Test design under voluntary participation
Supplementary material

Frank Rosar
University of Bonn

April 27, 2016

1. Parametrization of HARA utility

In the paper, I impose the assumptions $u' > 0$, $u'' < 0$ and $u''' \geq 0$ on the agent's reduced-form utility function. I argue here that these assumptions are satisfied for a broad class of HARA utility functions.

The parameterized class of utility functions

$$u(\mu) = \begin{cases} -\frac{1}{1-c_1}(c_1\mu + c_2)^{-\frac{1}{c_1}+1} & \text{if } c_1 \in \mathbb{R}\backslash\{0,1\} \text{ and } c_2 \in [\max\{-c_1, 0\}, \infty) \\ -c_2 \exp(-\frac{1}{c_2}\mu) & \text{if } c_1 = 0 \text{ and } c_2 \in (0, \infty) \\ \ln(\mu + c_2) & \text{if } c_1 = 1 \text{ and } c_2 \in (0, \infty) \end{cases}$$

exhibits hyperbolic absolute risk aversion (HARA). The Arrow-Pratt measure of absolute risk-aversion is given by $-u''(\mu)/u'(\mu) = 1/(c_1\mu + c_2)$. As special cases, this class includes quadratic utility ($c_1 = -1$), cubic utility ($c_1 = -1/2$), exponential/CARA utility ($c_1 = 0$), logarithmic utility ($c_1 = 1$), and CRRA utility ($c_2 = 0$). The parameter space is illustrated by the grey areas in Figure S-1; the special cases are indicated by the colored line segments. $u' > 0$ and $u'' < 0$ is true for the entire parameter space; $u''' \geq 0$ is true for $c_1 \geq -1$. That is, the assumptions in my reduced model are satisfied for any HARA utility with parameters in the dark grey area. This includes all mentioned special cases.

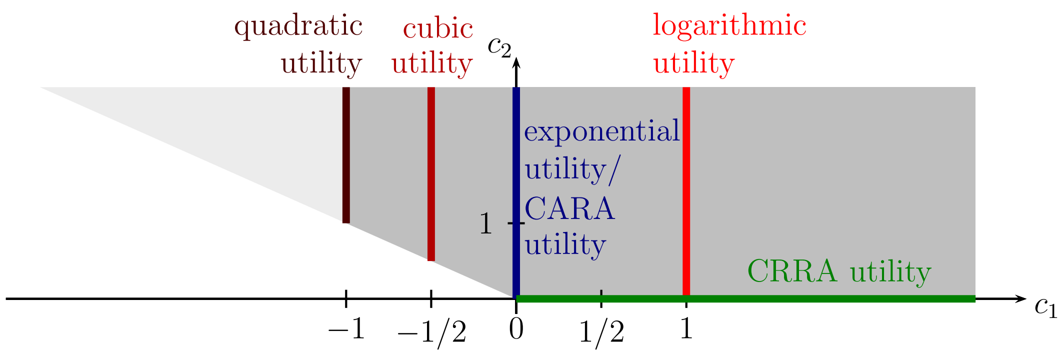

Figure S-1: Parameter space and special cases of HARA utility

2. Imperfectly informative, non-binary quality types

The analysis in Subsection 5.3 of the paper allows me also to draw conclusions about the case where the agent's quality is ultimately still binary but where it is only possible to generate non-binary stochastic information about this quality at the time of testing. As an example, think of a bank that is ultimately either viable or not. At the time of testing, it may only be possible to generate information about whether the bank is definitely viable, definitely non-viable or that it is only viable under certain circumstances.

In the simplest case, the test is with a probability $\psi \in (0,1)$ that is independent of the other random variables not capable of generating additional information.[1] There exist then three quality types: the agent is bad ($\widehat{\omega} = b$), the test is not capable of generating information ($\widehat{\omega} = n$), and the agent is good ($\widehat{\omega} = g$). The quality type $\widehat{\omega} = n$ reveals that the agent's quality ω is good with probability θ_Y. In this environment, a test maps each quality type $\widehat{\omega} \in \{b, n, g\}$ into a distribution over test results. Quality perceptions are computed according to the formula

$$\mu(p^b_\sigma, p^n_\sigma, p^g_\sigma; \theta_Y) \equiv \frac{(1-\psi)\theta_Y p^g_\sigma \cdot 1 + \psi p^n_\sigma \cdot \theta_Y + (1-\psi)(1-\theta_Y)p^b_\sigma \cdot 0}{(1-\psi)\theta_Y p^g_\sigma + \psi p^n_\sigma + (1-\psi)(1-\theta_Y)p^b_\sigma}. \tag{S-1}$$

As the additional, imperfectly informative quality type has to be mapped to at least one test result, it constrains how high μ_2 and how low μ_1 can be; that is, it affects which quality perception pairs can be induced by a binary test (i.e., Lemma 2 in the paper). However, the principal's and the agent's preferences over quality perception pairs as described in Lemma 3 and 4 of the paper stay unaffected. Hence, the modification in the model introduces only a constraint that limits the set of feasible quality perception pairs. After using (S-1) to compute the set of supportable quality perception pairs, I can apply a reasoning that is analogous to that in Remark 4 of the paper.

