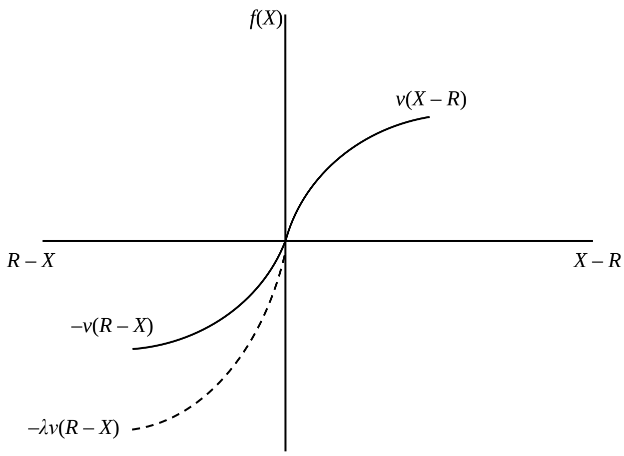

Figure 1 Psychic valuation function

The valuation function is sketched in Figure 1, and is assumed to be increasing ($v'(.) > 0$) and concave ($v''(.) < 0$). By using function *v*(.) for the psychic valuation in the gains region, and the negative of this function for the valuation in the loss region, it is assumed that the patient is risk averse in the gains region and risk loving in the loss region (the so-called *reflection effect* [1]), in line with experimental observations. The patient therefore has diminishing sensitivity to small changes in the outcome, the further away from the reference point.

λ is the patient's degree of *loss aversion*: by including in the valuation function a coefficient $\lambda > 1$ for losses (dashed part of the valuation function in Figure 1), one adapts the model to the empirical observation that decision makers with prospect-theoretic preferences care more about losses (say, a loss of one year in life expectancy) than about equally-sized gains (say, a gain of one year in life expectancy).

We finally assume that any patient *i* does not use the objective probabilities to calculated his expected valuation, but instead weighted probabilities, according to the probability weighting function $\pi_i(p)$, where *p* is any given objective probability. The weighted probabilities may differ from the objective probabilities for the following two reasons. *First*, following prospect theory, the patient may apply subjective weights to the objective probabilities. In particular, it is known that decision makers typically overweigh small probabilities (and thus underweight large probabilities). The weighting function $\pi_i(p)$ has a positive slope, but is such that there is a critical *p** such that for any $p < p^*$, $\pi_i(p) > p$ (overweighting), and for any $p > p^*$, $\pi_i(p) < p$ (with $\pi_i(p^*) = p^*$). *Second*, the probability of success and failure may be affected by personal characteristics of the patient, known to the patient but not to the physician. With some abuse of terminology, the objective probability of success is then interpreted as the average probability of success across patients, and the weighted probability is determined by the patient's deviation from the average probability.

It follows now that a patient *i* prefers the risky action if

$$\pi_i(p_L) f(L) + \pi_i(p_H) f(H) \geq f(M)$$
$$\Leftrightarrow$$
$$\alpha = \frac{f(H) - f(M)}{f(M) - f(L)} \geq \frac{\pi_i(p_L)}{\pi_i(p_H)} = o_i(p_L, p_H) \tag{2}$$

and prefers the safe action when the inequality sign in (2) is reversed. α denotes the benefit of getting the high outcome rather than the intermediate outcome, relative to the benefit of getting the intermediate outcome rather than the low outcome, and is thus an inverse measure of the degree of concavity of the valuation function, and at the same time a measure of the patient's risk tolerance. $o_i(p_L, p_H)$ denotes the patient's weighted odds of failure. Given the patient's prospect-theoretic preferences in (1), α is a function of *R*, denoted as $\alpha(R)$. It follows that for a patient with given $o_i(p_L, p_H)$, in the curative treatment scenario the physician should choose any *R* such that α exceeds $o_i(p_L, p_H)$, and in the preventive treatment scenario should choose any *R* such that α does not exceed $o_i(p_L, p_H)$.

We obtain more fine-tuned results, because we take the realistic assumption that the physician faces a distribution of patients. In particular, we assume all patients to have identical valuation functions *f*(.), and therefore identical α in (2), but to have different probability weighting functions. Each patient observes his own probability weighting function, but the physician only knows the distribution of such functions across patients.

[1] The reflection effect is in line with the well-known psychological principle (Weber-Fechner law, see Thaler, 1999, p. 185) stating that, formulated in health terms, the difference between a gain (respectively loss) of 1 year in life expectancy and a gain (loss) of 2 years in life expectancy appears larger than the difference between a gain (loss) of 10 years in life expectancy and a gain (loss) of 11 years in life expectancy.

Formally, we then add a stage 0 preceding stage 1 to the game, where Nature decides on the individual patient's probability weighting function $\pi_i(.)$. From the perspective of the physician, the patient's subjective odds of failure are now distributed over $[o_{\min}, o_{\max}]$. In the curative scenario, for any given α, all patients with $\alpha \geq o$ choose the curative treatment, whereas all patients with $\alpha < o$ choose no treatment. The physician therefore chooses the optimal reference point, denoted R^C, which maximises $\alpha(R)$ (= maximises risk tolerance). In the preventive scenario, for any given α, all patients with $\alpha \leq o$ choose the preventive treatment, whereas all patients with $\alpha > o$ choose no treatment. The physician now chooses the optimal reference point, denoted R^P, which minimises $\alpha(R)$ (= minimise risk tolerance).

3. Results

The main results are contained in Proposition 1. The technical proof is in the Appendix, but the intuition can be understood separately. As seen above, in order to induce the risky curative treatment (safe preventive treatment), the physician should maximise (minimise) the patient's risk tolerance. Risk tolerance α as a function of the reference point R is represented in Figures 2-4, where the solid curve represents the case without loss aversion ($\lambda = 1$), whereas the dashed curve and dash-dotted curves represent low (λ_L) and high (λ_H) loss aversion. As the patient is risk averse (risk loving) for pure gains (pure losses), risk tolerance is low (high) for low (high) reference points. For intermediate reference points, where the patient perceives mixed gains and losses, he is more risk averse the more loss averse he is, as loss aversion creates a kink in the valuation around the reference point. It follows that, roughly, when inducing the risky curative treatment the physician should set a high reference point, and, when inducing the safe preventive treatment, the physician should set a low reference point for low loss aversion, but an intermediate reference point for high loss aversion.

We now proceed with a detailed intuition. Risk tolerance is the inverse of risk aversion, and a standard measure of the latter is absolute risk aversion (henceforth ARA, equal to $-f''/f'$). As an increase in the patient's reference point is the analogue of a decrease in wealth, the manner in which R affects α at first sight depends on whether the patient has a DARA (decreasing), CARA (constant), or IARA (increasing ARA) valuation. In the pure gains and pure loss region, this is indeed the case. With CARA, α is low and flat for $R \leq L$, and is high and flat for $R \geq H$ (Figure 2). With DARA (Figure 3), as an increase in R is the analogue of a decrease in wealth, and as in the gains region a decrease in wealth means higher risk aversion and therefore lower risk tolerance, $\alpha(R)$ decreases in R for $R \leq L$. Because of the symmetry imposed by the reflection effect, in the loss region risk lovingness is higher for higher wealth, so that higher R means lower risk tolerance and lower $\alpha(R)$. By the same reasoning, with IARA (Figure 4), $\alpha(R)$ increases in R in both the gains and the loss region.

