Online appendix for

Bias in patient satisfaction surveys:

a threat to measuring health care quality

by Felipe Dunsch, David K. Evans, Mario Macis, and Qiao Wang

This appendix includes more details on the data and estimation for the article "Bias in patient satisfaction surveys: a threat to measuring health care quality."

1. Data

As discussed in the paper, patients were randomly assigned to one of three experimental conditions: all positively framed statements, all negatively framed statements, or a random mix of the two. Enumerators visited clinics without providing advance notice, and they invited all patients who visited the clinic to participate. The random assignment of individual patients to treatments was generated by software ("SurveyCTO") on the tablets at the time of interview. The enumerators did not know in advance which set of statements would be presented, the surveys were anonymous, and the interviews were conducted with spatial separation from the PHCs to ensure confidentiality.

For the negatively framed statements, we avoided statements with the word "not", as deciding whether you disagree with the statement "You did not have enough privacy during your visit" can be confusing to respondents due to the double negative (Lietz 2010). As such, in that case, we framed the statement as "You had too little privacy during your visit" in the negatively framed statements. All questions were asked in two stages. In the first stage, the respondent had to decide whether to "agree", "neither agree nor disagree", or "disagree" with the presented statement. Then, in the second stage, the respondent decided – conditional on having chosen to agree or disagree – whether to agree or disagree strongly or not (see Figure 1). For the analysis, we reversed the sign on the negatively framed questions, so that we are comparing the people who agreed with positively framed statements to people who disagreed with negatively framed statements.

Appendix Figure 1: Experiment decision structure

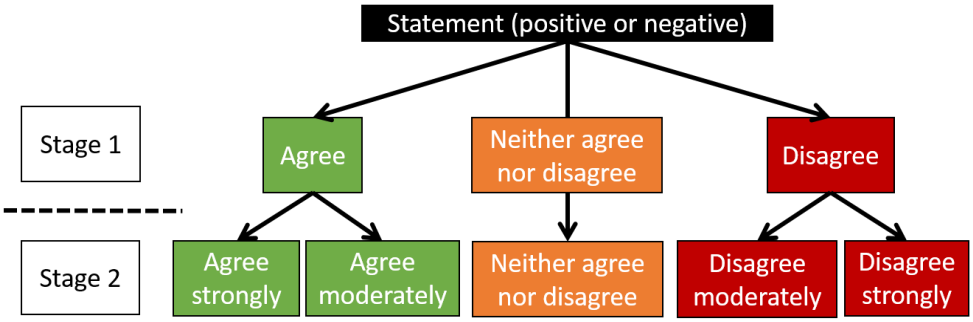


Table 1 shows the distribution of participants across treatment groups, by state and overall. In total, 42 percent of patients received the positively framed questions, 42 percent received the negatively framed questions, and 16 percent answered the random mix.[1]

Appendix Table 1: Distribution of participants across treatment groups, by state and overall

State	N.	Positive framing (%)	Negative framing (%)	Positive-Negative Mixed Framing (%)
Anambra	346	43%	44%	14%
Bauchi	456	40%	42%	18%
Cross River	265	43%	38%	19%
Ekiti	325	44%	43%	14%
Kebbi	444	45%	39%	16%
Niger	386	38%	47%	15%
Total	**2,222**	**42%**	**42%**	**16%**

In Table 2, we present average patients' characteristics, overall and by treatment condition. The average age of patients was 30.3 years. 72% of the patients interviewed were between 19 and 34 years old, 19% were between 35 and 54, 5% were 55 or older, and 3% were 18 or younger. Only 39% of the patients had at least some secondary school education, 83% report being self-employed, 10% were unemployed, and 90% were married. 72% of the patients had never been to a private health care facility. The random allocation of treatment conditions had the desired effect of achieving balance across all of these characteristics.

[1] The third treatment condition, a mix of positively- and negatively-framed statements, was used only during the first three rounds of data collection (of eight total); this explains the fact that they account for a smaller share of the observations.