If ψ is sufficiently low, the constraint is not binding such that a quality perception pair $(\mu_1, 1)$ is optimal. The optimal binary test is then not subject to false positives: $\widehat{\omega} = g$ is revealed with a certain probability and all other cases are pooled. Intuitively, the limits to information generation stay effectively without consequence as the test that is optimal for the unconstrained problem (i.e., for $\psi = 0$) is inaccurate anyway and the imperfectly informative quality type can be used to generate the required amount of inaccuracy.

3. Relation to problems with non-binary quality

Consider a problem where the agent's true quality is $q \in \{q_1, q_2, q_3\}$ with $q_1 < q_2 < q_3$. In an important class of economic environments, the principal's and the agent's payoff depend only on the expected value of the true quality conditional on the information released through testing/not testing. Let $m(q)$ denote the information that is released. The reduced-form payoff of

[1]For a formal analysis of this extension, see the previous working paper version of this article, Rosar (2014).

the principal (agent) can then be written as $\check{v}(\mathbb{E}[q|m(q)])$ ($\check{u}(\mathbb{E}[q|m(q)])$). For example, payoffs with such a structure arise when the principal takes a decision to minimize the quadratic distance between her decision and the actual quality, and the agent cares only about the principal's decision.

When I define $\theta(q) \equiv (q-q_1)/(q_3-q_1)$ and $\mu_{m(q)} \equiv \mathbb{E}_\theta[\theta(q)|m(q)]$, I can rewrite the reduced-form payoffs as $v(\mu_{m(q)}) \equiv \check{v}(q_1 + (q_3 - q_1)\mu_{m(q)})$ and $u(\mu_{m(q)}) \equiv \check{u}(q_1 + (q_3 - q_1)\mu_{m(q)})$. This shows that, technically, the problem with non-binary quality $q \in \{q_1, q_2, q_3\}$ is equivalent to a problem with binary quality, say high (q_3) or low (q_1), where it is only possible to reveal stochastic information about the agent's quality: $\theta(q)$ describes then the probability that the agent's reinterpreted quality is high when the true quality is q. A test can reveal information about whether $\theta(q)$ assumes the value $\theta(q_1) = 0$, $\theta(q_2) \in (0, 1)$ or $\theta(q_3) = 1$. $\mu_{m(q)}$ describes the principal's perception of the reinterpreted quality; that is, it describes the expected probability that the agent's reinterpreted quality is high conditional on the observation $m(q)$.

By this reasoning, the problem with non-binary quality is related to the problem with binary quality and limits to information generation that I described in Section 2 of this supplementary material. What makes the analysis in my paper specific is not the assumption of binary quality but the considered nature of private information: I assumed that private information is one-dimensional and that private signals and test results are independent conditional on quality. In reasonable problems with non-binary quality, this assumption is violated. Therefore, allowing for more general forms of private information would be interesting, but it would give rise to additional issues that make the analysis less tractable (e.g., the existence of tests that induce non-monotonic participation behavior). The reinterpretation of variables described above allows me to use the solution approach proposed in this article also for the problem with non-binary quality. The additional issues that arise for more general assumptions on private information may, however, require the imposition of additional structure.

4. Discussion of the optimal participation behavior in the uniform-quadratic case

In the paper, I argue that the participation threshold, which the principal does optimally induce, solves the following problem PART:

$$\begin{aligned}\max_{s\in\mathcal{S}} \quad & (1-F(s))\left[p^{\theta_Y(s)}(1, 1-\rho(s))v(\mu(1, 1-\rho(s); \theta_Y(s))) + p^{\theta_Y(s)}(0, \rho(s))v(1)\right] \\ & +F(s)v(\theta_N(s))\end{aligned}$$

In this supplementary material, I discuss the choice of the optimal participation threshold in the uniform-quadratic case:

Assumption S-1 *$u(\mu) = -(1-\mu)^2$, $v(\mu) = (\mu - \theta_0)^2$, and $\theta \sim U[\underline{\theta}(\delta), \overline{\theta}(\delta)]$ with $\underline{\theta}(\delta) \equiv (1-\delta)/2$, $\overline{\theta}(\delta) \equiv (1+\delta)/2$ and $\delta \in (0, 1)$.*

Figure 2 in the paper displays which participation thresholds are inducible under this assumption.