Strictly speaking, the ARA is a measure of local risk aversion, defined for infinitesimal gambles. While the risky action of our patient is not an infinitesimal gamble (L and H are not arbitrarily close to M), in the pure gains and pure loss region, results still depend only on whether we have DARA, CARA or IARA. For R such that the patient perceives both gains and losses, however, we need a measure of global risk aversion. An inverse measure of global risk aversion is found precisely in α, as it is an inverse measure of global concavity. Leaving out loss aversion for the moment and focusing purely on the reflection effect, the following is now clear. First, the slope of $\alpha(R)$ just above L is the same as just below L; the same applies for the slopes around M and H. Furthermore, it is clear that an R between L and H exists such that R is globally risk neutral, meaning that while α is low for $R \leq L$ and high for $R \geq H$, it is intermediate for intermediate R. E.g., if M happens to lie exactly between L and H, the patient is globally risk neutral for $R = M$. Finally, $\alpha(R)$ is increasing R approaching $R = M$ from the right. To see why, note that for $R = M$, a slight increase in R has two effects. Because of diminishing sensitivity, first, the difference in the valuation of M and L becomes smaller. Second, the difference in the valuation of H and M becomes larger; this effect is further reinforced because the convex region of the valuation function is entered.

For CARA and IARA, these results suggest a monotonous increase from low to high risk tolerance in the region of mixed gains and losses, as depicted by the solid curves in Figures 2 and 4. If follows that with CARA the physician should set R^P in any way such that $R^P \leq L$ and R^C in any way such that $R^C \geq L$; with IARA the physician should set $R^P = R_{\min}$ and $R^P = R_{\max}$. For DARA, the picture is more complex. By the previous, α is now low and decreasing in R around L, is intermediate and increasing in R around M, and is high but decreasing in R around H. This suggests that for DARA, α as a function of R takes the form of the solid curve in Figure 3. It follows that $L < R^P < M$, and $M < R^C < H$. As these results apply in the absence of loss aversion, it is intuitive that similar results apply for sufficiently low loss aversion, which explains in Proposition 1 for IARA the cases (i)(a) and (ii)(d), and for DARA the cases (i)(b) and (ii)(a). $\alpha(R)$ with low loss aversion (λ_L) is represented by the dashed curves in Figures 2-4.

When adding loss aversion to the analysis, the following facts are clear. First, loss aversion does not impact $\alpha(R)$ in the pure gains or pure loss region. Thus, when we add to Figures 2-4 the dashed and dash-dotted curves for $\alpha(R)$ with loss aversion (λ_L and λ_H), for $R \leq L$ and $R \geq H$, they coincide with the solid curves for $\alpha(R)$ without loss aversion. This is because loss aversion only has an impact for mixed gains and losses. Second, as we let loss aversion become very large, global risk tolerance at $R^P = M$ becomes minimal, as the kink in the valuation function around the reference point becomes extreme. It follows that as long as loss aversion is sufficiently high, $R^P = M$ (see Proposition 1(i)(c)). Third, the slope of α in R approaching M from the right continues to be positive when there is loss aversion. Loss aversion has the additional effect that the difference between the valuation of M and L is further increased as the loss region is entered.

Fourth, the kink in the valuation function around the reference point is translated into a kink in α for R around L, M and H. Starting at $R = L$, as R is increased and the loss region is entered, the difference in the valuation of M and L increases because of loss aversion. By (2), it follows that the slope of α is smaller approaching L from the right than from the left. This means that at $R = L$, $\alpha(R)$ never has a peak for CARA or DARA, but does have a peak for IARA when loss aversion is sufficiently high. Starting at $R = M$ and lowering R, the difference in valuation between M and L decreases as one partially enters the gains region. For sufficiently large loss aversion, this has a large impact on the denominator of α (see (2)), so that for sufficiently large loss aversion, α decreases in R approaching M from the left, and has a peak at $R = M$. Finally, starting at $R = H$ and lowering R, the presence of loss aversion decreases the difference in valuation of H and M. For sufficiently large loss aversion, this has a large impact on the numerator of α (see (2)). It follows that loss aversion makes the slope of α larger approaching H from the right than from the left. This means that at $R = H$, $\alpha(R)$ never has a peak for CARA or IARA, but does have a peak for DARA when loss aversion is sufficiently high.

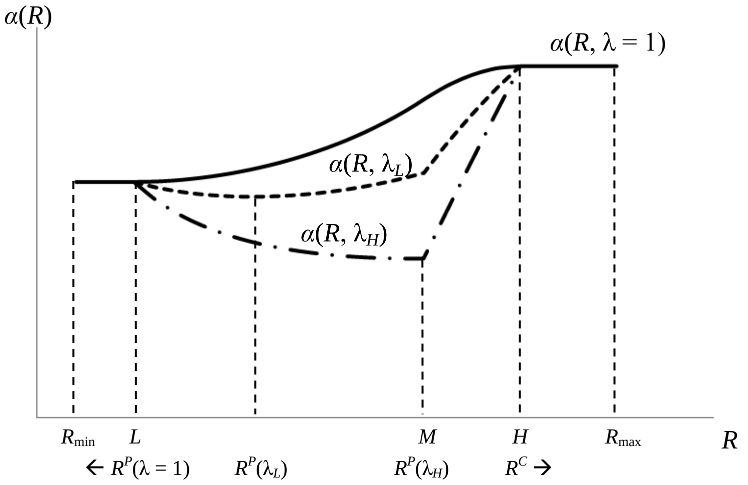

Figure 2. Global risk tolerance (α) as function of patient's reference point (R) for CARA valuation function. Solid curve: no loss aversion (λ = 1); dashed curve: low loss aversion (λ_L); dash-dotted curve: high loss aversion (λ_L). R^C (R^P) indicates optimal reference point in curative (preventive) scenario.

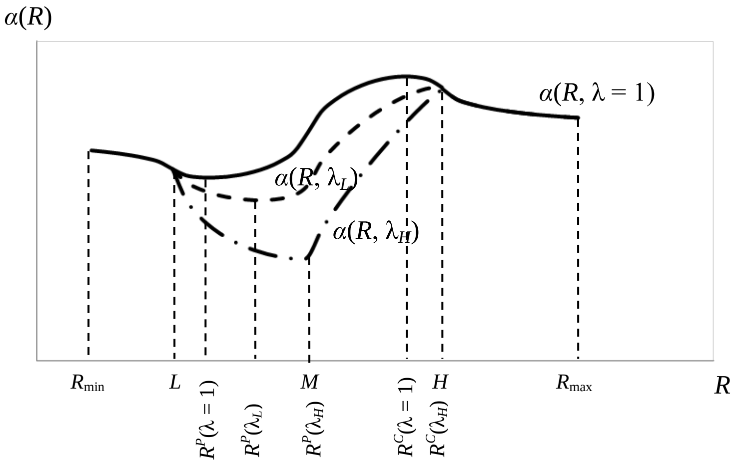

Figure 3. Idem as Figure 2, but for DARA valuation function.

