Supplemental Appendix for "The Power of Anecdotes"

By ARJADA BARDHI AND NINA BOBKOVA*

A. OMITTED PROOFS

A1. Belief updating

Suppose case j has been observed, and the receiver forms beliefs about case k. Note that $\Pr[x_j = 1, x_k = 1] = \mathbb{E}[x_j x_k]$ because of binary outcomes, and moreover $\operatorname{cov}(x_j, x_k) = \mathbb{E}[x_j, x_k] - \mathbb{E}[x_j]\mathbb{E}[x_k]$. By applying Bayes' rule, the belief after $x_k = 1$ is

$$\Pr[x_k = 1 \mid x_j = 1] = \frac{\Pr[x_j = 1, x_k = 1]}{p_j} = \frac{\operatorname{cov}(x_j, x_k) + p_j p_k}{p_j} = p_k + \frac{\operatorname{cov}(x_j, x_k)}{p_j}.$$

Since $\operatorname{cov}(x_j, x_k) = \rho^{|j-k|}\sigma_j\sigma_k$ and $\sigma_j/p_j = r_j$, this yields $p_k + \rho^{|j-k|}\sigma_k \cdot r_j$. The expression for $x_j = 0$ follows analogously, replacing p_j with $1 - p_j$ in the denominator and noting that $\sigma_j/(1-p_j) = 1/r_j$. Updating about statistic s follows by summing over k:

$$\mathbb{E}[s \mid x_j = 1] = \sum_{k=1}^{N} \Pr[x_k = 1 \mid x_j = 1] = \sum_{k=1}^{N} p_k + r_j \sum_{k=1}^{N} \rho^{|j-k|}\sigma_k = \mathbb{E}[s] + r_j \cdot R_j.$$

The calculation for $x_j = 0$ is analogous.

A2. Informativeness of an anecdote about the statistic

The variance of posterior beliefs is

$$\operatorname{var}(\mathbb{E}[s \mid x_j]) = p_j(H_j - \mathbb{E}[s])^2 + (1-p_j)(L_j - \mathbb{E}[s])^2 = p_j \cdot r_j^2 R_j^2 + (1-p_j) \cdot \frac{R_j^2}{r_j^2}.$$

Substituting success rarity $r_j^2 = (1-p_j)/p_j$ reduces the variance expression to $p_j \cdot \frac{1-p_j}{p_j} R_j^2 + (1 - p_j) \cdot \frac{p_j}{1-p_j} R_j^2 = (1-p_j)R_j^2 + p_j R_j^2 = R_j^2$.

B. LARGE N

With finite N, revealing one case is a nontrivial fraction of the whole. To isolate the pure effect of correlation—where a single case is negligible in and of itself but informative only through its correlation with others—we consider cases indexed by $k \in [0,1]$ with binary outcomes $X(k) \in \{0,1\}$, prior success probabilities $p_k = \Pr[X(k) = 1]$, and aggregate statistic $s = \int_0^1 X(k)\,\mathrm{d}k$. As in our baseline model, the receiver accepts if and only if $\mathbb{E}[s \mid \text{evidence}] \geq \tau$, where now the prior belief about the statistic is

$$\mathbb{E}[s] = \int_0^1 p_k\,\mathrm{d}k.$$

Let the distance-based correlation over $[0,1]$ be of the exponential decay form:

$$\operatorname{corr}[X(j), X(k)] = e^{-\lambda|j-k|}$$

* Bardhi: Department of Economics, New York University, arjada.bardhi@nyu.edu; Bobkova: Department of Economics, Rice University, nina.bobkova@rice.edu.

for some decay rate $\lambda > 0$.[1] Note that $\lambda \to 0$ corresponds to perfect correlation and $\lambda \to +\infty$ corresponds to independence.

A parametric family of priors that is convenient to consider in conjuction with exponential decay correlation is that of the logistic form:

$$p_k = \frac{1}{1+e^{-\mu(k-1/2)}},$$

where $\mu \geqslant 0$ governs prior heterogeneity. When $\mu = 0$, priors are uniform: $p_k = 1/2$ for all k. As μ increases, priors range from low at $k = 0$ to high at $k = 1$, with $p_{1/2} = 1/2$ always.

Not all (μ, λ)-pairs yield a well-defined statistical structure. As in Assumption 1, correlation needs to respect the Fréchet upper bound. For a logistic prior with exponential decay, the feasibility condition takes a convenient and simple form: $e^{-\lambda|j-k|} \leqslant e^{-\mu|j-k|/2}$, or equivalently $\lambda \geqslant \mu/2$. Greater prior heterogeneity (higher μ) demands faster correlation decay (higher λ). Alternatively, for any fixed level of correlation, prior heterogeneity cannot be too large.

Next, we turn to the discussion of reach, rarity, and the optimal anecdote for this parametric family. Define $a_j := e^{-\mu(j-1/2)}$, so that $p_j = 1/(1+a_j)$. Hence, success **rarity** is $r_j = \sqrt{a_j}$; it is weakly decreasing in j for $\mu \geqslant 0$. Prior uncertainty $\sigma_j = \sqrt{p_j(1-p_j)}$, on the other hand, is symmetric and single peaked around $j = 1/2$, with $\sigma_{1/2} = 1/2$. That is, cases further away from the center case are symmetrically more uncertain.