Appendix Table 2: Patient characteristics, overall and by treatment group

	Total		Positive framing		Negative framing		Positive-Negative Mixed Framing	
	N.	*mean*	*n.*	*mean*	*n.*	*mean*	*n.*	*mean*
Age	2,211	30.3	923	30.5	938	29.9	350	30.5
Age group:								
<=18 years	72	3%	27	3%	34	4%	11	3%
19-34 years	1600	72%	668	72%	685	73%	247	71%
35-54 years	424	19%	173	19%	177	19%	74	21%
>=55 years	115	5%	55	6%	42	4%	18	5%
Gender								
% female	1,859	84%	772	83%	802	85%	285	81%
Employment								
Employed	150	7%	72	8%	56	6%	22	6%
Self-employed	1,840	83%	749	81%	791	84%	300	85%
Unemployed	230	10%	108	12%	92	10%	30	9%
Education Level								
Low	1,365	61%	577	62%	569	61%	219	62%
High	855	39%	352	38%	370	39%	133	38%
Marital Status								
Married	1,991	90%	831	89%	842	90%	318	90%
Single	182	8%	80	9%	78	8%	24	7%
Widowed	42	2%	18	2%	16	2%	8	2%
Divorced	5	0%	2	0%	2	0%	1	0%
Ever been to a private health care facility	611	28%	259	28%	242	26%	110	31%

Notes: Low education = primary school or less (no completed education, adult literacy education, arabic, vocational, other); High education refers to secondary school and higher, including college and higher (university, master's degree, Msc/MA, Ordinary National Diploma, Higher National Diploma, Nigeria Certificate in Education.

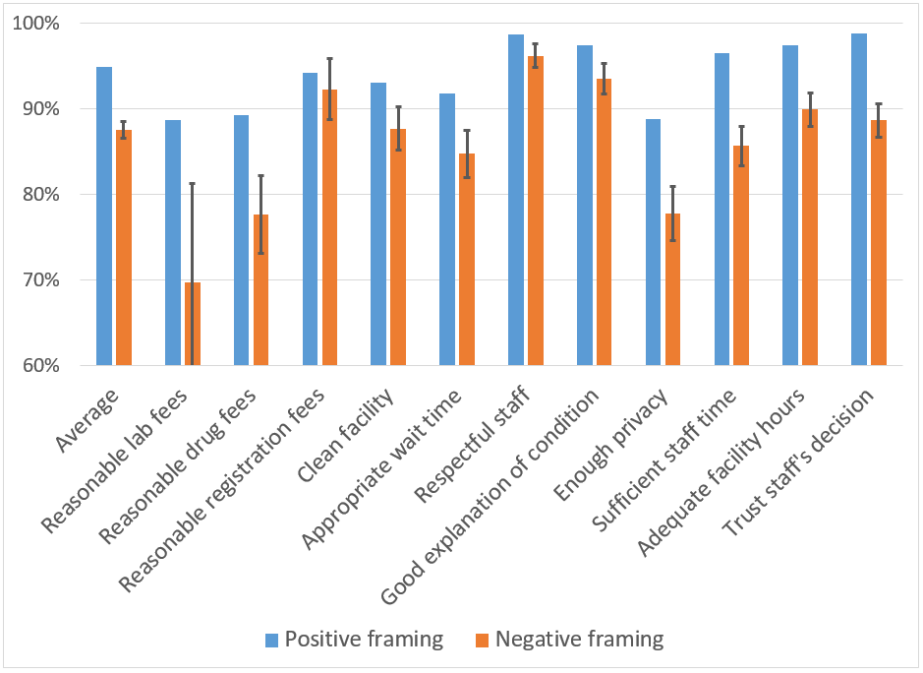
2. Analysis

We estimate three linear probability models. We have estimated ordinal logit models with similar results. Here, we use linear probability models both because it is one of the most common methods of estimation with patient satisfaction survey analysis (Evans and Welander Tärneberg 2017) and for ease of interpretation (Angrist and Pischke 2008).

(1) $favorable_{ik} = \beta_0 + \beta_1 neg_k + \varepsilon_{ik}$

where $favorable_{ik}$ takes the value 1 if patient i gave a favorable response to statement k, and 0 otherwise, and neg denotes negatively framed statements. Because we have balance across observed characteristics (gender, education, age, and income), we do not control for them in our main specification, although we do so as a robustness check in section 3. The results of this specification are reported in Table 2 of the main article. Figure 1 in the main article shows the results visually, and Appendix Figure 2 provides confidence intervals around the estimates.