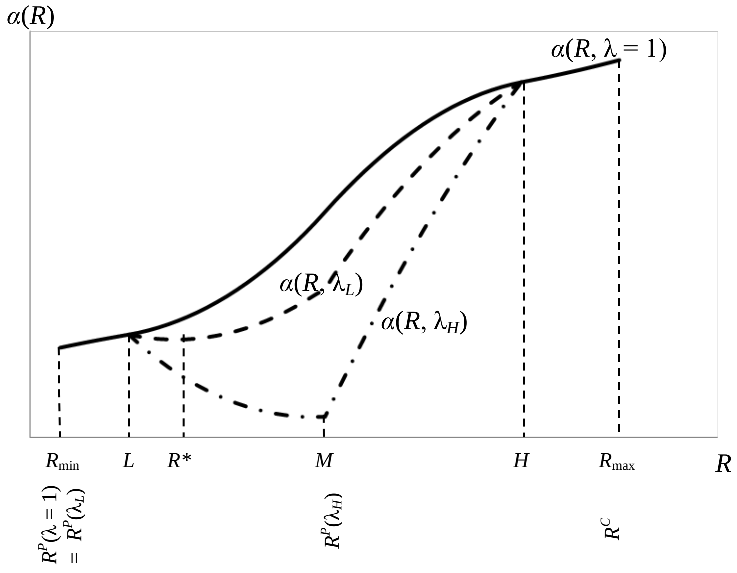

Figure 4. Idem as Figure 2, but for IARA valuation function.

All results in Proposition 1 are now readily obtained when assuming that, first, the sign of the slope of $\alpha(R)$ changes at most once in each of the ranges $L < R < M$ and $M < R < H$, and second, that the patient is everywhere risk averse in the range $L < R < M$.[2] For CARA, as soon as there is loss aversion, $R^P > L$ (dashed and dash-dotted curves in Figure 2); when loss aversion is sufficiently high, $R^P = M$ (dash-dotted curve). With or without loss aversion, it remains the case that R^C can be put anywhere equal to or larger than *H*. For both DARA and IARA, if loss aversion is small, the results are qualitatively the same as without loss aversion; if loss aversion is sufficiently large, $R^P = M$. For DARA, for sufficiently large loss aversion, $\alpha(R)$ has a peak at *H*, so that $R^C = H$ (see dashed and dash-dotted curves in Figure 3). For IARA, for sufficiently large loss aversion, $\alpha(R)$ has a peak at *L*. Yet at $R = H$, the patient is risk averse rather than risk loving, so that it continues to be the case that R^C should be as large as possible ($R^C = R_{\max}$). Given the peak at $R = L$, as long as loss aversion is not so large as to minimise $\alpha(R)$ for $R = M$, the level of R^P depends on the relative height of $\alpha(R_{\min})$. For instance, in the dashed curve in Figure 4, $\alpha(R)$ has a local minimum at R^*; however, as $\alpha(R_{\min}) < \alpha(R^*)$, we have $R^P = R^*$. With a large level of loss aversion, in the dash-dotted curve in Figure 4, $\alpha(R)$ has a local minimum at *M*; as $\alpha(R_{\min}) > \alpha(M)$, we have $R^P = M$.

Proposition 1. Assume that $\alpha(R)$ is single-peaked in both the range $L < R < M$ and the range $M < R < H$, and that the patient is risk averse in the range $L < R < M$. Then the physician's optimal reference point is the following:

(i) *Preventive scenario*: $R^P \le M$, and in particular:

a) $R^P = R_{\min}$: IARA with $\alpha(R_{\min})$ sufficiently low and/or loss aversion sufficiently low;
b) a unique R^P such that $L < R^P < M$: sufficiently low loss aversion, and either CARA/DARA, or IARA with $\alpha(R_{\min})$ sufficiently high;
c) $R^P = M$: loss aversion sufficiently high, and either CARA/DARA, or IARA with $\alpha(R_{\min})$ sufficiently high.

(ii) *Curative scenario*: $R^C > M$, and in particular:

a) a unique R^C such that $M < R^C < H$: DARA with loss aversion sufficiently low;
b) $R^C = H$: DARA with los aversion sufficiently high;
c) *any* R^C such that $H \le R^C \le R_{\max}$: CARA;
d) $R^C = R_{\max}$: IARA.

Proof: See Appendix.

We next look in Proposition 2 at how the optimal reference point is affected by the patient's degree of risk aversion and loss aversion for two valuation functions commonly used in applied theory, namely the CARA valuation function and the CRRA valuation function[3]

---

[2] These assumptions can be checked to be valid for the HARA valuation function (of which CARA and CRRA (constant relative risk aversion) valuation functions are special cases).

[3] For the CRRA valuation function, we only consider $v(.) = (.)^{1-RRA}$ for $0 \le RRA < 1$. This excludes higher levels of *RRA* (Wakker, 2008), for which the valuation function takes the form $v(.) = -(.)^{1-RRA}$ for $RRA > 1$ ($v(.) = \ln(.)$ for $RRA = 1$). The problem is that $v(0)$ is then not defined, which is problematic as what happens at the reference point plays a crucial role in our analysis. Still, in the overview of estimates of the

(constant relative risk aversion (henceforth RRA), where the RRA equals $-Xf''/f'$), where the latter has the DARA form. For both valuation functions, R^P increases in λ (see Proposition 2(i)) because a physician who wants to discourage a patient from taking the risk of not buying preventive treatment, has more reason to set the reference point such that losses are perceived when the patient is more loss averse, as the extent to which the patient is risk averse for mixed gains and losses is now larger (see Figures 2 and 3). For the same reason, the physician who wants to induce a risky curative treatment rather than no treatment has more incentives to stay out of the gains region when the patient is more loss averse. For CRRA, this means again an increase in R^C (see Figure 3 and Proposition 2(i)(a)), for CARA it does not change R^C as in this case there is no incentive to enter the loss region even without loss aversion (see Figure 2 and Proposition 2(i)(b)).

We next look at the effect of changes in ARA and RRA, and look first at the effect on R^P. In the absence of loss aversion and focusing purely on the reflection effect, with CARA, as shown in Figure 2 by the solid curve, the physician sets $R^P = L$, i.e. frames all outcomes as gains; with CRRA, as shown in Figure 3 more generally for DARA valuation function, the physician sets $L < R^P < M$. To focus purely on the effect of loss aversion and leave the reflection effect out, consider a linear valuation function v(.), and consider $\lambda > 1$ (note that the linear valuation function can be considered as a limit case with extremely small curvature of both CARA and CRRA). In this case, $R^P = M$, as the patient is now only risk averse when perceiving mixed gains and losses. As increasing ARA or RRA makes the impact of the reflection larger, intuitively, it follows that increasing ARA or RRA will makes the R^P smaller (see Proposition 2(i)).

We next look at the effect of ARA and RRA on R^C. As with CARA, $\alpha(R)$ is flat for $R \geq H$, and increases in R approaching H from the left as soon as there is at least some loss aversion (see Figure 2), it is clear R^C is not affected by ARA (see Proposition 2(ii)(a)). For CRRA, take as a starting point a case with intermediate RRA, and with loss aversion small enough for it to be the case that $M < R^C < H$. Compare this, first, to the limit case where v(.) approaches a linear function (RRA approaches zero). Then only the effect of loss aversion is at work, and the physician wants to avoid the patient perceiving any gains, as this reduces risk tolerance, so that $R^C = H$. As a decrease in RRA means an increase in R^C, this means that R^C decreases in RRA. Second, compare to the limit case where RRA becomes extremely large. Then by the fact that ARA = RRA/Y, the ARA is large just above the reference point. Again, the physician wants to avoid the patient perceiving ay gains, and sets $R^C = H$. As this is the result of an increase in ARA, it follows that R^C increases in RRA. This intuition explains the ambiguous result in Proposition 2(ii)(b).

