Technical Appendix

A. 1 Distribution of the maximum belief

The dynamic model outlined in the paper involves the distribution, *F*, of the contestant's assessment of her chance of answering the question successfully. Without lifelines, $F \equiv F_n$ is the distribution of $\max_{\mathbf{x}\in\Delta_n}(\mathbf{x})$ if $\mathbf{x}$ has the probability density function, $\psi_n(\mathbf{x})$. Indeed $\max_{\mathbf{x}\in\Delta_n}(\mathbf{x})$ measures the individual assessment of her likelihood of answering the question correctly when faced with *n* alternatives. In this section, we describe formulae for the distribution of the highest order statistic $F_n(z) \equiv \Pr\left[\bigcap_{i=1}^{n}\{X_i < z\}\right]$, given that $\mathbf{X}$ is distributed with density function ψ_n. In particular we can show that

$$F_2(z) = \begin{cases} 0 & \text{if } z \le 1/2 \\ 2\Phi[z]-1 & \text{if } 1/2 \le z \le 1, \\ 1 & \text{otherwise.} \end{cases}$$

and, as a consequence, the density function $f_2(z)$ has support $\left[\frac{1}{2},1\right]$ where is satisfies $f_2(z) = 2\phi[z]$. The distribution function at higher orders can be obtained from F_2 recursively. Whenever $z \in (0,1)$, we have $F_3(z) = 2\int_{1-z}^{1} F_2\left(\frac{z}{y}\right) y\phi(y)\,dy$, and $F_4(z) = \frac{1}{\mu_2}\int_{1-z}^{1} F_3\left(\frac{z}{y}\right) y^2\phi(y)\,dy$, and the relevant density functions, say f_3 and f_4, can be shown to exist and to be continuous everywhere inside $(0,1)$. For example, in the uniform case where $\phi(x)=1$ for $x \in [0,1]$, and 0 elsewhere, we find that

$$F_4(z) = \begin{cases} 0 & \text{if } 0 \le z \le 1/4 \\ (4z-1)^3 & \text{if } 1/4 \le z \le 1/3 \\ -44z^3+60z^2-24z+3 & \text{if } 1/3 \le z \le 1/2 \\ 1-4(1-z)^3 & \text{if } 1/2 \le z \le 1 \end{cases}$$

In this latter case it is easy to verify that the density function is continuous and that the derivatives match at the boundaries of each segment. The distribution functions F_n do depend on the density ϕ in an important fashion. We interpret ϕ as a description of the individual's knowledge. When ϕ is diffuse over $[0,1]$ (e.g. uniform) all points on the simplex Δ_n are equally likely and in some instances the individual will have the belief that she can answer the question correctly while in some cases the beliefs will be relatively uninformative, while if ϕ is concentrated around, or in the limit at, ½ the individual is always indecisive. Finally, when ϕ's modes are located around 0 and 1, the individual is always relatively informed about the correct answer.

A. 2 Lifelines

Extending the model above to allow for the lifelines makes the analysis more difficult but also enables us to exploit more aspects of the data. We show how, in the first sub-section below, the model can be modified when only one lifeline is allowed for. In a second sub-section we show how the model can be modified for all three lifelines. We then present the precise assumptions that allow the modelling of each lifeline in particular.

A. 2.1 The complete game

There are three lifelines available which can be played at most once. As above each lifeline generates a new belief, **q**, which is used in the decision process. Given the initial belief **p**, the new belief is drawn from a separate distribution for each lifeline, say $H_1(\mathbf{q}\,|\,\mathbf{p})$ for 50:50, $H_2(\mathbf{q}\,|\,\mathbf{p})$ for ATA and $H_3(\mathbf{q}\,|\,\mathbf{p})$ for PAF. We write $\boldsymbol{\gamma}=(\gamma_1,\gamma_2,\gamma_3)$ for the "lifeline state" vector where $\gamma_i=0$ if the i^{th} lifeline has been played and 1 otherwise. We use $W_n(\mathbf{p};\boldsymbol{\gamma})$ to denote the optimal expected value of the game at stage n, when the probability vector of the current question is **p** and the lifeline state is $(\gamma_1,\gamma_2,\gamma_3)$. As above, $V_n(\mathbf{p})$ is used as a shorthand for $W_n(\mathbf{p};(0,0,0))$ and $V_n(\mathbf{p})$ satisfies the recursive dynamic programming equations set out in section 5.1 above. Below we write the dynamic programming equations using the notation $f_n((\gamma_1,\gamma_2,\gamma_3))=\mathrm{E}\left[W_n(\mathbf{P}_n;(\gamma_1,\gamma_2,\gamma_3))\right]$, where the expectation is with respect to $\mathbf{P}_n$, the distribution of the belief vector **p** at stage *n*. When there are one or more lifeline left, i.e. $\gamma_1+\gamma_2+\gamma_3\geq 1$, the contestant has three options: (i) quit, (ii) answer the question, (iii) use one of the remaining lifelines. The recursive equation is $W_n(\mathbf{p};\boldsymbol{\gamma})=\max\left\{a_{n-1},\,p\left(f_{n-1}(\boldsymbol{\gamma})-b_n\right)+b_n,\,k_n(\mathbf{p};\boldsymbol{\gamma})\right\}$ where $k_n(\mathbf{p};\boldsymbol{\gamma})$ denotes the maximum expected value from using a lifeline when the belief is p and the lifeline state vector is $\boldsymbol{\gamma}$.

Here, $k_n(\mathbf{p};\boldsymbol{\gamma}) = \max_{i \in I(\boldsymbol{\gamma})} \left\{ \mathrm{E}\left[W_n\left(\mathbf{Q}_i; \boldsymbol{\gamma} - \mathbf{e}_i\right) | \mathbf{p} \right] \right\}$ where $I(\boldsymbol{\gamma}) = \left\{ j : \mathbf{e}_j \le \boldsymbol{\gamma} \right\}$ using $\mathbf{e}_i$ to denote the i^{th} unit vector in R^3 and $\mathbf{Q}_i$ is distributed according to $H_i(\mathbf{q} | \mathbf{p})$. This formulation does not preclude an individual from using more than one lifeline on the same question, a behaviour we observe in some contestants.

