Framing Effects in Survey Research: Supplementary Materials

by Jacob Goldin and Daniel Reck

Appendix A: Proofs of Propositions

Proof of Proposition 2.2 Let $\alpha = p[(y_i(d_0), y_i(d_1)) = (1, 0)]$. Frame monotonicity imposes $\alpha = 0$, which leads to the result in 2.1. Otherwise, it follows from the four possibilities for $(y_i(d_0), y_i(d_1))$ that

$$Y_0 = \pi_c E[c_i] + \alpha, \tag{1}$$

and also that

$$1 - Y_1 = (1 - \pi_c)E[c_i] + \alpha. \tag{2}$$

To obtain bounds for $E[c_i]$, note that by (1) and (2),

$$Y_0 + 1 - Y_1 = E[c_i] + 2\alpha. \tag{3}$$

The upper bound for $E[c_i]$ is immediately implied by (3) and the fact that $\alpha \geq 0$. The lower bounds for $E[c_i]$ are implied by the fact that $\alpha \leq Y_0$, and $\alpha \leq 1 - Y_1$ from (1) and (2). Combining these with (3) we obtain $E[c_i] \geq 1 - Y_1 - Y_0$ and $E[c_i] \geq Y_0 + Y_1 - 1$. Combining these yields desired the given lower bound.[1]

To obtain bounds for π_c, we combine 1 and 2 in the definition of Y_c:

$$Y_c = \frac{\pi_c E[c_i] + \alpha}{E[c_i] + 2\alpha}. \tag{4}$$

Subtract $\frac{1}{2}$ from both sides of equation (4) and simplify to obtain

$$Y_c - \frac{1}{2} = \left(\pi_c - \frac{1}{2}\right) \frac{E[c_i]}{E[c_i] + 2\alpha}. \tag{5}$$

Noting that $\frac{E[c_i]}{E[c_i]+2\alpha} \in [0, 1]$, equation 5 implies the desired bounds for π_c. ■

Proof of the Corollary to Proposition 2.2 Using the expression for Y_c in (4) and the expression for Y_p in Proposition 1, it follows that

$$Y_c - Y_p = \frac{(\pi_c - \frac{1}{2})P(c_i = 1)[1 - P(c_i = 1) - 2\alpha]}{P(c_i = 1) + 2\alpha}. \tag{6}$$

One can derive that $\alpha < \frac{1}{2}(1 - E[c_i])$ from the definitions of Y_0 and Y_1,[2] so $1 - P(c_i = 1) - 2\alpha > 0$. When $Y_c \geq \frac{1}{2}$, we know that $\pi_c \geq Y_c$ from 2.2 and

[1] Note that which of these two constraints binds depends on whether $Y_0 \gtrless 1 - Y_1$, i.e. whether $Y_c \gtrless \frac{1}{2}$, so the two cases match the two cases of the second part of the result.

[2] Specifically, we assumed $Y_1 > Y_0$. Plugging the formulas for Y_1 and Y_0 in the proof of 2.1 into this inequality gives the result. An interesting special case occurs in the limiting case where $\alpha = \frac{1}{2}(1 - E[c_i])$, which also implies $Y_0 = Y_1$. In this case there are just as many frame defiers as frame compliers, and the two cancel each other out in the observations of Y_1 and Y_0. Then, although $Y_1 = Y_0$ and the aggregtae data show no inconsistency, π_c and $E[c_i]$ can take *any value* from zero to one. We rule this out when we assume that framing effects are observed, i.e. $Y_1 \neq Y_0$, but it is important to note the implication of this reasoning: framing effects can be present even when they are unobserved.