Proposition 2.
(i) Consider the *preventive scenario*, and let there be a *unique* R^P with $L < R^P < M$. Then
  a) under CARA, $\partial R^P / \partial \lambda > 0$ and $\partial R^P / \partial ARA < 0$;
  b) under CRRA, $\partial R^P / \partial \lambda > 0$, and $\partial R^P / \partial RRA < 0$.

(ii) Consider the *curative scenario*, and let there be a *unique* R^C with $M < R^C < H$. Then
  a) under CARA, $\partial R^C / \partial \lambda = 0$, and $\partial R^C / \partial ARA = 0$;
  b) under CRRA, $\partial R^C / \partial \lambda > 0$, and $\partial R^C / \partial RRA < 0$ for small RRA, whereas $\partial R^C / \partial RRA > 0$ for large RRA.

Proof:
See Appendix.

CRRA valuation functions in the context of prospect theory, Booij et al. (2010) report only a single study where *RRA* is estimated larger than 1.

4. Discussion

Several simplifying assumptions of our model deserve attention. *First*, the risky action in our model has only two possible outcomes. As soon as there are more outcomes, the patients' probability weighting function plays a role in the optimal reference point. *Second*, we have considered the same valuation function for gains and for losses. However, literature which estimates the valuation function using the power function (CRRA) often finds the valuation function to be closer to linearity in the loss part than in the gains part (for a recent overview see Booij et al., 2010). Still, the differences in the estimated powers are small, so that the effect on our results is limited. *Third*, we have assumed that patients value healthcare decisions along a single dimension. Yet, patients may consider several dimensions, such as health outcome, monetary cost, or unpleasantness of the treatment. Separate framing may then take place for each dimension, where framing may be easier for dimensions that are more salient (Rothman et al., 2006). Using the techniques of our analysis, it is possible to extend the model along these lines. *Fourth*, prospect theory is based on observed decisions over lotteries with monetary outcomes. Yet, prospect-theoretic preferences may have a specific form for healthcare decisions, because of the large stakes involved (Attema et al., 2012), and because outcomes are obtained in the distant future (Van der Pol and Ruggeri, 2008). Further empirical research is needed here to guide plausible estimates of patients' prospect-theoretic preferences. *Fifth*, we have assumed that the physician can freely set the patient's reference point, even if it deviates considerably from the patient's expectations. In Kőszegi and Rabin's (2006) influential theory of reference-dependent preferences, however, a decision maker's reference point is fully determined by his or her expectations. For a physician inducing a safe treatment, this does not pose a problem, as the optimal reference point coincides with the outcome expected by the patient. A physician inducing a risky treatment, however, should set the reference point equal to the best possible outcome, where this outcome deviates from the patient's expectations. The question arises then whether the physician can set a reference point remote from the patient's expectations. Solving this issue requires setting up an experiment that can test the framing-based approach to reference points against the expectations-based approach. Our theory may guide such an experiment, as we make predictions for the framing-based approach.

This brings us to the relevance of the present theory for empirically testing the SID hypothesis. While the vast empirical literature on SID shows that physicians respond to financial incentives (see e.g. recently Van Dijk et al., 2013), doubt remains as to whether this proves the existence of SID: simply, we cannot hope to observe what patients with the same information as the physicians would do. The theory in this paper has the benefit of predicting *how* persuasive SID could take place. Yet, this still does not mean that empirically, one is able to catch physicians in the act of inducing demand. As an alternative, one may offer hypothetical and simplified treatment decisions to participants in an experiment, frame these in different manners, and look whether participants respond in the manner we predict, also as a function of their characteristics. This may then at least provide indication of physicians' ability to induce demand – if not evidence that they actually induce demand.

References

Attema, A.E., Brouwer, W.B.F., and L'Haridon, O. (2012) Prospect theory in the health domain: a quantitative assessment. *Unpublished manuscript*.

Barigozzi, F. and Levaggi, R. (2008) Emotions in physician agency. *Health Policy* 88, 1-14.

Booij, A.S., Praag, B.M.S. van, and Kuilen, G. van de (2010) A parametric analysis of prospect theory's functionals for the general population. *Theory and Decision* 68, 115-148.

Calcott, P., 1999. Demand inducement as cheap talk. *Health Economics* 8, 721-733.

De Jaegher, K. (2012) The value of private patient information in the physician-patient relationship: a game-theoretic account. *Computational and Mathematical Methods in Medicine,* vol. 2012, Article ID 847396.

De Jaegher, K. and Jegers, M. (2000) A model of physician behaviour with demand inducement. *Journal of Health Economics* 19, 231-258.

De Jaegher, K. and Jegers, M. (2001) The physician-patient relationship as a game of strategic information transmission. *Health Economics* 10, 651-668.

Dijk C.E. van, Berg B. van den, Verheij, R.A., Spreeuwenberg, P., Groenewegen P.P., and Bakker D.H. de (2013) Moral hazard and supplier-induced demand: empirical evidence in general practice. *Health Economics*. 22, 340-52.

Dorfman, R. and Steiner, P.O. (1954). Optimal Advertising and Optimal Quality, *American Economic Review* 44, 826-836.

Dranove, D. (1988) Demand inducement and the physician/patient relationship. *Economic Inquiry* 26, 281-298.

Evans, R.G. (1974) Supplier-induced demand: some empirical evidence and implications. In: Perlman, M. (ed.) *The Economics of Health and Medical Care*. Macmillan: London; 1974, pp.162-201.

Frank R. (2001) Behavioral economics and health economics. NBER working paper no. 10881; 2004.

Hilken, K., De Jaegher, K. and Jegers, M. (2013) Strategic framing in contracts – contracts under hidden action. TKI discussion paper no. 13-04.

Kahneman, D. and Tversky, A. (1979) Prospect theory: an analysis of decision under risk. *Econometrica* 47, 263-291.

Kőszegi, B. and Rabin, M. (2006) A model of reference-dependent preferences. *Quarterly Journal of Economics* 121, 1133-1165.

Loewenstein, G., Brennan, T. and Volpp, K.G. (2007) Asymmetric paternalism to improve health behaviors. *Journal of the American Medical Association* 298, 2415-2417.

Peacock, S. and Richardson, J. (2007) Supplier-induced demand re-examining identification and misspecification in cross-sectional analysis. *The European Journal of Health Economics*, 8, 267-277.

Pol, M. van der and Ruggeri, M. (2008) Is risk attitude outcome specific within the health domain? *Journal of Health Economics* 27, 706-717.

Rothman, A.J., Bartels, R.D., Wlaschin, J. and Salovey, P. (2006) The strategic use of gain- and loss-framed messages to promote healthy behaviour: how theory can inform practice. *Journal of Communication* 56, S202-S220.

Rothman, A.J. and Salovey, P. (1997) Shaping perceptions to motivate healthy behaviour: the role of message framing. *Psychological Bulletin* 121, 3-19.

Schwartz, A. and Goldberg, J. (2008) Prospect theory, reference points, and health decisions. *Judgment and Decision Making* 3, 174-180.