A. 2.2 "50:50"

This is the simplest lifeline to model. It provides the contestant with "perfect information" since two incorrect answers are removed. Ex-ante (i.e. before the lifeline is played) the contestant believes that the correct answer is i (=1,…,4) with probability p_i. The 50:50 lifeline removes two of the incorrect answers, retaining $j \neq i$, say, with equal probability (1/3). By Bayes Theorem, the probability that answers *i,j* survive this elimination process is $p_i/3$. The answers *i* and *j* can also be retained if *j* is correct and *i* survives elimination. This occurs with probability $p_j/3$. Applying Bayes Theorem gives the updated belief vector $\mathbf{q}^{\{i,j\}}$, where $\mathbf{q}_k^{\{i,j\}}$ equals $\dfrac{p_i}{p_i + p_j}$ if $k = i$, $\dfrac{p_j}{p_i + p_j}$ if $k = j$, and 0 otherwise. Hence $H_1(\mathbf{q};\mathbf{p})$ is a discrete distribution with support $\left\{\mathbf{q}^{\{i,j\}}\right\}_{\{i,j\}\in\{1,2,3,4\}}$ such that $H_1\left(\mathbf{q}^{\{i,j\}};\mathbf{p}\right) = \left(p_i + p_j\right)/3$, and 0 elsewhere.

A. 2.3 "Ask the Audience"

Modelling ATA requires more than simply applying Bayes' rule to the current belief draw. In particular we must allow the contestant to learn from the information provided by the lifeline, i.e. here the proportions of the audience's votes in favour of each alternative answer. The difficulty here is to understand why and how should a "perfectly informed" rational individual revise his/her prior on the basis of someone else's opinion? The route we follow here was proposed by French (1980) in the context of belief updating after the opinion of an expert is made available. French suggests that the updated belief that some event A is realised after some information inf has been revealed should be obtained from the initial belief, $\Pr[A]$, the marginal probability that a given realisation of the information is revealed, $\Pr[\inf]$, and the individual's belief about the likelihood that the information will arise if A subsequently occurs, $\Pr[\inf | A]$ according to the following rule, related to Bayes theorem: $\Pr[A|\inf] = \Pr[\inf \mid A]\Pr[A]/\Pr[\inf]$. In this expression $\Pr[\inf \mid A]$ is understood as another component of the individual's belief - her assessment of the likelihood of the signal given that the relevant event subsequently occurs. Introducing $\bar{A}$, A's alternative event, this is rewritten as

$$\Pr[A|\inf] = \frac{\Pr[\inf \mid A]\Pr[A]}{\Pr[\inf \mid A]\Pr[A] + \Pr[\inf \mid \bar{A}]\Pr[\bar{A}]}.$$

In our context we treat asking the audience as an appeal to an expert, and assume that the events of interest are the four events "answer k is correct", k=1,2,3,4. We assume that contestants "learn" some information about the quality of the expert in particular the distribution of the quantities $\Pr\left[\mathbf{q} = (q_1, q_2, q_3, q_4) \mid \text{answer } k \text{ is correct}\right] \equiv \theta_k$, where q_k is the proportion of votes allocated to the k^{th} alternative. Following French's proposal, the k^{th} component of the updated belief π given the information q is $\pi_k = \theta_k p_k \Big/ \sum_{j=1}^{4} \theta_j p_j$. Let us assume for now that each contestant knows the joint distribution of the vector $\theta = (\theta_1, \theta_2, \theta_3, \theta_4)$. In fact the above expression implies that, without loss of generality, we can normalise the θ_k to sum to one. Denote $I(\theta)$ the density function of θ given some initial belief p. Given p, the density of the updated belief $H_2(\pi;\mathbf{p})$ can be calculated as $H_2(\pi;\mathbf{p}) = I(\theta(\pi;\mathbf{p}))\left(\prod_{k=1}^{4} p_k\right)\left(\sum_{k=1}^{4} \pi_k p_k^{-1}\right)^4$, with $\theta_i(\pi;\mathbf{p}) = \pi_i p_i^{-1} \Big/ \sum_{k=1}^{4} \pi_k p_k^{-1}$. The term $\left(\prod_{k=1}^{4} p_k\right)\left(\sum_{k=1}^{4} \frac{\pi_k}{p_k}\right)^4$ arises because of the change of variable from θ to π.

The quantities $\Pr\left[\mathbf{q} = (q_1, q_2, q_3, q_4) \mid \text{answer } k \text{ is correct}\right] \equiv \theta_k$ represent the added information obtained from using the lifeline and are estimable from the data provided we assume a form of conditional independence. In particular we require that the contestant's choice to ask the audience does not influence the audience's answer. Furthermore, we assume that there is no information contained in the position of the correct answer, hence we expect the following symmetry restrictions to hold :

$$\Pr\left[\mathbf{q}=(q_1,q_2,q_3,q_4)\,|\,\text{answer 1 is correct}\right]=\Pr\left[\mathbf{q}=\left(q_{\sigma(1)},q_1,q_{\sigma(2)},q_{\sigma(3)}\right)|\,\text{answer 2 is correct}\right]$$
$$=\Pr\left[\mathbf{q}=\left(q_{\sigma'(1)},q_{\sigma'(2)},q_1,q_{\sigma'(3)}\right)|\,\text{answer 3 is correct}\right]$$
$$=\Pr\left[\mathbf{q}=\left(q_{\sigma''(1)},q_{\sigma''(2)},q_{\sigma''(3)}q_1,\right)|\,\text{answer 4 is correct}\right],$$