$Y_c \geq Y_p$ from equation (6). Once we know that $\pi_c \geq Y_c \geq Y_p$, the desired result is established for this case. A similar line of reasoning establishes the result for the case when $Y_c \leq \frac{1}{2}$. In the knife-edge case where $Y_c = \frac{1}{2}$, we know from 2.2 that $\pi_c = \frac{1}{2}$, and then Proposition 1 implies that $Y_p = \frac{1}{2}$. ■

Proof of Proposition 3 First note that by the same logic as in Proposition 2.1, conditional frame exogeneity and frame monotonicity – along with the law of iterated expectations – imply that

$$P(c_i = 1|g_i = g) = E_g[Y_0(g) + 1 - Y_1(g)]. \tag{7}$$

Bayes' rule implies that

$$P(g_i = g|c_i = 1) = \frac{P(c_i = 1|g_i = g)P(g_i = g)}{P(c_i = 1)}. \tag{8}$$

Substituting (7) and the expression for $P(c_i = 1)$ obtained in Proposition 2.1 yields 3.1. The proof of 3.2 is analogous. ■

Proof of Proposition 4.1 The law of iterated expectations implies

$$E[y_i^*] = E[y_i^*|c_i = 1]P(c_i = 1) + E[y_i^*|c_i = 0]P(c_i = 0). \tag{9}$$

Under Assumption 4, $\pi_c = E[y_i^*|c_i = 1]$. Applying what we know about π_c and $P(c_i = 1)$ from Proposition 2.1 and simplifying yields $E[y_i^*] = Y_0 + E[y_i^*|c_i = 0](Y_1 - Y_0)$. The bounds then follow from the fact that $0 \leq E[y_i^*|c_i = 0] \leq 1$. ■

Proof of 4.2 We first consider the lower bound. Because $E[y_i^*|c_i = 0] \geq 0$, equation (9) implies that$E[y_i^*] \geq \pi_c P(c_i = 1)$. This is equivalent to $E[y_i^*] \geq Y_0 - \alpha$ by equation (1). Finding a suitable maximum value for α will yield a lower bound. Re-arranging equation (2) yields $1 - \pi_c = \frac{1 - Y_1 - \alpha}{E_c}$, and $1 - \pi_c \leq 1$, so $\alpha \leq 1 - Y_1$. Substituting our maximum value for α gives the desired lower bound. Note that this lower bound is non-trivial only when $Y_c \geq \frac{1}{2}$; otherwise $Y_0 - (1 - Y_1) \leq 0$.

We next derive an upper bound using similar reasoning. Because $E[y_i^*|c_i = 0] \leq 1$, equation (9) implies that$E[y_i^*] \leq \pi_c P(c_i = 1) + 1 - P(c_i = 1)$. This is equivalent to $E[y_i^*] \leq Y_1 + \alpha$ by equation(2). Once again we require a suitable maximum value for α. In this case the non-trivial bound comes from equation (1), which implies $\pi_c = \frac{Y_0 - \alpha}{E[c_i]} \geq 0$, so $\alpha \leq Y_0$. Note that this upper bound is non-trivial only when $Y_c \leq \frac{1}{2}$; otherwise $Y_0 + Y_1 \geq 1$. ■

Proof of Proposition 5 By the law of iterated expectations, $E[y_i^*] = E_g[E[y_i^*|g_i = g]]$. Proposition 2.1 implies that for any g, $Y_c(g) = E[y_i^*|g_i = g, c_i = 1]$. Conditional consistency independence then implies that $Y_c(g) = E[y_i^*|g_i = g]$.[3] Substituting this into the first expression yields the result. ■

[3]Recall that two-valued variables are independent if they have zero covariance. One can also show formally that $E[y_i^*] = Y_c - \frac{cov(c_i, y_i^*)}{P(c_i = 1)}$. Conditioning this on g and taking its expectation also leads to the desired result.

Appendix B: Finite Sample Estimation and Standard Errors

The body of the paper ignores finite sample concerns, but any empirical application should account for finite sample concerns and report standard errors for the estimators we propose. This section describes estimation procedures in greater detail and derives estimators of asymptotic varianceand standard errors of the estimators of the population parameters of interest. Because all variables we work with in this paper are discrete and the data are presumed to come from independent random sampling of the population, our derivations here rely only on the characteristics of the binomial and multinomial distributions, along with the delta method. The standard errors we derive are incorporated into Stata program files that are available upon request from the authors.

Throughout this appendix, we let $\overline{X}$ denote the finite sample sample analogue of some population-level statistic X.