Stano, M. (1987) A clarification of theories and evidence on supplier-induced demand for physicians' services. *The Journal of Human Resources* 22, 611-620.

Thaler, R.H. (1999) *Mental accounting matters*. Journal of Behavioral Decision Making 12, 183-206.

Tversky, A. and Kahneman, D. (1981) The framing of decisions and the psychology of choice. *Science* 211, 453-458.

Tversky, A. and Kahneman, D. (1992) Advances in prospect theory: cumulative representation of uncertainty. *Journal of Risk and Uncertainty* 5, 297–323.

Wakker, P.P. (2008) Explaining the characteristics of the power (CRRA) utility family. *Health Economics* 17, 1329-144.
Wakker, P.P. (2010) *Prospect Theory for Risk and Ambiguity*. Cambridge University Press, Cambridge UK.
Xie, B., Dilts, D.M. and Shor, M. (2006) The physician-patient relationship: the impact of patient-obtained information. *Health Economics* 15, 813-833.

Appendix

Using the psychic valuation function in (2), we can see that as a function of the relation between *R* and the outcomes, the expression for *α* can take on four possible forms:

$$\alpha(R \le L) = \frac{v(H-R)-v(M-R)}{v(M-R)-v(L-R)} \tag{3}$$

$$\alpha(L < R \le M) = \frac{v(H-R)-v(M-R)}{v(M-R)+\lambda v(R-L)} \tag{4}$$

$$\alpha(M < R \le H) = \frac{v(H-R)+\lambda v(R-M)}{\lambda v(R-L)-\lambda v(R-M)} \tag{5}$$

$$\alpha(R > H) = \frac{v(R-M)-v(R-H)}{v(R-L)-v(R-M)} \tag{6}$$

Proof of Proposition 1.
Step 1. Let us first compare all the cases where *R* is equal to exactly one of the values *L*, *M*, *H*:

$$\alpha(R = L) = \frac{v(H-L)-v(M-L)}{v(M-L)} \tag{7}$$

$$\alpha(R = M) = \frac{v(H-M)}{\lambda v(M-L)} \tag{8}$$

$$\alpha(R = H) = \frac{v(H-M)}{v(H-L)-v(H-M)} \tag{9}$$

It is clear from (7)-(9) that for $\lambda = 1$, given diminishing sensitivity, $v(H-L)-v(M-L) < v(H-M)$ and $v(H-L)-v(H-M) < v(M-L)$, so that $\alpha(R=L) < \alpha(R=M) < \alpha(R=H)$. The same result is obtained as long as $\lambda < \dfrac{v(H-M)}{v(H-L)-v(M-L)} = \lambda^* < 1$ (note that the LHS of λ^* is larger than 1 as soon as there is diminishing sensitivity). If $\lambda^* > \dfrac{v(H-M)}{v(H-L)-v(M-L)}$, then $\alpha(R=M) < \alpha(R=L)$ $< \alpha(R=H)$.

We next check how expressions (3)-(6) change as a function of *R*, where we pay particular attention to the cases where *R* lies close to one of the values *L*, *M*, *H*. In Step 2, we start with the cases $R \le L$ and $R > H$.

Step 2.

$$\frac{\partial\alpha(R\le L)}{\partial R}=\frac{-\frac{[v'(H-R)-v'(M-R)][v(M-R)-v(L-R)]}{[v(M-R)-v(L-R)]^2}+}{\frac{[v(H-R)-v(M-R)][v'(M-R)-v'(L-R)]}{[v(M-R)-v(L-R)]^2}} \tag{10}$$

It follows that the sign of (10) depends on the sign of

$$\frac{|v'(H-R)-v'(M-R)|}{v(H-R)-v(M-R)}-\frac{|v'(M-R)-v'(L-R)|}{v(M-R)-v(L-R)} \tag{11}$$

Furthermore:

$$\frac{\partial\alpha(R>H)}{\partial R}=\frac{\begin{array}{c}[v'(R-M)-v'(R-H)][v(R-L)-v(R-M)]-\\ [v(R-M)-v(R-H)][v'(R-L)-v'(R-M)]\end{array}}{[v(R-L)-v(R-M)]^2} \tag{12}$$

It follows that the sign of (12) depends on the sign of

$$-\frac{|v'(R-H)-v'(R-M)|}{v(R-M)-v(R-H)}+\frac{|v'(R-M)-v'(R-L)|}{v(R-L)-v(R-M)} \tag{13}$$

In (11) and (13), each of the terms have a direct relation to the ARA. To see why, note that the ARA can be seen as the slope at a particular point of the function $f'(X^{-1}(f))$ that expresses marginal valuation as a function of the valuation itself (rather than as a function of the outcome). Instead, each of the four terms in (11) and (13) measure the slope of the secant through two distinct points on this same function, and therefore measure average ARA between these two points. It immediately follows that if it is everywhere the case that ARA increases, decreases or remains constant for higher outcomes (and therefore also for higher utility), then average ARA as reflected in these four terms will also increase, decrease or remain constant. We therefore obtain the following cases:

(i) With CARA, $\frac{\partial\alpha(R\le L)}{\partial R}=\frac{\partial\alpha(R>H)}{\partial R}=0$.

(ii) With DARA, $\frac{\partial\alpha(R\le L)}{\partial R}<0$ and $\frac{\partial\alpha(R>H)}{\partial R}<0$.

(iii) With IARA, $\frac{\partial\alpha(R\le L)}{\partial R}>0$ and $\frac{\partial\alpha(R>H)}{\partial R}>0$.

We further have:

$$\frac{\partial\alpha(L<R\le M)}{\partial R}=\frac{\begin{array}{c}[-v'(H-R)+v'(M-R)][v(M-R)+\lambda v(R-L)]-\\ [v(H-R)-v(M-R)][-v'(M-R)+\lambda v'(R-L)]\end{array}}{[v(M-R)+\lambda v(R-L)]^2} \tag{14}$$

$$\frac{\partial\alpha(M<R\leq H)}{\partial R}=\frac{\begin{array}{c}[\lambda v'(R-M)-v'(H-R)][\lambda v(R-L)-\lambda v(R-M)]-\\ [\lambda v(R-M)+v(H-R)][\lambda v'(R-L)-\lambda v'(R-M)]\end{array}}{[\lambda v(R-L)-\lambda v(R-M)]^2} \tag{15}$$

In the following steps, we use (14) and (15) to derive further results about the shape of α as a function of R.

Step 3. Evaluating (10) and (14) around $R=L$, we obtain:

$$\left.\frac{\partial\alpha(R\leq L)}{\partial R}\right|_{R=L}=\frac{\begin{array}{c}-[v'(H-L)-v'(M-L)][v(M-L)]+\\ [v(H-L)-v(M-L)][v'(M-L)-v'(0)]\end{array}}{[v(M-L)]^2} \tag{16}$$

and

$$\left.\frac{\partial\alpha(L<R\leq M)}{\partial R}\right|_{R\to L}=\frac{\begin{array}{c}-[v'(H-L)-v'(M-L)][v(M-L)]+\\ [v(H-L)-v(M-L)][v'(M-L)-\lambda v'(0)]\end{array}}{[v(M-R)+\lambda v(R-L)]^2} \tag{17}$$

It follows that

$$\left.\frac{\partial\alpha(R\leq L)}{\partial R}\right|_{R=L}>\left.\frac{\partial\alpha(L<R\leq M)}{\partial R}\right|_{R\to L}. \tag{18}$$

The following therefore applies:

(i) With CARA, by (18) and Step 2(i), $\left.\frac{\partial\alpha(L<R\leq M)}{\partial R}\right|_{R\to L}<0=\left.\frac{\partial\alpha(R\leq L)}{\partial R}\right|_{R=L}$.