As in the baseline model, the **reach** of case j is $R_j = \int_0^1 e^{-\lambda|j-k|}\sigma_k \,\mathrm{d}k$, where $\sigma_k = \sqrt{p_k(1-p_k)}$. With uniform priors ($\mu = 0$), we have $\sigma_k = 1/2$ for all k, and reach admits a closed form:

$$R_j = \frac{1}{\lambda} - \frac{e^{-\lambda j} + e^{-\lambda(1-j)}}{2\lambda}.$$

This is symmetric around $j = 1/2$ and maximized there: $R_{1/2} = \frac{1}{\lambda}(1-e^{-\lambda/2})$. At the boundaries, $R_0 = R_1 = \frac{1}{2\lambda}(1-e^{-\lambda})$. Central cases have greater reach because they have neighbors on both sides. As $\lambda \to 0$ (perfect correlation), $R_j \to \int_0^1 \sigma_k \,\mathrm{d}k = 1/2$ for all j: every case is equally connected to every other, so position becomes irrelevant. As $\lambda \to \infty$ (independence), $R_j \to 0$ for all $j \in (0,1)$: every case is negligible in terms of informativeness about the statistic.

Lemma B.1. *For any (μ, λ) such that $\lambda \geqslant \mu/2$, R_j is symmetric and single-peaked around $j = 1/2$ on $[0,1]$.*

PROOF:

Symmetry. The logistic prior satisfies $p_{1-k} = 1 - p_k$, so $\sigma_{1-k} = \sigma_k$. Substituting $u = 1-k$ in R_{1-j}:

$$R_{1-j} = \int_0^1 e^{-\lambda|1-j-k|}\sigma_k \,\mathrm{d}k = \int_0^1 e^{-\lambda|u-j|}\sigma_{1-u}\,\mathrm{d}u = \int_0^1 e^{-\lambda|u-j|}\sigma_u \,\mathrm{d}u = R_j.$$

Single-peakedness. Differentiating:

$$R'_j = \lambda\left[\int_0^{1-j} e^{-\lambda u}\sigma_{j+u}\,\mathrm{d}u - \int_0^{j} e^{-\lambda v}\sigma_{j-v}\,\mathrm{d}v\right].$$

For $j < 1/2$, split the first integral at $u = j$:

$$R'_j = \lambda\left[\int_0^{j} e^{-\lambda u}(\sigma_{j+u} - \sigma_{j-u})\,\mathrm{d}u + \int_j^{1-j} e^{-\lambda u}\sigma_{j+u}\,\mathrm{d}u\right].$$

The second term is positive. For the first, since σ_k is symmetric and single-peaked at $k = 1/2$, case

[1] Ichihashi, Li, and Zou (2025) consider exclusively the case of uniform prior in a related Poisson setting.

$j+u$ is closer to $1/2$ than $j-u$ when $j<1/2$, so $\sigma_{j+u}>\sigma_{j-u}$. Both terms are positive, hence $R'_j>0$ for $j<1/2$. By symmetry, $R'_j<0$ for $j>1/2$.

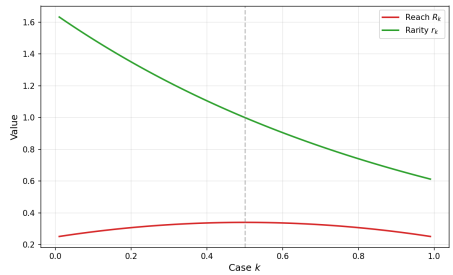

FIGURE B1. REACH (RED) VS. RARITY (GREEN) FOR $\mu=2,\lambda=1.5$.

Optimal anecdote. The difference between the success posterior and the prior belief is $H_j-\mathbb{E}[s]=r_j\cdot R_j$, where rarity $r_j=\sqrt{(1-p_j)/p_j}$ and reach R_j are as defined above. With uniform priors ($\mu=0$), success rarity is constant across cases: $r_j=1$ for all j. Persuasive power is determined entirely by reach, so the most persuasive anecdote is the most central one, $j^*=1/2$. This is similar to our baseline model.

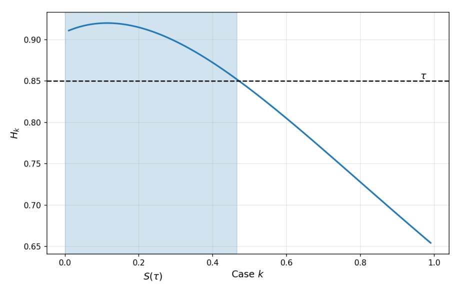

FIGURE B2. PERSUASION SET $S(\tau)$ FOR $\mu=2,\lambda=1.5,\tau=0.85$

With heterogeneous priors ($\mu>0$), rarity decreases in r_j whereas R_j is symmetric and single-peaked. See Figure B.1 above for an illustration. For this numerical example, the success posterior belief is single-peaked at a case to the left of the center case. Figure B.2 illustrates the persuasion set $S(\tau)$, which consists of an interval of leftmost cases. The optimal anecdote is the rightmost case in this persuasion set, which in this example corresponds to $j^*\approx 0.465$ with $p_{j^*}\approx 0.47$.

Therefore, the logic of the discrete baseline model goes through to a continuum of cases—and the logistic prior with exponential decay gives further structure to the rarity-reach tradeoff. Feasibility constrains prior heterogeneity: for any given level of correlation, priors cannot vary too much. When priors are uniform, rarity is constant and reach determines the optimal anecdote. We conjecture that with prior heterogeneity, (i) success posterior is always single-peaked, hence $S(\tau)$ is a single interval, (ii) the optimal anecdote is always the rightmost case in this interval $S(\tau)$, and (iii) the greater the prior heterogeneity, the further to the left of $1/2$ is the case with the highest H_j—and hence, for sufficiently high thresholds, the optimal anecdote.