where $(\sigma(1),\sigma(2),\sigma(3))$, $(\sigma'(1),\sigma'(2),\sigma'(3))$ and $(\sigma''(1),\sigma''(2),\sigma''(3))$ are some permutations of $(2,3,4)$. The symmetry restrictions, the conditional independence assumption, and the uniform random allocation of the correct answer among four alternative answers allow us to estimate the likelihood of the information given the position of the correct answer, and therefore provide empirical estimates for $\Pr\left[\mathbf{q}=(q_1,q_2,q_3,q_4)\,|\,\text{answer } k \text{ is correct}\right]$. In practice we assume that, given answer k is correct, information **q** has a Dirichlet density $D\left(\mathbf{q};\gamma_k(\lambda,\nu)\right)$, k=1…4, defined over Δ_4 such that $D\left(\mathbf{q};\gamma_k(\lambda,\nu)\right)=\frac{\Gamma(3\nu+\lambda)}{\Gamma(\lambda)\Gamma(\nu)^3}\left(\prod_{i=1}^{4}q_i^{\nu-1}\right)q_k^{\lambda-\nu}$, where the symmetry assumption is imposed through the parameter vector $\gamma_k(\lambda,\nu)=\nu+\mathbf{e}_k(\lambda-\nu)$ with $\mathbf{e}_k$ is a vector of zeros with a 1 in position k. This vector of parameters for the Dirichlet density depends on two free parameters only, λ and ν . These two parameters can be estimated (independently from the other parameters of the model) by maximum likelihood from the observation of the information obtained from the audience (i.e. the histograms) whenever the lifeline is used, and the observation of which answer is the correct answer (even when the contestant chooses an incorrect answer, the compére always reveals the correct one). For completeness note that θ_k can be defined in terms of the elements of q as $\theta_k=q_k^{\lambda-\nu}\Big/\sum_{j=1}^{4}q_j^{\lambda-\nu}$. The information density which the contestant expects is therefore the mixture $D(\mathbf{q};\mathbf{p},\lambda,\nu)$ of the previous densities $D\left(\mathbf{q};\gamma_k(\lambda,\nu)\right)$, k=1…4, conditional on a given answer being correct, we have: $D(\mathbf{q};\mathbf{p},\lambda,\nu)=\sum_{i=1}^{4}p_iD\left(\mathbf{q};\gamma_i(\lambda,\nu)\right)=\frac{\Gamma(3\nu+\lambda)}{\Gamma(\lambda)\Gamma(\nu)^3}\left(\prod_{i=1}^{4}q_i^{\nu-1}\right)\left(\sum_{i=1}^{4}p_iq_i^{\lambda-\nu}\right)$, where the mixing weights are the initial beliefs $p_i, i=1...4$.

A. 2.4 "Phone a Friend"

To use this lifeline the contestant determines, ahead of the game, six potential experts ("friends") and when she plays the lifeline she chooses one from this list of six. We imagine that the contestant engages in some diversification when drawing up the list (i.e. the range and quality of "expert knowledge" of the friends on the list is in some way optimised), and at the time of playing the lifeline the contestant chooses the expert to call optimally. There is however little information available to us about this process. As a consequence our model for this particular lifeline is somewhat crude. We assume that the entire process can be modelled as an appeal to an expert who knows the answer with some probability κ, and is ignorant with the probability $1-\kappa$. We assume that the expert informs the contestant of his confidence (contestants invariably ask the friend how confident they feel – although the answer is usually not quantitative). Hence either the contestant knows the answer and her opinion "swamps" the contestant's belief, or the expert is ignorant and conveys no information and the contestant's belief is left unchanged. The density of the updated belief is therefore $H_3(\pi;\mathbf{p}) = \kappa \mathbf{1}_{[\pi=(1,0,0,0)]} + (1-\kappa)\mathbf{1}_{[\pi=\mathbf{p}]}$.

A. 3 Proposition **(factorisation of $\chi_4(x_1,x_2,x_3,x_4)$):**

The joint density: $\chi_4(x_1,x_2,x_3,x_4) = \dfrac{2}{\mu_2}\phi(x_1)\phi\left(\dfrac{x_2}{1-x_1}\right)\phi\left(\dfrac{x_3}{1-x_1-x_2}\right)$, such that $\sum_{i=1}^{4} x_i = 1$, $x_i \geq 0$ for all i, can be factorised as follows: $\chi_4(x_1,x_2,x_3,x_4) = f_{U_1}(x_1)\, f_{U_2|U_1}(x_2;x_1)\, f_{U_3|U_1,U_2}(x_3;x_1,x_2)$,

$f_{U_1}(u)$, $f_{U_2|U_1}(v;u)$, $f_{U_3|U_1,U_2}(w;u,v)$, are (conditional) densities such that

$f_{U_1}(u) = \dfrac{(1-u)^2\phi(u)}{\mu_2}\mathbf{1}_{[0\leq u\leq 1]}$, with $\mu_2 = \int_0^1 (1-x)^2 \phi(x)\,dx$,

$f_{U_2|U_1}(v;u) = 2\dfrac{(1-u-v)}{(1-u)^2}\phi\left(\dfrac{v}{1-u}\right)\mathbf{1}_{[0\leq v\leq 1-u]}$ and

$f_{U_3|U_1,U_2}(w;u,v) = \dfrac{1}{1-u-v}\phi\left(\dfrac{w}{1-u-v}\right)\mathbf{1}_{[0\leq w\leq 1-u-v]}$.