Appendix Figure 2: The Impact of Positive and Negative Framing on Patient Satisfaction



Notes: The bars on the negative framing indicate the 95 percent confidence interval around the "negative" coefficient, relative to the default, which is positive framing.

The second specification captures the full array of treatments.

(2) $favorable_{ik} = \beta_0 + \beta_1 neg\_w\_neg_k + \beta_2 pos\_w\_mix_k + \beta_3 neg\_w\_mix_k + X_i + \varepsilon_{ik}$

In this second specification, we examine whether including a negatively framed statement within a mix of positively and negatively framed statements affects reporting. *neg_w_neg* denotes negatively framed statements in sets of all negative statements, and *pos_w_mix* and *neg_w_mix* denotes positively and negatively framed statements, respectively, in sets of mixed positive and negative statements (the omitted (or reference) category thus consists of positively framed statements in sets of all positive statements). The results of this specification are reported in Table 3 of the main article.

(3) $favorable_{ik} = \beta_0 + \beta_1 neg_k + \beta_2 X_i + \beta_3 neg_k \times X_i + \varepsilon_{ik}$

In the third specification, we examine whether the impact of negative framing differs by patient characteristic, where *X* represents a patient characteristics such as gender, education, or assets. The results of this specification are reported in Appendix Table 3. In all cases we obtain very similar results to our main specification. We see no statistically significant differences of framing by these characteristics, as demonstrated in the coefficients of the interaction terms. That is, the pattern of acquiescence bias that we uncovered seems to affect patients irrespective of their gender, income, or education.

We find the same result – that the positive or negative framing is crucial to patient responses – if we focus on the more detailed "stage 2" patient responses, when they are asked – conditional on agreement with each statement – if they *strongly* agree or disagree (Appendix Table 4). Of the 11 items, 8 are significant for the *neg_w_neg* group and 7 out of 11 in the *neg_w_mix* group. The effects are slightly smaller for the *neg_w_neg* group when compared to the stage 1 results and about the same for the *neg_w_mix* group. In the *neg_w_mix* group, statement 4 (drug fees) is insignificant for stage 2. For the *neg_w_neg* group, statements 2 (cleanliness) and 5 (respect) become insignificant. The largest effect in this group can be observed for the "lab fees" item.

Appendix Table 3: Impact of framing – Interaction with patient characteristics

	(1)	(2)	(3)	(4)	(5)
	Baseline (no control)	*Baseline + Gender interaction terms*	*Baseline + Age group interaction terms*	*Baseline + Education interaction terms*	*Baseline + Wealth Quintile interaction terms*
Depedent Var.	overall effect	overall effect	overall effect	overall effect	overall effect
Independent Var.					
Neg with Neg	-0.0665	-0.0603	-0.0644	-0.052	-0.0548
	[<0.001]	[<0.001]	[<0.001]	[<0.001]	[<0.001]
Neg with Mix	-0.124	-0.147	-0.16	-0.151	-0.119
	[<0.001]	[<0.001]	[<0.001]	[<0.001]	[<0.001]
Pos with Mix	-0.00528	0.0143	-0.0189	-0.00477	-0.000419
	[0.454]	[0.237]	[0.204]	[0.650]	[0.974]
Female		-0.00628			
		[0.382]			
Female * Neg with Neg		-0.00714			
		[0.624]			
Female * Neg with Mix		0.0275			
		[0.335]			
Female * Pos with Mix		-0.0251			
		[0.087]			
Age group (24-44)			0.00384		
			[0.568]		
Age group (>=45)			0.0282		
			[0.007]		
Age group (24-44) * Neg with Neg			-0.00236		
			[0.851]		
Age group (24-44) * Neg with Mix			0.0437		
			[0.152]		
Age group (24-44) * Pos with Mix			0.0196		
			[0.256]		
Age group (>=45) * Neg with Neg			0.000149		
			[0.995]		
Age group (>=45) * Neg with Mix			0.0885		
			[0.019]		
Age group (>=45) * Pos with Mix			0.0229		
			[0.302]		
Education (Low)				-0.00788	
				[0.218]	
Education (Low) * Neg with Neg				-0.0237	
				[0.040]	
Education (Low) * Neg with Mix				0.0431	
				[0.082]	
Education (Low) * Pos with Mix				-0.000812	
				[0.954]	