Standard Errors for Proposition 2

Proposition 2 introduced the statistic $Y_c = \frac{Y_0}{Y_0+1-Y_1}$ as an estimator for the consistent response share $\pi_c = E[y_i^*|c_i = 1]$. Note that $y_i(d_0)$ and $y_i(d_1)$ are two-valued random variables. From the properties of the binomial distribution, we therefore know that

$$\sqrt{n}\begin{pmatrix} Y_0 - \overline{Y}_1 \\ Y_1 - \overline{Y}_1 \end{pmatrix} \overset{a}{\sim} N(0, \Sigma)$$

where

$$\Sigma = \begin{pmatrix} \frac{1}{\alpha_0}\overline{Y}(0)(1-\overline{Y}(0)) & 0 \\ 0 & \frac{1}{\alpha_1}\overline{Y}(1)(1-\overline{Y}(1)) \end{pmatrix}$$

and $\alpha_j = \frac{n_j}{n}$ for $j \in \{0, 1\}$, where n_j denotes the number of individuals observed under frame d_j.

The finite-sample analogue of Y_c is

$$\overline{Y}_c = \frac{\overline{Y}_0}{\overline{Y}_0 + 1 - \overline{Y}_1}$$

Using the delta method, we know that $V(\overline{Y}_c) \simeq \frac{1}{n}\nabla Y_c' \, \Sigma \, \nabla Y_c$, where $\nabla \overline{Y}_C$ is evaluated at (Y_0, Y_1). Taking derivatives of the expression for Y_c yields

$$\nabla Y_c = \begin{pmatrix} \frac{1-Y_1}{(Y_0+1-Y_1)^2} \\ \frac{Y_0}{(Y_0+1-Y_1)^2} \end{pmatrix}.$$

Simplifying the expression for $V(\overline{Y}_c)$ yields

$$V(\overline{Y}_c) \simeq \frac{(1-Y_1)^2 Y_0(1-Y_0)}{\alpha_0(Y_0+1-Y_1)^4 n} + \frac{Y_0 Y_1(1-Y_1)}{\alpha_1(Y_0+1-Y_1)^4 n}$$

which, letting $C = Y_0 + 1 - Y_1$ be the fraction of the population that are consistent, and using that $V(\overline{Y}_j) = \frac{Y_j(1-Y_j)}{n_j}$ for $j = 0, 1$, simplifies to

$$V(\overline{Y}_c) = \left(\frac{1 - Y_c}{C}\right)^2 V(\bar{Y}_0) + \left(\frac{Y_c}{C}\right)^2 V(\bar{Y}_1) \tag{10}$$

We can consistently estimate the variance of the asymptotic distribution of our estimator $\hat{y}_c$ by replacing all the terms in (10) with the corresponding sample means, i.e. replacing Y_0 with $\overline{Y}_0$, and Y_1 with $\overline{Y}_1$.

Note also that we can estimate the variance of the asymptotic distribution of our estimator for $\bar{c} \equiv E[C_i]$. Write the estimator itself as:

$$\bar{C} = \bar{Y}_0 + 1 - \bar{Y}_1$$

and the variance of this estimator is simply:

$$V(\overline{C}) = V(\bar{Y}_0) + V(\bar{Y}_1) \tag{11}$$

Standard Errors for Propositions 3 and 5

The standard errors for the results splitting the sample by observable characteristics are considerably more complicated, due to the presence of multiple demographic groups and the use of probabilities that must themselves be estimated from data and are then used as weights in the re-weighting prodedure in Prpoosition 5. We discuss two solutions here, one built around linear regression and a non-parametric bootstrap, and another for directly calcualting analytic standard errors. We use the latter of these in the illustration in the paper.