Proof: It is easy to verify, by simple integration for $f_{U_1}(u)$, $f_{U_2|U_1}(v;u)$, and by construction for $f_{U_3|U_1,U_2}(w;u,v)$, all three are well defined densities over the relevant ranges. Moreover their product is equal to $\chi_4(.)$. This implies that if U_1, U_2 and U_3 are three random variables each distributed with densities $f_{U_1}(u)$, $f_{U_2|U_1}(v;u)$, and $f_{U_3|U_1,U_2}(w;u,v)$, then the random vector $P = \begin{pmatrix} U_1 & \bar{U}_1 U_2 & \bar{U}_1 \bar{U}_2 U_3 & \bar{U}_1 \bar{U}_2 \bar{U}_3 \end{pmatrix}$, with $\bar{U}_i = 1 - U_i$ for all i=1..3, is distributed with joint density: $\chi_4(x_1, x_2, x_3, x_4) = \frac{2}{\mu_2} \phi(x_1) \phi\left(\frac{x_2}{1-x_1}\right) \phi\left(\frac{x_3}{1-x_1-x_2}\right)$. Note that, by construction, $P'\mathbf{e} = 1$, and $P \geq 0$. Since $\chi_4(\mathbf{x})$ and $\psi_4(\mathbf{x})$ share the same joint density for the order statistics, i.e. $4!\psi_4(\tilde{\mathbf{x}})$ where $\tilde{\mathbf{x}}$ has elements in descending order, to sample from $4!\psi_4(\tilde{\mathbf{x}})$ we sample first from $\chi_4(.)$ and sort the resulting vector in descending order.

A.4 Probabilities and Simulated Likelihood

Here we describe the evaluation of some of the probabilities that lead to the log likelihood. A complete description of the calculations can be obtained from the authors.

Calculating the probabilities when only one lifeline is available.

When the contestant has used all her lifelines, the events of interest are quitting or losing (and, for the last question, the event that the contestant wins the million prize). The probabilities of these events can be calculated directly from the analytical expressions given in section 5.1 using the formulation for *F* we derive in section 5.3. When one or more lifelines are available the calculations are made more complicated because of the information which is gained when the lifeline is used and which allows the contestant to update her belief. Hence, given the initial draw of the belief we determine whether this particular draw leads to the use of the lifeline and, if so, whether the updated belief, or the original belief, if the lifeline is not played, is informative enough to lead to an answer attempt. Finally, we evaluate the probability that the answer is correct (under the original or the updated belief). We will write $\Omega_{k,n}^{ijk}(\mathbf{p})$ as the probability that given **p** at stage *n* event *k* (which is defined precisely below) is observed, given that the contestant is in the lifeline-state *ijk*, where *i*, (respectively *j* or *k*) is one if the first (respectively second or third) lifeline is yet to be played and zero otherwise. Let, $\Omega_{k,n}^{ijk}$ be the expected value of $\Omega_{k,n}^{ijk}(\mathbf{p})$ over all possible realisations of **p**, i.e. $\Omega_{k,n}^{ijk} = \mathrm{E}\left[\Omega_{k,n}^{ijk}(\mathbf{P})\right]$. Finally $\Omega_{k,n}^{ijk,i'j'k'}(\mathbf{p})$ stand for the probability that given **p** at stage *n* event *k* is observed given that the contestant starts the question in the lifeline-state *ijk* and transit to lifeline-state *i'j'k'*. We consider below representative events for each lifeline.

“50:50” is the only lifeline available at stage n.

The contestant uses “50:50”, plays and wins (moves to the next stage or wins the million prize).

First, define the probability that 50:50 is used, she plays and wins, given a draw (ordered in decreasing order) **p** from the belief distribution, as $\Pr\left[\{50{:}50\}\cap\{\text{plays}\}\cap\{\text{wins}\}\,|\,\{\text{stage}_{n}\},\mathbf{p}\right] \equiv$

$\mathbf{\Omega}_{1,n}^{100}(\mathbf{p}) = \mathbf{1}_{\left[k_n^1(\mathbf{p},0,0,0)\geq p_1\left(f_{n-1}(1,0,0)-b_n\right)+b_n\right]}\mathbf{\Omega}_{1,n}^{100,000}(\mathbf{p}),$ where

$$\mathbf{\Omega}_{1,n}^{100,000}(\mathbf{p}) = \frac{1}{3}\sum_{j=1}^{3} p_j \sum_{k=j+1}^{4} \mathbf{1}\left[\pi_{jk}(\mathbf{p})\left(f_{n-1}(0,0,0)-b_n\right)+b_n \geq a_{n+1}\right], \quad \text{with}$$

$\pi_{jk} = \dfrac{p_j}{p_j+p_k}$. This last expression is the probability that, given **p**, the contestant correctly answers after using the lifeline. Hence, the unconditional probability satisfies

$$\Pr\left[\{50{:}50\}\cap\{\text{plays}\}\cap\{\text{wins}\}\,|\,\{\text{stage}_{n}\}\right] \equiv \mathbf{\Omega}_{1,n}^{100}$$
$$= \int_{\bar{\Delta}_4} \mathbf{1}_{\left[k_n^1(\mathbf{p},0,0,0)\geq p_1\left(f_{n-1}(1,0,0)-b_n\right)+b_n\right]}\mathbf{\Omega}_{1,n}^{100,000}(\mathbf{p})\,\psi_4^o(\mathbf{p})\,d\mathbf{p},$$

where $\bar{\Delta}_4$ is the subset of the 4-simplex where $p_1 \geq p_2 \geq p_3 \geq p_4 \geq 0$. In order to determine the probabilities we have used the fact that a contestant with a lifeline available will either use it (and perhaps then quit), or play. It is straightforward to verify that the five expressions above sum to unity; in particular the sum of the first three expressions is the probability that the contestant uses the lifeline and this is the complement of the sum of the last two probabilities. Each term of the sum that determines $\mathbf{\Omega}_1^{100}(\mathbf{p})$ (and similarly $\mathbf{\Omega}_2^{100}(\mathbf{p})$ and $\mathbf{\Omega}_3^{100}(\mathbf{p})$) is the product of the probability that a given two of the four options remain after the lifeline is played, with probability $(p_j+p_k)/3$, multiplied by the probability that the remaining alternative with the largest updated belief is correct, with probability $\pi_{jk}(\mathbf{p}) = \dfrac{p_j}{p_j+p_k}$ with $p_j \geq p_k$, multiplied by the indicator that, given the updated belief, the contestant decides to play.