(ii) With DARA, by (18) and Step 2(ii), $0>\left.\frac{\partial\alpha(R\leq L)}{\partial R}\right|_{R=L}>\left.\frac{\partial\alpha(L<R\leq M)}{\partial R}\right|_{R\to L}$.

(iii) With IARA, by (18) and Step 2(iii), $\left.\frac{\partial\alpha(R\leq L)}{\partial R}\right|_{R=L}>\left.\frac{\partial\alpha(L<R\leq M)}{\partial R}\right|_{R\to L}$, where $\left.\frac{\partial\alpha(R\leq L)}{\partial R}\right|_{R=L}>0$. By (17), $\left.\frac{\partial\alpha(L<R\leq M)}{\partial R}\right|_{R\to L}>0$ when $\lambda<\lambda_L$, and $\left.\frac{\partial\alpha(L<R\leq M)}{\partial R}\right|_{R\to L}<0$ when $\lambda>\lambda_L$, with

$$\lambda_L=\frac{v(H-L)v'(M-L)-v(M-L)v'(H-L)}{[v(H-L)-v(M-L)]v'(0)}, \tag{19}$$

where it can be checked that for IARA, $\lambda_L>1$.

Step 4. Evaluating (14) and (15) around $R=M$, we obtain:

$$\left.\frac{\partial\alpha(L<R\leq M)}{\partial R}\right|_{R=M}=\frac{[v'(0)-v'(H-M)]\lambda v(M-L)-v(H-M)[\lambda v'(M-L)-v'(0)]}{[\lambda v(M-L)]^2} \tag{20}$$

and

$$\left.\frac{\partial\alpha(M<R\leq H)}{\partial R}\right|_{R\to M}=\frac{[\lambda v'(0)-v'(H-M)]\lambda v(M-L)-v(H-M)[\lambda v'(M-L)-\lambda v'(0)]}{[\lambda v(M-L)]^2} \tag{21}$$

Given diminishing sensitivity, it follows by (21) that it is always the case that $\left.\frac{\partial\alpha^i(M<R\leq H)}{\partial R}\right|_{R\to M}>0$. (20) and (21) imply that $\left.\frac{\partial\alpha(M<R\leq H)}{\partial R}\right|_{R\to M}>\left.\frac{\partial\alpha(L<R\leq M)}{\partial R}\right|_{R=M}$. Furthermore, by (20), $\left.\frac{\partial\alpha(L<R\leq M)}{\partial R}\right|_{R=M}>0$ if $v'(0)>[v'(M-L)v(H-M)+v'(H-M)v(M-L)]/v(M-L)$. For $v'(0)<[v'(M-L)v(H-M)+v'(H-M)v(M-L)]/v(M-L)$, $\left.\frac{\partial\alpha(L<R\leq M)}{\partial R}\right|_{R=M}>0$ if $\lambda<\lambda_M$, and $\left.\frac{\partial\alpha(L<R\leq M)}{\partial R}\right|_{R=M}<0$ if $\lambda>\lambda_M$, with

$$\lambda_M=\frac{v'(0)v(H-M)}{v'(M-L)v(H-M)-[v'(0)-v'(H-M)]v(M-L)}, \tag{22}$$

where it can be checked that it is always valid that $\lambda_M>1$.

Step 5. Evaluating (12) and (15) around $R=H$, we obtain:

$$\left.\frac{\partial\alpha(M<R\leq H)}{\partial R}\right|_{R=H}=\frac{[\lambda v'(H-M)-v'(0)][\lambda v(H-L)-\lambda v(H-M)]-[\lambda v(H-M)][\lambda v'(H-L)-\lambda v'(H-M)]}{[\lambda v(H-L)-\lambda v(H-M)]^2} \tag{23}$$

$$\left.\frac{\partial\alpha(R>H)}{\partial R}\right|_{R\to H}=\frac{[v'(H-M)-v'(0)][v(H-L)-v(H-M)]-[v(H-M)][v'(H-L)-v'(H-M)]}{[v(H-L)-v(H-M)]^2} \tag{24}$$

It follows from (23) and (24) that

$$\left.\frac{\partial\alpha(R>H)}{\partial R}\right|_{R\to H}<\left.\frac{\partial\alpha(M<R\leq H)}{\partial R}\right|_{R=H}. \tag{25}$$

The following now applies:

(i) For CARA, by (25) and Step 2(i), $\left.\frac{\partial\alpha(R>H)}{\partial R}\right|_{R\to H}=0<\left.\frac{\partial\alpha(M<R\leq H)}{\partial R}\right|_{R=H}$.

(ii) For DARA, by (25) and Step 2(ii), $\left.\frac{\partial\alpha(R>H)}{\partial R}\right|_{R\to H} < \left.\frac{\partial\alpha(M<R\leq H)}{\partial R}\right|_{R=H}$ and $\frac{\partial\alpha(R>H)}{\partial R}<0$. In this case, $\left.\frac{\partial\alpha(M<R\leq H)}{\partial R}\right|_{R=H}<0$ iff $\lambda<\lambda_H$, and $\left.\frac{\partial\alpha(M<R\leq H)}{\partial R}\right|_{R=H}>0$ iff $\lambda>\lambda_H$, with

$$\lambda_H=\frac{v'(0)[v(H-L)-v(H-M)]}{\{v'(H-M)v(H-L)-v'(H-L)v(H-M)\}}, \tag{26}$$

where it can be checked that for DARA, $\lambda_H>1$.

(iii) For IARA, by (25) and Step 2(iii), $0<\left.\frac{\partial\alpha(R>H)}{\partial R}\right|_{R\to H} < \left.\frac{\partial\alpha(M<R\leq H)}{\partial R}\right|_{R=H}$.

Step 6. In this final step, we put all the results of the previous steps together.

(i) For CARA, by Step 2, $\alpha(R)$ is flat in R both for $L\leq R$ and $R\geq H$, where by diminishing sensitivity $\alpha(R)$ is higher for $H\geq R$. By Steps 4 and 5, $\alpha(R)$ increases in R just above M and just below H, so that by single-peakedness, $\alpha(R)$ increases everywhere in R for $M<R<H$. By Step 3, $\alpha(R)$ decreases in R just above L. By Step 4, if $v'(0)>[v'(M-L)v(H-M)+v'(H-M)v(M-L)]/v(M-L)$, $\alpha(R)$ decreases in R just below M. By single-peakedness it follows that $\alpha(R)$ decreases in R for $L<R<M$. Yet, by (8) in Step 1, as loss aversion approaches infinity, $\alpha(M)$ approaches zero, which is the lowest value that $\alpha(R)$ can reach. If $v'(0)<[v'(M-L)v(H-M)+v'(H-M)v(M-L)]/v(M-L)$, $\alpha(R)$ decreases in R just below M if $\lambda<\lambda_M$, and $\alpha(R)$ increases in R just below M if $\lambda>\lambda_M$. By single-peakedness, in the former case, $\alpha(R)$ has a local minimum for $R=M$; in the latter case, $\alpha(R)$ reaches a local minimum for $L<R<M$. It follows that R^C is any value of at least H. R^P lies strictly between L and M for sufficiently small loss aversion, and lies precisely at M for sufficiently high loss aversion.