First, one can simply estimate the following via least squares regression:

$$E[y_i|d_i = d, g_i = g] = f_1(g, \theta_1) + 1\{d_i = d_1\} f_2(g, \theta_2).$$

where $f_1()$ and $f_2()$ are specified up to vectors of parameters θ_1 and θ_2, which are estimated from data. For example, we could implement a linear model with uni-dimensional observables g:

$$E[y_i|d, g] = \alpha + \beta g + 1\{d_i = d_1\}(\gamma + \delta g)$$

This equation can be estimated by a least squares linear probability model, and then the ingredients of the matching estimator are given by:

$$\overline{C}(g) = 1 - \hat{\gamma} - \hat{\delta} g$$
$$\overline{C} = \frac{1}{n} \sum_i (1 - \hat{\gamma} - \hat{\delta} X_i)$$

These two , along with the finite-sample estimators for the distribution of g in the population (which are also simple sample means when G is discrete), can be combined straighforwardly to estimate the statistics used in Proposition 3:

$$\overline{P(g_i = g|c_i = 1)} = \frac{\overline{C}(g)}{\overline{C}} P(g_i = g),$$

$$\overline{P(g_i = g|c_i = 0)} = \frac{1 - \overline{C}(g)}{1 - \overline{C}} P(g_i = g).$$

When $P(g_i = g)$ is known with negligible sampling variation, the variance of these estimators can be obtained directly using post-regression-estimation commands available in most statistical packages, with associated analytic delta method standard errors. Otherwise, one should use a non-parametric bootstrap to obtain standard errors for the estimators for Proposition 3 that properly correct for uncertainty about $P(g_i = g)$ for each value of g.

For the matching-and-reweighting approach described by Proposition 5, we can estimate the parameters of interest in our regression model as follows:

$$\overline{Y}_c(g) = \frac{\hat{\alpha} + \hat{\beta} g}{1 - \hat{\gamma} - \hat{\delta} g}$$

$$\overline{Y}[y_i^*] = \frac{1}{n} \sum_i \overline{Y}_c(g_i)$$

$$\overline{Y}[Y_i^*|C_i = 0] = \frac{1}{n} \sum_i \frac{1 - \overline{C}(g_i)}{1 - \overline{C}} \hat{Y}_c(g_i)$$

One can then obtain standard errors for these estimators via a non-parametric bootstrap.

Note that all of the above can be applied regardless of the properties of the set of observables G. When G is discrete, taking values $g_1, ..., g_J$ one may derive delta-method-based standard errors directly. We provide analytical formulae for the variance and the gradients of parameters of interest, which may be straightforwardly incorporated into a matrix-based programming language, such as MATLAB, to calculate the variance of the estimators with discrete demographic groups. MATLAB code illustrating this procedure is also available upon request from the authors. This technique is employed in the illustration in the paper.

The primitive parameters of the discrete-characteristics model are, for each g,Y_{0g}, Y_{1g}, and $p(g_i = g) = p(g_i = g)$. When the g's are non-stochastic, such as when the researcher wishes to estimate preferences for a population with a known distribution of observable characteristics, the last of these may be excluded; the resulting modification of the variance estimation procedure below is straightforward. We denote the estimators of these quantities by $\bar{Y}_{0g}$, $\bar{Y}_{1g}$, and $\overline{p}_g$. Now we construct the variance covariance matrix of the vector primitive parameters. Letting $\theta = (Y_{0w_1}, Y_{0w_2}, ..., Y_{0w_J}, Y_{1w_1}, ..., Y_{1w_J}, p_{g_1}, ..., p_{g_J})'$, and $\hat{\theta} = (\bar{Y}_{0w_1}, \bar{Y}_{0w_2}, ..., \bar{Y}_{0w_J}, \bar{Y}_{0w_1}, ..., \bar{Y}_{0w_J}, \bar{p}_{g_1}, ..., \bar{p}_{g_J})$, we know from the central limit theorem that

$$\sqrt{n}(\hat{\theta} - \theta) \overset{a}{\sim} N(\overrightarrow{0}, \Sigma).$$