“Ask the Audience” is the only lifeline left at stage n.

The contestant uses “Ask the Audience”, plays and loses,

$\Pr\left[\{\text{ATA}\}\cap\{\text{plays}\}\cap\{\text{loses}\}\,|\,\{\text{stage}_{\text{n}}\}\right]\equiv\Omega_{2,n}^{010}=\int_{\bar{\Delta}_4}\Omega_{2,n}^{010}(\mathbf{p})\psi_4^o(\mathbf{p})\,d\mathbf{p},$ where

$\boldsymbol{\Omega}_{2,n}^{010}(\mathbf{p})=\mathbf{1}\left[k_n^2(\mathbf{p},0,0,0)\geq p_1\left(f_{n-1}(0,1,0)-b_n\right)+b_n\right]\boldsymbol{\Omega}_{2,n}^{010,000}(\mathbf{p}),$ and

$\boldsymbol{\Omega}_{2,n}^{010,000}(\mathbf{p})=\int_{\Delta_4}\left(1-\pi_1(\mathbf{q};\mathbf{p})\right)\mathbf{1}\left[\pi_1(\mathbf{q};\mathbf{p})\left(f_{n-1}(0,0,0)-b_n\right)+b_n\geq a_{n+1}\right]D(\mathbf{q};\mathbf{p},\lambda,\nu)\,d\mathbf{q}$ where $\pi(\mathbf{q};\mathbf{p})$ stands for the revised belief after information vector $\mathbf{q}$ is made available and $\pi_1(\mathbf{q};\mathbf{p})$ is the largest element in $\pi(\mathbf{q};\mathbf{p})$.

"Phone a Friend" is the only lifeline left at stage n.

<u>The contestant uses "Phone a Friend" and quits,</u>

$\Pr\left[\{\text{PAF}\}\cap\{\text{quits}\}\,|\,\{\text{stage}_{\text{n}}\}\right]\equiv\Omega_{3,n}^{001}=\int_{\bar{\Delta}_4}\Omega_{3,n}^{001}(\mathbf{p})\psi_4^o(\mathbf{p})\,d\mathbf{p},$ where $\Omega_{3,n}^{001}(\mathbf{p})=\mathbf{1}_{\left[k_n^3(\mathbf{p},0,0,0)\geq p_1\left(f_{n-1}(0,0,1)-b_n\right)+b_n\right]}\boldsymbol{\Omega}_{3,n}^{001,000}(\mathbf{p})$ and $\boldsymbol{\Omega}_{3,n}^{001,000}(\mathbf{p})=(1-\kappa)\mathbf{1}_{\left[p_1\left(f_{n-1}(0,0,0)-b_n\right)+b_n<a_{n+1}\right]}$.

General Case: all the lifelines are available

When more than one lifeline is available at a given stage, the number of elementary events of interest increases, since not only can the contestants decide to play one lifeline among many but the contestant can play more than one lifeline to answer a single question. Hence while there are only five elementary events of interest when only one given lifeline is left there are nine such events when two lifelines are available and seventeen when all three lifelines are available, ignoring the order in which the contestant uses the lifeline and not counting events with zero probability ex-ante (for example observing an event such as quitting while the three lifelines are available). In this section we present the relevant expressions needed to obtain the probabilities of a few selected elementary events. All other probabilities can be obtained in a similar fashion.

<u>The contestant uses the three lifelines (in any order), plays and loses.</u>

$$\Pr\left[\{\text{uses all life lines}\}\cap\{\text{plays}\}\cap\{\text{loses}\}\,|\,\{\text{stage}_{\text{n}}\}\right]\equiv\boldsymbol{\Omega}_{2,n}^{111}$$
$$=\int_{\bar{\Delta}_4}\mathbf{1}_{\left[k_n^1(\mathbf{p},0,1,1)\geq\max\left\{p_1\left(f_{n-1}(1,1,1)-b_n\right)+b_n,\,k_n^2(\mathbf{p},1,0,1),k_n^3(\mathbf{p},1,1,0)\right\}\right]}\boldsymbol{\Omega}_{2,n}^{111,011}(\mathbf{p})$$
$$+\mathbf{1}_{\left[k_n^2(\mathbf{p},1,0,1)\geq\max\left\{p_1\left(f_{n-1}(1,1,1)-b_n\right)+b_n,\,k_n^1(\mathbf{p},0,1,1),k_n^3(\mathbf{p},1,1,0)\right\}\right]}\boldsymbol{\Omega}_{2,n}^{111,101}(\mathbf{p})$$
$$+\mathbf{1}_{\left[k_n^3(\mathbf{p},1,1,0)\geq\max\left\{p_1\left(f_{n-1}(1,1,1)-b_n\right)+b_n,\,k_n^1(\mathbf{p},0,1,1),k_n^2(\mathbf{p},1,0,1)\right\}\right]}\boldsymbol{\Omega}_{2,n}^{111,110}(\mathbf{p})\psi_4^o(\mathbf{p})\,d\mathbf{p}.$$