Quintile (Poorest)					-0.0326
					[0.001]
Quintile (Less poor)					-0.00997
					[0.257]
Quintile (Average)					-0.0164
					[0.077]
Quintile (Less poor)					-0.0032
					[0.731]
Quintile (Poorest) * Neg with Neg					-0.0206
					[0.232]
Quintile (Poorest) * Neg with Mix					0.0364
					[0.273]
Quintile (Poorest) * Pos with Mix					0.0107
					[0.613]
Quintile (Less poor) * Neg with Neg					-0.0186
					[0.263]
Quintile (Less poor) * Neg with Mix					0.0168
					[0.603]
Quintile (Less poor) * Pos with Mix					-0.0241
					[0.258]
Quintile (Average) * Neg with Neg					-0.00425
					[0.810]
Quintile (Average) * Neg with Mix					-0.0235
					[0.523]
Quintile (Average) * Pos with Mix					-0.0092
					[0.655]
Quintile (Less poor) * Neg with Neg					-0.0224
					[0.212]
Quintile (Less poor) * Neg with Mix					-0.0587
					[0.126]
Quintile (Less poor) * Pos with Mix					-0.00668
					[0.739]
Pos with Pos (Control Mean)	0.949	0.949	0.949	0.949	0.949
Obs. (N)	19586	19586	19222	19568	19361

Appendix Table 4: Impact of framing on patient satisfaction – Second stage ("strongly agree")

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Lab fees	Drugs fees	Registration fees	Clean	Wait time	Respect	Explain	Privacy	Staff time	Open hours	Trust	Overall
Neg with Neg	-0.178	-0.0859	-0.0306	-0.0264	-0.0572	-0.00178	-0.0263	-0.0845	-0.114	-0.0452	-0.0867	-0.0571
	[0.028]	[0.005]	[0.175]	[0.175]	[0.002]	[0.880]	[0.046]	[<0.001]	[<0.001]	[0.002]	[<0.001]	[<0.001]
Neg with Mix	-0.0867	-0.103	-0.191	0.0347	-0.124	-0.0209	-0.0896	-0.196	-0.130	-0.117	-0.223	-0.112
	[0.463]	[0.068]	[0.003]	[0.296]	[0.001]	[0.371]	[0.003]	[<0.001]	[<0.001]	[<0.001]	[<0.001]	[<0.001]
Pos with Mix	-0.0802	-0.0335	-0.181	0.0622	-0.0255	0.00152	0.00767	0.0414	0.0155	-0.0223	0.0188	0.00479
	[0.571]	[0.550]	[0.019]	[0.034]	[0.433]	[0.940]	[0.712]	[0.170]	[0.526]	[0.392]	[0.149]	[0.695]
Pos with Pos (Control Mean)	0.723	0.770	0.917	0.785	0.830	0.932	0.925	0.802	0.892	0.911	0.958	0.874
Obs. (N)	178	1004	784	2219	2219	2213	2204	2209	2219	2144	2193	19586
N. of missing response	2	7	37	3	3	9	18	13	3	78	29	202
Obs with perfect response rate	180	1011	821	2222	2222	2222	2222	2222	2222	2222	2222	19788

Dependent variable = 1 if the patient responded strongly favorably in stage 2 (i.e., "strongly agree" on positively framed questions or "strongly disagree" on negatively framed questions), 0 otherwise.

References

Angrist JD, Pischke J. *Mostly harmless econometrics: An empiricist's companion*. Princeton university press 2008.

Evans D, Welander Tärneberg A. Health Care Quality and Information Failure: Evidence from Nigeria. *Health Economics* 2017.

Lietz P. Research into questionnaire design – A summary of the literature. *International Journal of Market Research* 2010; **52**: 249-272.