(ii) For DARA, by Step 2, $\alpha(R)$ is decreasing in R both for $L\leq R$ and $R\geq H$, where by diminishing sensitivity $\alpha(R)$ is everywhere higher for $H\geq R$. For $L<R<M$, by Steps 3 and 4, the analysis is analogous to CARA. By Step 4, $\alpha(R)$ increases in R just above M. By Step 5, $\alpha(R)$ decreases in R just below H if $\lambda<\lambda_H$, and increases in R just below H if $\lambda>\lambda_H$; by single-peakedness, $\alpha(R)$ reaches a local maximum for $M<R<H$ in the former case, and has a local maximum for $R=H$ in the latter case. It follows that R^C lies strictly between M and H for sufficiently small loss aversion, and precisely at H for sufficiently high loss aversion. R^P lies strictly between L and M for sufficiently small loss aversion, and precisely at M for sufficiently high loss aversion.

(ii) For IARA, by Step 2, $\alpha(R)$ is increasing in R both for $L\leq R$ and $R\geq H$, where by diminishing sensitivity $\alpha(R)$ is everywhere higher for $R\geq H$. By Steps 4 and 5, $\alpha(R)$ increases in R just above M and just below H, so that by single-peakedness, $\alpha(R)$ increases everywhere in R for $M<R<H$. For $L<R<M$, by Step 2, just above L, $\alpha(R)$ increases in R for $\lambda<\lambda_L$, and decreases in R for $\lambda>\lambda_L$. The analysis for R just above M is identical to the one for IARA. By single-peakedness, for sufficiently low loss aversion, $\alpha(R)$ increases everywhere in R for $L<R<M$; for sufficiently high loss aversion, it decreases everywhere; for intermediate loss aversion, however, $\alpha(R)$ may either have a local minimum, or a local

maximum strictly between L and M. As the patient can never be more risk tolerant for $L < R < M$ than for $R \geq H$, it follows that $R^C = R_{\max}$. For sufficiently small loss aversion, $R^P = R_{\min}$; it is also the case that $R^P = R_{\min}$ if loss aversion is sufficiently high for there to be a local minimum at $R = M$, but if it is the case that $\alpha(R_{\min}) < \alpha(M)$. For sufficiently high loss aversion, $R^P = M$; it is also the case that $R^P = M$ if loss aversion is high enough for there to be a peak at M, and if $\alpha(R_{\min})$ is higher than $\alpha(M)$. For loss aversion high enough to make $\alpha(R)$ decreasing in R just above L, but low enough to make $\alpha(R)$ increasing in R just below M, R^P lies strictly between L and M if $\alpha(R_{\min})$ is sufficiently high, but continues to equal $R_{\min}$ otherwise.

Proof of Proposition 2.

Step 1. Consider the generic CARA valuation function $v(X) = 1 - \exp[-ARA * X]$. Then it is easily checked, by putting the RHS of (14) equal to zero, that $R^P = L - \frac{1}{ARA} \ln \frac{1+\lambda}{2\lambda}$, so that $\partial R^P / \partial \lambda > 0$, and for $\lambda > 1$, we have $\partial R^P / \partial ARA < 0$.

Step 2. The sign of $\partial R^P / \partial \lambda$ can be found by taking the total differential of $\left.\frac{\partial \alpha(L < R \leq M)}{\partial R}\right|_{R=R^P} = 0$ with respect to R^P and λ, and equating to zero, or

$$\left.\frac{\partial^2 \alpha(L < R \leq M)}{\partial R^2}\right|_{R=R^P} dR_P + \left.\frac{\partial^2 \alpha(L < R \leq M)}{\partial R \partial \lambda}\right|_{R=R^P} d\lambda = 0 \tag{27}$$

In the same way, the sign of $\partial R^P / \partial RRA$ can be found by taking the total differential of $\left.\frac{\partial \alpha(L < R \leq M)}{\partial R}\right|_{R=R^P} = 0$ with respect to R^P and RRA, and equating to zero, or

$$\left.\frac{\partial^2 \alpha(L < R \leq M)}{\partial R^2}\right|_{R=R^P} dR_P + \left.\frac{\partial^2 \alpha(L < R \leq M)}{\partial R \partial RRA}\right|_{R=R^P} dRRA = 0 \tag{28}$$

Looking at (14), one obtains $\left.\frac{\partial \alpha(L < R \leq M)}{\partial R}\right|_{R=R_P} = 0$ when the numerator of the RHS of (14) equals zero, a condition which under the generic CRRA valuation function $v(X) = X^{1-RRA}$ can be written as

$$(1-RRA)(H-R^P)^{-RRA} \lambda (R^P - L)^{-RRA} (H-L)\{X^P\} = 0 \tag{29}$$

where $X^P = \left(\frac{M-R^P}{R^P-L}\right)^{-RRA} \lambda^{-1} \frac{H-M}{H-L} + \left(\frac{M-R^P}{H-R^P}\right)^{-RRA} \frac{M-L}{H-L} - 1$

Clearly, the condition $X^P = 0$ implicitly defines R^P. For RRA approaching 0, X^P approaches $\lambda^{-1} \frac{H-M}{H-L} + \frac{M-L}{H-L} - 1$, which is smaller than zero given that $\lambda > 1$. It follows that for RRA

approaching 0, in the range of R^P from L to M, R^P should be put as large as possible to achieve a minimum α, so that R^P approaches M. For RRA approaching 1, X^P approaches $X^P = \left(\frac{R^P - L}{M - R^P}\right)(\lambda^{-1}+1)\frac{H-M}{H-L} > 0$. It follows that for RRA approaching 1, in the range of R^P from L to M, R^P should be put as small as possible to achieve a minimum α, so that R^P approaches L. This suggests that $\partial R^P / \partial RRA < 0$ for RRA from 0 to 1, as we will show. Furthermore, for R^P such that $X^P = 0$, as we increase λ, X^P becomes negative, suggesting that R^P should be increased, as will also be shown.