$\Omega_{2,n}^{111,011}(\mathbf{p}) = \frac{1}{3}\sum_{j=1}^{3}\sum_{k=j+1}^{4}\Omega_{2,n}^{011}\left(\left(\pi_{j,k}(\mathbf{p}),\pi_{k,j}(\mathbf{p}),0,0\right)\right)$, $\Omega_{2,n}^{111,101}(\mathbf{p}) = \int_{\Delta_4}\Omega_{2,n}^{101}\left(\pi(\mathbf{q};\mathbf{p})\right)D(\mathbf{q};\mathbf{p},\lambda,\nu)\,d\mathbf{q}$, and $\Omega_{2,n}^{111,110}(\mathbf{p}) = \kappa\Omega_{2,n}^{110}(1,0,0,0)+(1-\kappa)\Omega_{2,n}^{110}(\mathbf{p})$. Inspection of these expressions reveals that the probabilities of events in which more than one lifeline is available, here $\Omega_{2,n}^{111}$, can be defined recursively in terms of the conditional probability of events with one fewer lifeline, given the initial belief draw, here $\Omega_{2,n}^{011}(\mathbf{p})$, $\Omega_{2,n}^{101}(\mathbf{p})$ and $\Omega_{2,n}^{110}(\mathbf{p})$. In turn, each of these conditional probabilities can be calculated from conditional probabilities involving only one lifeline, i.e. $\Omega_{2,n}^{001}(\mathbf{p})$, $\Omega_{2,n}^{100}(\mathbf{p})$ and $\Omega_{2,n}^{010}(\mathbf{p})$. This property is a consequence of the recursive definition of the value function over the lifeline part of the state space (see section 5.4.b). Recall, however, that the number of events of interest when the three lifelines are available is larger than when only two or less are available. Hence the definition of 17 probabilities with three lifeline at stage *n*, i.e. $\Omega_{m,n}^{111}$, *m*=1…17, will involve the 27 conditional probabilities with two lifelines, i.e. $\Omega_{m,n}^{011}(\mathbf{p})$, $\Omega_{m,n}^{101}(\mathbf{p})$ and $\Omega_{m,n}^{110}(\mathbf{p})$, *m*=1…9. In turn each of these conditional probabilities will depend on the 15 probabilities with one lifeline as defined in the previous section, i.e. $\Omega_{m,n}^{100}(\mathbf{p})$, $\Omega_{m,n}^{010}(\mathbf{p})$ and $\Omega_{m,n}^{001}(\mathbf{p})$ m=1…5.

<u>The three lifelines are available, the contestant uses “50:50” , plays and wins.</u>

$$\Pr\left[\{\text{50:50 only}\}\cap\{\text{plays}\}\cap\{\text{wins}\}\,|\,\{\text{stage}_{\text{n}}\}\right]\equiv\mathbf{\Omega}_{10,n}^{111}$$
$$=\int_{\bar{\Delta}_4}\mathbf{1}_{\left[k_n^1(\mathbf{p},0,1,1)\geq\max\left\{p_1\left(f_{n-1}(1,1,1)-b_n\right)+b_n,\,k_n^2(\mathbf{p},1,0,1),\,k_n^3(\mathbf{p},1,1,0)\right\}\right]}\mathbf{\Omega}_{10,n}^{111,011}(\mathbf{p})\,\psi_4^o(\mathbf{p})\,d\mathbf{p}.$$

with $\Omega_{10,n}^{111,011}(\mathbf{p}) = \frac{1}{3}\sum_{j=1}^{3}\sum_{k=j+1}^{4}\Omega_{8,n}^{011}\left(\left(\pi_{j,k}(\mathbf{p}),\pi_{k,j}(\mathbf{p}),0,0\right)\right)$ where $\Omega_{8,n}^{011}(\mathbf{p})$ is the probability that with ATA and PAF available, for some belief **p**, the individual plays and wins.

<u>Three lifelines are available, the contestant does not use any, plays and loses.</u>

$$\Pr\left[\{\text{does not use any of the 3 life lines}\}\cap\{\text{plays}\}\cap\{\text{loses}\}\,|\,\{\text{stage n}\}\right]\equiv\mathbf{\Omega}_{17,n}^{111}$$
$$=\int_{\bar{\Delta}_4}\mathbf{1}_{\left[p_1\left(f_{n-1}(1,1,1)-b_n\right)+b_n\geq\max\left\{k_n^1(\mathbf{p},0,1,1),\,k_n^2(\mathbf{p},1,0,1),\,k_n^3(\mathbf{p},1,1,0)\right\}\right]}(1-p_1)\,\psi_4^o(\mathbf{p})\,d\mathbf{p}.$$

Simulation and smoothing

The evaluation of the probabilities $\Omega_{m,n}^{rst}(\mathbf{p})$, *n*=1..15, *m*=1..17, $(r,s,t)\in\{0,1\}^3$ and of the conditional expectations $k_n^j(\mathbf{p},r,s,t)$, *n*=1..15, *j*=1..3, and $(r,s,t)\in\{0,1\}^3$ requires the use multidimensional integration techniques. If $\Omega_{m,n}^{rst}$ is not defined for some *m*, and some *r,s,t* we assume $\Omega_{m,n}^{rst}=0$. Simulation methods (as described in Gouriéroux and Monfort (1996) and Train (2003)) are well suited and have been applied successfully in similar context (see the examples discussed in Adda and Cooper, (2003)).