Denoting the fraction of individuals with observable characteristic w observed in frame d_j by $\alpha_{jw} \equiv \frac{n_{jw}}{n}$, we can write the variance matrix as:

$$\Sigma(\theta) = \begin{pmatrix} \frac{Y_{0w_1}(1-Y_{0g_1})}{\alpha_{0g_1}} & & & & & & & & & \\ & \cdots & & & & & & & & \\ & & \frac{Y_{0g_J}(1-Y_{0g_J})}{\alpha_{0g_J}} & & & & & & & \\ & & & \frac{Y_{1g_1}(1-Y_{1g_1})}{\alpha_{1g_1}} & & & & & & \\ & & & & \cdots & & & & & \\ & & & & & \frac{Y_{1g_J}(1-Y_{1g_J})}{\alpha_{1g_J}} & & & & \\ & & & & & & p_{g_1}(1-p_{g_1}) & -p_{g_1}p_{g_2} & \cdots & -p_{g_1}p_{g_J} \\ & & & & & & -p_{g_1}p_{g_2} & p_{g_2}(1-p_{g_2}) & \cdots & -p_{g_2}p_{g_J} \\ & & & & & & \cdots & \cdots & \cdots & \cdots \\ & & & & & & -p_{g_1}p_{g_J} & -p_{g_2}p_{g_J} & \cdots & p_{g_J}(1-p_{g_J}) \end{pmatrix}$$

where all blank entries of the Σ matrix are zeroes.[4]

The matching approach employs many different combinations of these primitive parameters. We begin with the weights for the subset of inconsistent choosers,

$$s(g) \equiv p(g_i = g | c_i = 0) = \frac{Y_{1g} - Y_{0g}}{\sum_{g=h}(Y_{1h} - Y_{oh})p_h} p_g$$

Carefully taking derivatives of this function and simplifying using the definition of $s(g)$, we obtain the following for any g and g':

$$\frac{\partial s(g)}{\partial Y_{0g'}} = -\frac{1\{g = g'\} - s(g)}{\sum_{g=h}(Y_{1h} - Y_{oh})p_h} p_{g'}$$

$$\frac{\partial s(g)}{\partial Y_{1g'}} = \frac{1\{g = g'\} - s(g)}{\sum_{g=h}(Y_{1h} - Y_{oh})p_h} p_{g'}$$

$$\frac{\partial s(g)}{\partial p_{g'}} = \frac{1\{g = g'\}s(g) - s(g')s(g)}{p_{g'}}$$

were $1\{\}$ is an indicator function equal to 1 when the expression inside the square brackets is true and zero otherwise. These three expressions can be used to generate the entire gradient of $s(g)$ for all values of g.

Next we consider the weights for the subset of consistent choosers, $r(g) \equiv p(g_i = g | c_i = 1)$. Proceeding similarly to before, we can write $r(g)$ as

$$r(g) = \frac{Y_{0g} + 1 - Y_{1g}}{\Sigma_{g=h}(Y_{0h} + 1 - Y_{1h})p_h} p(g_i = g)$$

Taking derivatives and simplifying, we obtain the following for any g and g':

$$\frac{\partial r(g)}{\partial Y_{0g'}} = \frac{1\{g = g'\} - r(g)}{\Sigma_{g=h}(Y_{0h} + 1 - Y_{1h})p_h} p_{g'}$$

[4]To be specific, we know that the off-diagonal elements in the first $2 * J$ rows and the first $2 * J$ columns, which govern the covariance of the various $\bar{Y}_{jg}$ estimates, are zero because the estimation sample for every $\bar{Y}_{jg}$ is distinct. We know that the entries of Σ governing the covariance of $\bar{Y}_{jg}$ and $p(g_i = g)$, for some j and g, are zero because of the unconfoundedness assumption.

$$\frac{\partial r(g)}{\partial Y_{1g'}} = -\frac{1\{g = g'\} - r(g)}{\Sigma_{g=h}(Y_{0h} + 1 - Y_{1h})p_h} p_{g'}$$
$$\frac{\partial r(g)}{\partial p_{g'}} = \frac{1\{g = g'\}r(g) - r(g)r(g')}{p_{g'}}$$