In Section 4 of the paper, I have argued that an accurate test induces less participation than any sufficiently inaccurate test. Although the switch from a very inaccurate test to an accurate test causes clearly less participation, participation may in general locally increase as the accuracy of the test increases. Responsible for this is that an increase in the participation threshold improves the pool of participants as well as the pool of non-participants. Depending on the relative strength of these improvements, participation may get more or less attractive in response to an increase in the supposed participation threshold. This implies that it is a priori not clear whether the equilibrium participation threshold increases or decreases in response to a small change in the test accuracy. The subsequent proposition establishes that there is under Assumption S-1 also locally for any intermediate participation threshold a trade-off between accuracy and participation.

Proposition S-1 (Trade-off between test accuracy and participation) *Suppose that Assumption S-1 holds and consider tests $T^{NFP}(\rho(s))$ where $\rho(s)$ is as defined in Proposition 4 of the paper. Then, $\rho(s)$ is increasing; that is, more participation is induced by a less accurate test.*

Proof. It follows from Proposition 2 in the paper that $\mathcal{S}$ is under Assumption S-1 a convex set (see also the discussion of the example in Section 5.1 of the paper). By Proposition 4 in the paper, the optimal test accuracy $\rho(s)$ solves $U^s(T^{\text{NFP}}(\rho(s)), \mu_Y(T^{\text{NFP}}(\rho(s)); \theta_Y(s))) = u(\theta_N(s))$ if $s \in \mathcal{S}\backslash\mathcal{S}$. Moreover, it follows from Proposition 2 (b) in the paper that the same equation holds if $s \in \mathcal{S}_a$. Under Assumption S-1, the solution is also in this case unique. Thus, for any $s \in \mathcal{S}$, I can write the equation that uniquely determines the test accuracy ρ as

$$g(\rho, s) \equiv (1 - \rho s)u(\mu(1, 1 - \rho; \theta_Y(s))) + \rho s u(1) - u(\theta_N(s)) = 0$$

By applying the Implicit Function Theorem, I obtain that $\rho(s)$ is a continuously differentiable function with $\rho'(s) = -(\partial g(\rho, s)/\partial s)/(\partial g(\rho, s)/\partial \rho)$. Since $\mu(1, 1 - \rho; \theta_Y(s))$ is decreasing in ρ, it follows from Lemma 3 (d) in the paper that $\partial g(\rho, s)/\partial \rho < 0$. Thus, to conclude the proof, it remains to show that $\partial g(\rho, s)/\partial s > 0$. I have,

$$\begin{aligned}
\frac{\partial g(\rho, s)}{\partial s} &= \rho(u(1) - u(\mu(1, 1 - \rho; \theta_Y(s))) \\
&\quad + (1 - \rho s)\frac{1}{2}\frac{1 - \rho}{(1 - \theta_Y(s)\rho)^2}u'(\mu(1 - \rho, \rho; \theta_Y(s))) - \frac{1}{2}u'(\theta_N(s)) \\
&> \rho(u(1) - u(\mu(1, 1 - \rho; \theta_Y(s))) \\
&\quad + \frac{1}{2}\left(\frac{1 - \rho s}{1 - \theta_Y(s)\rho}\frac{1 - \rho}{1 - \theta_Y(s)\rho} - 1\right)u'(\mu(1 - \rho, \rho; \theta_Y(s))) \\
&> \rho(u(1) - u(\mu(1, 1 - \rho; \theta_Y(s))) \\
&\quad + \frac{1}{2}\left(\frac{1 - \rho}{1 - \theta_Y(s)\rho} - 1\right)u'(\mu(1 - \rho, \rho; \theta_Y(s))) \\
&= \rho(u(1) - u(\mu(1, 1 - \rho; \theta_Y(s)))
\end{aligned}$$