Clearly the specification of the belief lends itself to a simulation based likelihood methodology since simulations of Beta variates are obtained simply from Gamma variates (see for example Poirier (1995)). In turn, Gamma variates can be obtained directly, using the inverse of the incomplete Gamma function. Numerically accurate methods to evaluate the inverse of the incomplete Gamma function are detailed in Didonato and Morris (1996). This is implemented in Gauss in the procedure **gammaii** (contained in the file cdfchic.src). The main advantage of their results is that it allows for simulations that are continuous in the parameters of the Gamma distributions. Evaluation by simulation of an integral involving the density of a 4 dimensional Dirichlet random vector, $D(\mathbf{q};\mathbf{p},\lambda,\nu)$, is obtained directly by the simulation of each of its component. For example $\Omega_{1,n}^{100}=\int_{\bar{\Delta}_4}\Omega_{1,n}^{100}(\mathbf{p})\psi_4^o(\mathbf{p})\,d\mathbf{p}=\int_{\bar{\Delta}_4}\mathbf{1}_{\left[k_n^1(\mathbf{p},0,0,0)\geq p_1\left(f_{n-1}(1,0,0)-b_n\right)+b_n\right]}\Omega_{1,n}^{100,000}(\mathbf{p})\psi_4^o(\mathbf{p})\,d\mathbf{p},$ can be approximated by $\widehat{\Omega}_{1,n}^{100}(S)=\frac{1}{S}\sum_{s=1}^{S}\Omega_{1,n}^{100}(\mathbf{p}_s)=\frac{1}{S}\sum_{s=1}^{S}\mathbf{1}_{\left[k_n^1(\mathbf{p}_s,0,0,0)\geq p_{1,s}\left(f_{n-1}(1,0,0)-b_n\right)+b_n\right]}\Omega_{1,n}^{100,000}(\mathbf{p}_s),$ where $\mathbf{p}_s$ is one of S (the number of simulations) independent draws from the distribution of the order statistics of the belief, $\psi_4^o(.)$. In fact the accuracy of this simulated probability, and of all others which involve draws from $\psi_4^o(.)$, can be improved through antithetic variance reduction techniques which involve the permutations of the gamma variates used to generate each individual beta variate (as explained, for example, in Train (2003)). Moreover, $\Omega_{10,n}^{111}=\int_{\bar{\Delta}_4}\mathbf{1}_{\left[k_n^1(\mathbf{p},0,1,1)\geq\max\left\{p_1\left(f_{n-1}(1,1,1)-b_n\right)+b_n,k_n^2(\mathbf{p},1,0,1),k_n^3(\mathbf{p},1,1,0)\right\}\right]}\Omega_{10,n}^{111,011}(\mathbf{p})\psi_4^o(\mathbf{p})\,d\mathbf{p}.$ is a quantity that can be evaluated using the simpler formula $\hat{\Omega}_{10,n}^{111}(S)=\frac{1}{S}\sum_{s=1}^{S}\mathbf{1}_{\left[k_n^1(\mathbf{p}_s,0,1,1)\geq\max\left\{p_{1,s}\left(f_{n-1}(1,1,1)-b_n\right)+b_n,k_n^2(\mathbf{p}_s,1,0,1),k_n^3(\mathbf{p}_s,1,1,0)\right\}\right]}\Omega_{10,n}^{111,011}(\mathbf{p}_s),$ or any improvement of it. Similarly $\Omega_{2,n}^{111,101}(\mathbf{p})=\int_{\Delta_4}\Omega_{2,n}^{101}\left(\pi(\mathbf{q};\mathbf{p})\right)D(\mathbf{q};\mathbf{p},\lambda,\nu)\,d\mathbf{q}$ can be evaluated by $\hat{\Omega}_{2,n}^{111,101}(\mathbf{p};S)=\frac{1}{S}\sum_{i=1}^{4}p_i\sum_{s=1}^{S}\left[\Omega_{2,n}^{101}\left(\pi\left(\mathbf{q}_{s,i};\mathbf{p}\right)\right)\right]$, where $\mathbf{q}_{s,i}$ is one of *S* independent draws from $D\left(\mathbf{q};\gamma_i(\lambda,\nu)\right)$. In practice, we use S=96.

Finally all quantities $k_n^2(\mathbf{p},r,s,t) \equiv \mathrm{E}_{\pi_2|\mathbf{p}}\left[W_n(\underset{\sim}{\Pi},r,s,t)\,|\,\mathbf{p}\right]$ which involve a multi dimensional integral and the joint density $D(\mathbf{q};\mathbf{p},\lambda,\nu)$ can be obtained in a similar fashion: for example, using $\hat{k}_n^2(\mathbf{p},r,s,t;S) = \frac{1}{\mathrm{S}}\sum_{\mathrm{i}=1}^{4} p_i \sum_{s=1}^{S} W_n(\mathbf{q}_{s,i},r,s,t)$, where $\mathbf{q}_{s,i}$ is one of S independent draws from $D(\mathbf{q};\gamma_i(\lambda,\nu))$. In practice these expression are modified in order to smooth out the discontinuities that are created by the indicator terms. Hence, the indicator functions $\mathbf{1}_{[v_1 \geq \max\{v_2,v_3,v_4\}]}$, $\mathbf{1}_{[v_1 \geq \max\{v_2,v_3\}]}$, or $\mathbf{1}_{[v_1 \geq v_2]}$, are replaced by their smoothed versions, $\frac{1}{1+\mathrm{e}^{(\eta(v_2-v_1))}+\mathrm{e}^{(\eta(v_3-v_1))}+\mathrm{e}^{(\eta(v_4-v_1))}}$, $\frac{1}{1+\mathrm{e}^{(\eta(v_2-v_1))}+\mathrm{e}^{(\eta(v_3-v_1))}}$, and $\frac{1}{1+\mathrm{e}^{(\eta(v_2-v_1))}}$ respectively, where η is a smoothing constant. In the limit as $\eta \to +\infty$ the smoothed versions tend to the indicators.

References

Adda, J. and R. Cooper (2003), Dynamic Economics, MIT Press: Cambridge, MA.

Didonato, A. and A.H. Morris (1986), “Computation of the Incomplete Gamma Function Ratios and their Inverse”, ACM Transactions on Mathematical Software, 12, 377-393. DOI:10.1145/22721.23109.