Given that the denominator of (14) is positive, and given that $R = R^P$ when the numerator of (14) is zero iff $X^P = 0$, we have $\operatorname{sgn}\left.\frac{\partial^2 \alpha(L < R \le M)}{\partial R^2}\right|_{R=R^P} = \operatorname{sgn}\frac{\partial X^P}{\partial R^P}$ (which is larger than zero by the second-order condition, as can indeed be checked). By the same reasoning, we have $\operatorname{sgn}\left.\frac{\partial^2 \alpha(L < R \le M)}{\partial R \partial \lambda}\right|_{R=R^P} = \operatorname{sgn}\frac{\partial X^P}{\partial \lambda}$. By (27), it follows that $\operatorname{sgn}\frac{\partial R^P}{\partial \lambda} = -\operatorname{sgn}\frac{\partial X^P}{\partial \lambda}$. Given that $\partial X^P / \partial \lambda < 0$, it follows that $\partial R^P / \partial \lambda > 0$. By the same reasoning, $\operatorname{sgn}\left.\frac{\partial^2 \alpha(L < R \le M)}{\partial R \partial RRA}\right|_{R=R^P} = \operatorname{sgn}\frac{\partial X^P}{\partial RRA}$. By (28), it follows that $\operatorname{sgn}\frac{\partial R^P}{\partial RRA} = -\operatorname{sgn}\frac{\partial X^P}{\partial RRA}$. Now

$$\frac{\partial X^P}{\partial RRA} = -\left(\frac{M-R^P}{R^P-L}\right)^{-RRA} \ln\left(\frac{M-R^P}{R^P-L}\right)\lambda^{-1}\frac{H-M}{H-L} - \left(\frac{M-R^P}{H-R^P}\right)^{-RRA} \ln\left(\frac{M-R^P}{H-R^P}\right)\frac{M-L}{H-L} \tag{30}$$

As $M - R^P < H - R^P$, the second term of the RHS of (30) is positive. Furthermore, for $R^P > (L+M)/2$, the first term is positive as well, so that $\partial X^P / \partial RRA > 0$, meaning that $\partial R^P / \partial RRA < 0$. This includes the case where *RRA* approaching 0, in which case by the above R^P approaches M, so that $R^P > (L+M)/2$, $\partial X^P / \partial RRA > 0$. Furthermore, taking the derivative of the RHS of (30) with respect to RRA, we obtain:

$$\frac{\partial^2 X^P}{\partial RRA^2} = \left(\frac{M-R^P}{R^P-L}\right)^{-RRA}\left[\ln\left(\frac{M-R^P}{R^P-L}\right)\right]^2 \lambda^{-1}\frac{H-M}{H-L} -$$

$$\left(\frac{M-R^P}{H-R^P}\right)^{-RRA}\left[\ln\left(\frac{M-R^P}{H-R^P}\right)\right]^2 \frac{M-L}{H-L} > 0$$

Given that $\partial X^P / \partial RRA > 0$ for the smallest possible *RRA*, and given that $\partial X^P / \partial RRA$ is everywhere increasing in RRA, it follows that it is everywhere the case that $\partial X^P / \partial RRA > 0$, meaning that $\partial R^P / \partial RRA < 0$. Indeed, for RRA approaching 1, as already shown above, R^P approaches L.

Step 3. By Proposition 1(ii)(c), for any CARA, *any* R^C such that $H \le R^C \le +\infty$ is optimal. It follows that R^C does not change a function of λ or *RRA*.

Step 4. The sign of $\partial R^C / \partial \lambda$ can be found by taking the total differential of $\left.\frac{\partial \alpha(M < R \le H)}{\partial R}\right|_{R=R^C} = 0$ with respect to R^P and λ, and equating to zero, or

$$\left.\frac{\partial^2\alpha(M<R\le H)}{\partial R^2}\right|_{R=R^C} dR^C + \left.\frac{\partial^2\alpha(M<R\le H)}{\partial R\partial\lambda}\right|_{R=R^C} d\lambda = 0 \tag{31}$$

In the same way, the sign of $\partial R^C / \partial RRA$ can be found by taking the total differential of $\left.\frac{\partial\alpha(M<R\le H)}{\partial R}\right|_{R=R^C} = 0$ with respect to R^P and RRA, and equating to zero, or

$$\left.\frac{\partial^2\alpha(M<R\le H)}{\partial R^2}\right|_{R=R^C} dR^C + \left.\frac{\partial^2\alpha(M<R\le H)}{\partial R\partial RRA}\right|_{R=R^C} dRRA = 0 \tag{32}$$

Looking at (15), one obtains $\left.\frac{\partial\alpha(M<R\le H)}{\partial R}\right|_{R=R^C} = 0$ when the numerator of the RHS of (15) equals zero, which under the generic CRRA valuation function $v(X) = X^{1-RRA}$ can be written as

$$(1-RRA)(H-R^C)^{-RRA}(R^C-M)^{-RRA}(H-M)\{X^C\} = 0 \tag{33}$$

where $X^C = \left(\frac{R^C-L}{H-R^C}\right)^{-RRA} \lambda \frac{M-L}{H-M} - \left(\frac{R^C-L}{R^C-M}\right)^{-RRA} \frac{H-L}{H-M} + 1$

Clearly, the condition $X^C = 0$ implicitly defines R^C. For RRA approaching 0, X^C approaches $\lambda\frac{M-L}{H-M} - \frac{H-L}{H-M} + 1$, which is larger than zero. It follows that in the range of R^C from M to H, R^C should be put as large as possible to achieve a maximum, so that R^C approaches H. For RRA approaching 1, it can be checked that X^C approaches $\left(\frac{H-R^C}{R^C-L}\right)(\lambda+1)\frac{M-L}{H-M} > 0$. It follows that R^C once more approaches H. This suggests that for small RRA we have $\partial R^C / \partial RRA < 0$, whereas for large RRA we have $\partial R^C / \partial RRA > 0$, which we now show. Furthermore, for R^C such that $X^C = 0$, as we increase λ, X^C becomes positive, suggesting that R^C should be increased, as will also be shown.

Given that the denominator of (15) is positive, and given that (33) is valid when $X^C = 0$, we have $\operatorname{sgn}\left.\frac{\partial^2\alpha(M<R\le H)}{\partial R^2}\right|_{R=R^C} = \operatorname{sgn}\frac{\partial X^C}{\partial R^C}$ (which is smaller than zero by the second-order condition, as can indeed be checked). In the same way, $\operatorname{sgn}\left.\frac{\partial^2\alpha(M<R\le H)}{\partial R\partial\lambda}\right|_{R=R^C} = \operatorname{sgn}\frac{\partial X^C}{\partial\lambda}$. By (31), it follows that $\operatorname{sgn}\frac{\partial R^C}{\partial\lambda} = \operatorname{sgn}\frac{\partial X^C}{\partial\lambda}$. Given that $\partial X^C / \partial\lambda > 0$, it follows that $\partial R^C / \partial\lambda > 0$. Furthermore, $\operatorname{sgn}\left.\frac{\partial^2\alpha(M<R\le H)}{\partial R\partial RRA}\right|_{R=R^C} = \operatorname{sgn}\frac{\partial X^C}{\partial RRA}$. By (32), it follows that $\operatorname{sgn}\frac{\partial R^C}{\partial RRA} = \operatorname{sgn}\frac{\partial X^C}{\partial RRA}$. Now

$$\frac{\partial X^C}{\partial RRA} = -\left(\frac{R^C - L}{H - R^C}\right)^{-RRA} \ln\left(\frac{R^C - L}{H - R^C}\right)\lambda \frac{M - L}{H - M} + \left(\frac{R^C - L}{R^C - M}\right)^{-RRA} \ln\left(\frac{R^C - L}{R^C - M}\right)\frac{H - L}{H - M} \tag{34}$$

As $R^C - L > R^C - M$, the second term of the RHS of (34) is positive. Furthermore, for $R^C > (L + H)/2$, the first term of the RHS of (34) is negative, so that $\partial R^C / \partial RRA$ and $\partial X^C / \partial RRA$ may be ambiguous. Indeed, for RRA approaching 0, as already shown, R^C approaches *H*, in which case it can be checked that $\partial X^C / \partial RRA$ is negative. Also, for RRA approaching 1, R^C again approaches *H*, in which case it can be checked that $\partial X^C / \partial RRA$ is positive.

QED