This information alone would be sufficient to construct estimators for the variance of the statistics in Proposition 3 via the delta method. We will consider the statistics in Proposition 3 and 5 together for simplicity; it is always trivial to discard the estimators that are not needed. First we consider the estimators for the preferences of various subgroups. First, the characteristic of interest in the inconsistent subgroup

$$Y_n \equiv E[y_i^* | c_i = 0] = \Sigma_g s(g) Y_{cg}$$

Taking derivatives of this – which may be done more easily using several expressions derived above – we obtain the following for any g:

$$\frac{\partial Y_n}{\partial Y_{0g}} = s(g)\frac{1 - Y_{cg}}{C_g} + \frac{p(g_i = g)}{1 - C}(Y_n - Y_{cg}),$$
$$\frac{\partial Y_n}{\partial Y_{1g}} = s(g)\frac{Y_{cg}}{C_g} - \frac{p(g_i = g)}{1 - C}(Y_n - Y_{cg}),$$
$$\frac{\partial Y_n}{\partial p_g} = \frac{s(g)}{p_g(1 - C)}[Y_{cg} - (1 - C)Y_n].$$

where $C_g \equiv Y_{0g} + 1 - Y_{1g} = E[c_i|g]$ similarly to before, and $C = \Sigma_{g=h} C_h p_h$. From these three expressions we construct the gradient of Y_n.

Proceeding similarly for the full population, $Y_{fp} \equiv E[y_i^*] = \sum_g p(g_i = g) Y_{cg}$, we obtain, for any g

$$\frac{\partial Y_{fp}}{\partial Y_{0g}} = p_g \frac{1 - Y_{cg}}{C_g}$$
$$\frac{\partial Y_{fp}}{\partial Y_{1g}} = p_g \frac{Y_{cg}}{C_g}$$
$$\frac{\partial Y_{fp}}{\partial p_g} = Y_{cg}$$

which allows us to construct the gradient of Y_{fp}.

We can also obtain the gradient of Y_c in terms of the primitive parameters of this model. This part is not necessary to obtain a standard error on $\overline{Y}_c$, because we know how to obtain a simpler formula for the asymptotic variance of our estimator of $\overline{Y}_c$ using the result in the previous subsection of this Appendix. Doing it the hard way here yields an identical standard error estimate. The usefulness of the expressions derived here is that these expressions may be used to estimate the (asymptotic) covariance of, say, the estimators for Y_c and Y_n, which is necessary for the statistical test of the null hypothesis of consistency independence against the alternative hypothesis of conditional consistency independence.

To express Y_c in terms of the primitives of the model, let $P(g_i = g|c_i = 1) = q(g) = \frac{C_g}{C} P(g_i = g)$. Taking derivatives of the expression for Y_c yields, for any g,

$$\frac{\partial Y_c}{\partial Y_{0g}} = q(g)\frac{1 - Y_{cg}}{C_g} + \frac{p_g}{C}(Y_{cg} - Y_c)$$

$$\frac{\partial Y_c}{\partial Y_{1g}} = q(g)\frac{Y_{cg}}{C_g} - \frac{p_g}{C}(Y_{cg} - Y_c)$$

$$\frac{\partial Y_c}{\partial p_g} = \frac{q(g)}{p_g C}(Y_{cg} - CY_c)$$

Using all of the above expressions, we can generate a gradient of each parameter of the matching-on-observables models. Putting all these expressions together, we construct a gradient matrix of the form:

$$G(\theta) = (\nabla s_{g_1}, ..., \nabla s_{g_J}, \nabla r_{g_1}, ..., \nabla r_{gJ}, \nabla Y_n, \nabla Y_{fp}, \nabla Y_c)$$

To be clear, each of the columns of $G(\theta)$ is the gradient of a particular (nonlinear) function of the primitive parameters of the model.

We can now estimate the full variance-covariance matrix of all the parameters of interest $(s_{g_1}, ..., s_{g_J}, r_{g_1}, ..., r_{g_J}, Y_n, Y_{fp}, Y_c)$ by

$$\hat{V}(\hat{\theta}) = \frac{1}{n} G' \Sigma G$$

where G and Σ are evaluated at $\hat{\theta}$. The square root of the diagonals of the matrix $\hat{V}(\hat{\theta})$ will be asymptotically correct standard errors for the parameter estimates themselves. The off-diagonal elements are the estimated covariance of different estimates, which are useful for tests of hypotheses involving more than one parameter of the model, such as tests of consistency independence in this framework.