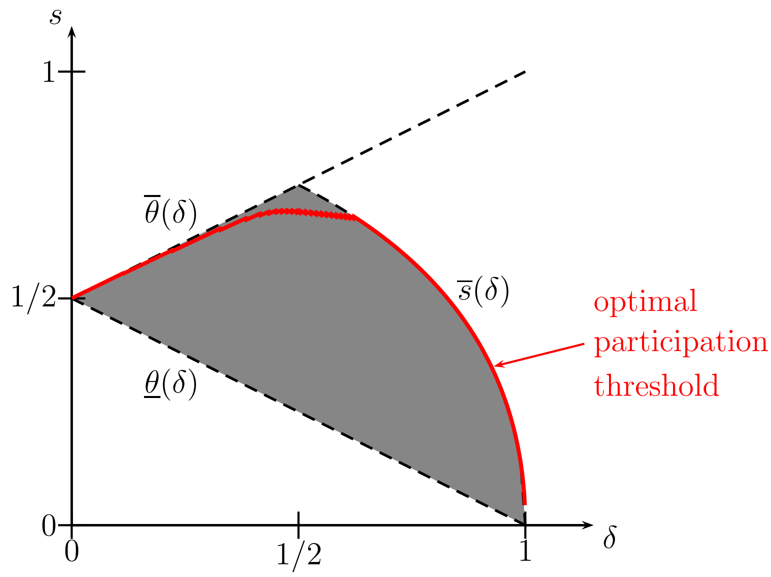

Figure S-2: Optimal participation threshold under Assumption S-1

$$
\begin{aligned}
& -\frac{1}{2}\rho(1-\mu(1-\rho,\rho;\theta_Y(s)))u'(\mu(1-\rho,\rho;\theta_Y(s))) \\
= \; & 0.
\end{aligned}
$$

The transformations arise as follows: The first equality follows from using that $\theta'_Y(s) = \theta'_N(s) = 1/2$ under Assumption S-1, and that $\mathrm{d}\mu(1, 1-\rho; \theta_Y)/\mathrm{d}\theta_Y = (1-\rho)/(1-\theta_Y \rho)^2$. To see why the first inequality holds, note first that necessary for $g(\rho, s) = 0$ to hold is $\mu(1, 1-\rho; \theta_Y(s)) < \theta_N(s)$. The inequality follows because concavity of u implies then that $u'(\mu(1-\rho, \rho; \theta_Y(s))) > u'(\theta_N(s))$. The second inequality follows from using that $s < \theta_Y(s)$ implies that $(1-\rho s)/(1-\theta_Y(s)\rho) > 1$. The second equality follows from using that I can rewrite the expression in parentheses as $-\rho(1-\mu(1, 1-\rho; \theta_Y(s)))$. The third equality follows from using that the structure of u in Assumption S-1 implies that $u(1) - u(\mu_1) = 1/2 \cdot (1-\mu_1)u'(\mu_1)$. q.e.d.

As motivated in Section 5.4 of the paper, I can interpret the optimal test as a pass-fail test. Decreasing the accuracy of a NFP test $T^{\text{NFP}}(\rho)$ makes "passing" less likely and increases the quality perception associated with "failing". This allows for the interpretation that more participation is optimally induced by "reducing the stigma of failure" and not by "inflating grades" (even though inducing more participation by inflating grades would also be possible).

I come now to the question of how the informativeness of the agent's private information δ affects the optimal participation threshold. The grey area in Figure S-2 indicates which participation thresholds are inducible; the red solid line illustrates the optimal participation threshold.[2] If the informativeness of the agent's private signal is low, any participation threshold $s \in [\underline{\theta}(\delta), \overline{\theta}(\delta)]$ is inducible. As signals are spread out in a small interval around 1/2, the

[2] I used Maple to derive the optimal participation threshold numerically.

worst possible quality perception associated to non-participation is quite high and the agent's signaling motive is weak. Fostering participation is costly for the principal in terms of test accuracy. The principal prefers little participation in a relatively accurate test to a higher participation in a much less accurate test. Thus, the optimal participation threshold lies close to $\overline{\theta}(\delta)$. If the informativeness of the agent's private signal is high, the agent's signaling motive is strong as the quality perception associated to non-participation is low. An unravelling effect kicks in causing relatively high participation in an accurate test. As inducing participation for private signals close to $\underline{\theta}(\delta)$ requires a very inaccurate test, fostering even more participation is, again, very costly in terms of test accuracy. An accurate test and therewith the highest inducible participation threshold is optimal. As inducing full participation is never optimal, there is always a role for indirect learning.

Remark S-1 (Voluntary vs. mandatory participation) A key assumption in my model was that the agent's participation in the test is voluntary. While this assumption is reasonable for some applications, the principal may be able to mandate participation in others. Suppose that the principal can choose between voluntary and mandatory participation but that she faces a cost $\kappa > 0$ of conducting the test; that is, consider $V(T, x, (\mu_N, \mu_Y)) = \mathbb{E}_\theta[x(\theta)(\sum_\sigma p^\theta(p^b_\sigma, p^g_\sigma)v(\mu_\sigma) - \kappa) + (1-x(\theta))v(\mu_N)]$. Under mandatory participation an accurate test is clearly optimal. It allows for perfect learning of the agent's quality. Nevertheless, voluntary participation may be optimal as it saves on testing costs. In particular, if κ is sufficiently large, voluntary participation is optimal. A nice feature of my two-step solution approach is that the analysis in Sections 5.3 and 5.4 of the paper remains valid for any κ. Hence, for any κ and for any participation threshold that shall be induced, a NFP test is optimal.

References

Rosar, F. (2014). Test design under voluntary participation and conflicting preferences. Mimeo, July 2014.