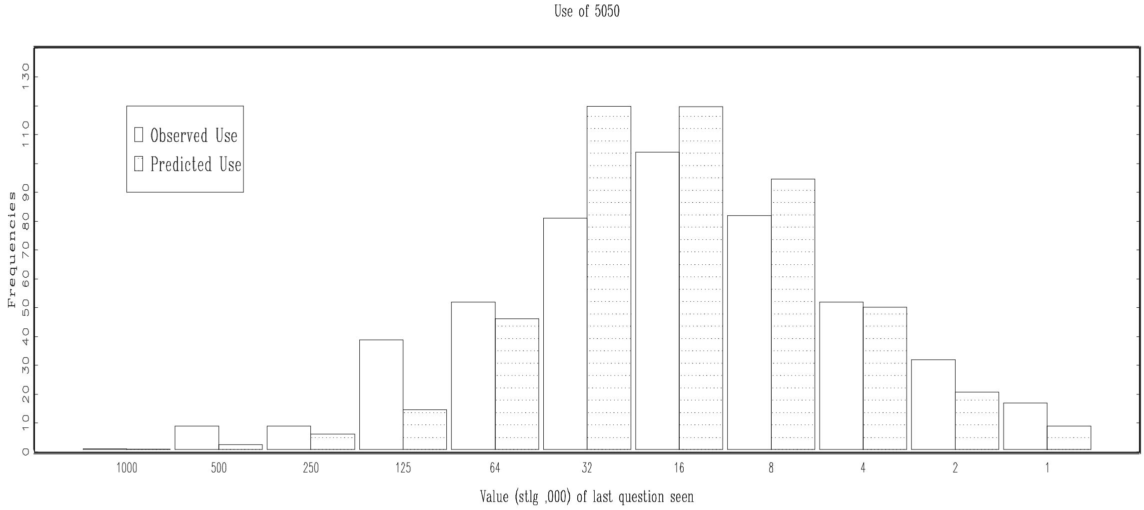
French, S. (1980), “Updating of Belief in the Light of Someone Else’s Opinion”, Journal of the Royal Statistical Society, 143, 43-48. DOI:10.2307/2981768.

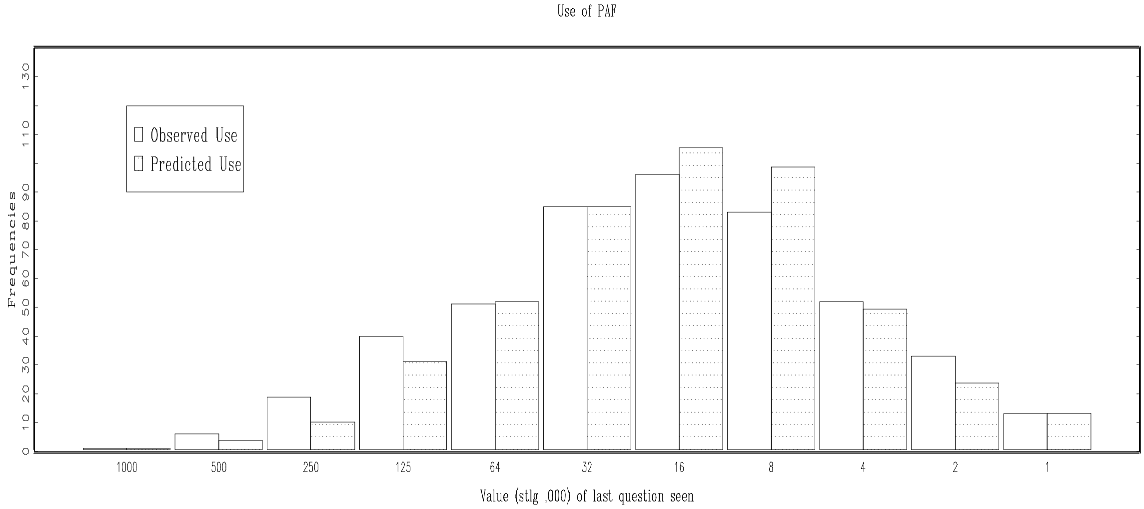
Gouriéroux, C. and A. Monfort (1996), Simulation-Based Econometric Methods, Oxford University Press: Oxford.

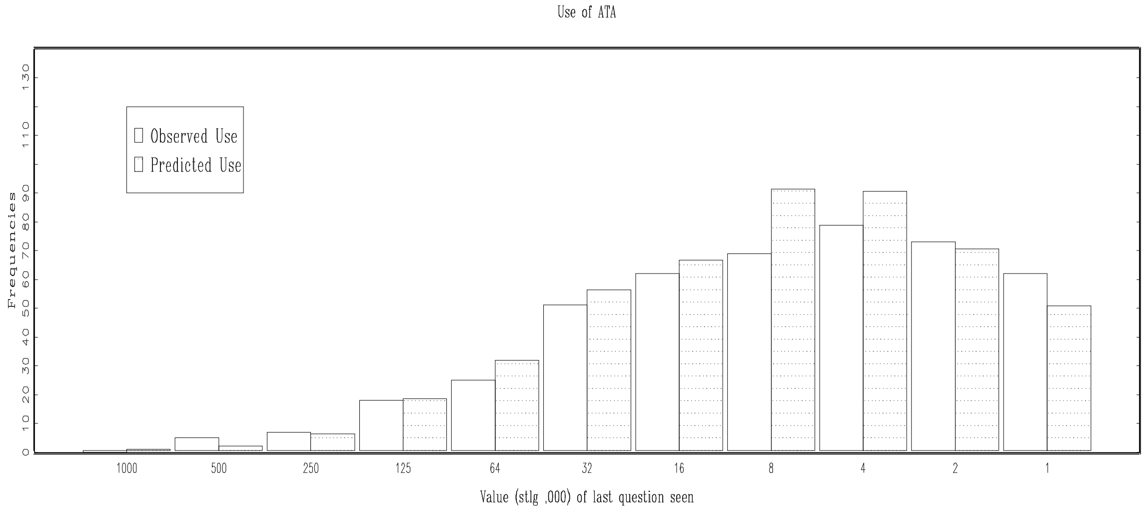
Poirer, D. (1995) Intermediate Statistics and Econometrics: A Comparative Approach, Cambridge, MA: MIT Press.

Train, K.E. (2003) Discrete Choice Methods with Simulation, Cambridge University Press: Cambridge.

Figure A1 Observed versus Predicted Use of Lifelines







Note: Predicted frequencies shown as shaded bars, observed frequencies as empty bars.