Supplementary Material A: Certainty

In this section we conduct a similar analysis to the one reported in the article but this time taking into account the certainty levels regarding the influenza vaccination as reported by the participants themselves (on a scale of 1 to 5). As discussed in the introduction, the comparative reasoning generated by the decoy option is more likely to show up when preferences are relatively weak. Table A.1 starts off by summarizing the numbers (and percentages) of individuals in each of the certainty levels regarding their attitude towards influenza vaccination (along the scale of 1 to 5). A non-negligible percentage of individuals rank themselves in each level.

Table A.1: Number and percentage of participants by certainty level.

Likert scale	Distribution
1 - Not certain at all	321 (9.8%)
2 - Low level of certainty	204 (6.2%)
3 - Medium level of certainty	575 (17.6%)
4 - High level of certain	1,053 (32.2%)
5 - Very high level of certainty	1,118 (34.2%)
Total	3,271 (100%)

In passing, we note that the answer to the question regarding the level of certainty may have been potentially influenced by the treatment. A decoy option, for example, may enhance a decision maker's confidence regarding his choice of the dominating option and, looking back, he may feel more certain about his decision compared to a similar choice with no decoy. To neutralize this potential bias we did not ask participants regarding their earlier decision. Instead, and without referring to the specific content of the message they received, we asked them about their certainty level regarding their willingness to receive the flu shot. While we cannot rule out that participants' responses to this question were influenced by the invitation message and by marking their intentions, the distributions of answers to this question were very similar across treatments. In Table A.2, we run an ordinal regression, in which the dependent variable is the degree of certainty (1-5) and the independent variables are the treatments. As is evident from the table, none of the treatment coefficients are significant. The results of this regression along with the random assignment of participants into treatments do not allow to reject the null hypothesis that reported certainty levels were independent of treatments (or similarly affected by them).

Table A.2: Ordinal regression examining how certainty levels are affected by treatments.

	Dependent Variable			
	Certainty levels			
Variable	Value	(Std. Error)	t-value	p-value
Recommendation	0.01	(0.10)	0.09	0.93
Stock	0.07	(0.10)	0.69	0.49
Cost	-0.01	(0.10)	-0.15	0.88
Benefit	0.06	(0.10)	0.65	0.52
1→2	-2.19	(0.09)	-25.66	0.00
2→3	-1.63	(0.08)	-20.84	0.00
3→4	-0.65	(0.07)	-9.08	0.00
4→5	0.68	(0.07)	9.42	0.00

In Table A.3 we report the percentages of reported intentions in the treatments and control for each certainty level. According to the table, participants who are mostly affected by the treatments are those with certainty levels 2 and 3 for whom we observe increases of roughly 20 percentage points compared to the control. Fisher's exact test for level 2 and Pearson's Chi Square test for level 3 allows us, after adjusting for multiple comparisons, to reject the null hypothesis that intentions in these groups are not affected by the treatments (level 2: adjusted p-value=0.022, level 3: $\chi^2 = 15.72$, df=4, adjusted p-value=0.022).[1] For the higher certainty levels (4 and 5) there are no significant differences (level 4: $\chi^2 = 5.673$, df=4, adjusted p-value=0.365, level 5: $\chi^2 = 5.986$, df=4, adjusted p-value=0.365). Interestingly, among those with the weakest preferences (certainty level 1), the percentage of those intending to take the shot is very low across treatments and there are no significant differences (adjusted p-value=0.537 according to Fisher's exact test). It is possible that being as uncertain as they are, the tendency to remain passive is very strong and may override any of our treatment effects. In support of this view, note that among this subgroup the percentage of those who did not receive even one vaccination in the past five years is 83% (in the other certainty levels, for comparison, this percentage lies between 34% and 67%).

[1] We use Fisher's exact test for those in the second certainty level due to the small numbers who intend to get the shot in some treatments. We do the same for those who reported the lowest certainty level.

Table A.3: Percentages of reported intentions to receive the vaccine in the treatments and control by certainty levels (1-5). In brackets is the actual number who intend to receive the shot divided by the total number of participants in the subgroup.

Certainty level	Control	Treatment			
		Recommendation	Stock	Cost	Benefit
1	2% (1/63)	7% (4/61)	6% (4/65)	8% (5/64)	3% (2/68)
2	0% (0/37)	19% (9/47)	21% (7/33)	22% (11/50)	32% (12/37)
3	37% (44/119)	56% (63/112)	58% (69/118)	56% (66/117)	58% (63/109)
4	67% (151/227)	61% (125/205)	71% (140/197)	66% (138/209)	70% (150/215)
5	64% (136/213)	64% (140/219)	63% (146/231)	67% (149/223)	72% (168/232)
Overall	50% (332/659)	53% (341/644)	57% (366/644)	56% (369/663)	60% (395/661)

Table A.4 reports the results of two logistic regressions that take the certainty level variable into account. In the left column of the table the specification takes into account the treatments, the certainty variable and the interaction between the two (we subtracted 1 from the certainty level measure so that the lowest value of this variable in the regression is 0 and the highest is 4). The specification in the right column controls for demographic variables. The table shows that all treatments have positive and highly significant effects on intentions (odds ratio ranging from 3.411 to 3.99). The interaction coefficients are negative meaning that the treatment effects weaken as the certainty level increases.

Including the certainty variable in a regression model alongside the treatments has two main disadvantages. The first, mentioned earlier, is that this variable may have been influenced to some extent by the message received by the participant and the intention he/she marked. The second disadvantage is that this variable is not observable by HMOs. Thus, while its interaction with the treatments is interesting from a theoretical and psychological point of view, it does not assist policy makers in targeting or understanding the groups in the population that should be approached with this intervention. For this reason, in the main text we make use of the number of vaccinations in the past five years which is observable by HMOs and therefore carries important policy implications. As reported in the main text, the correlation between the certainty level variable and the number of vaccinations in the past five years equals 0.39.

Table A.4: Logistic regression models with certainty levels.

	Dependent Variable:	
	Vaccination Intention	
	(1)	(2)
Recommendation	0.902^{***}	1.227^{***}
	(0.349)	(0.422)
Stock	1.109^{***}	1.303^{***}
	(0.346)	(0.417)
Cost	0.949^{***}	1.386^{***}
	(0.345)	(0.413)
Benefit	0.836^{**}	1.234^{***}
	(0.351)	(0.417)
Certainty	0.863^{***}	0.639^{***}
	(0.087)	(0.109)
Recommendation × Certainty	-0.270^{**}	-0.382^{***}
	(0.114)	(0.139)
Stock × Certainty	-0.289^{**}	-0.346^{**}
	(0.113)	(0.137)
Cost × Certainty	-0.239^{**}	-0.342^{**}
	(0.113)	(0.137)
Benefit × Certainty	-0.139	-0.280^{**}
	(0.115)	(0.138)
Last Year Vaccination		3.662^{***}
		(0.163)
Gender (Male)		0.567^{***}
		(0.099)
Age		-0.014^{***}
		(0.004)
Income (Above Average)		0.221
		(0.248)
Income (Slightly Above)		0.199
		(0.199)
Income (Average)		0.525^{***}
		(0.195)
Income (Slightly Under)		0.077
		(0.191)
Income (Under Average)		0.148
		(0.184)
Education (Undergraduate)		0.009
		(0.154)
Education (Vocational)		0.008
		(0.167)
Education (High School)		0.022
		(0.164)
Education (Elementary)		-0.013
		(0.450)
Constant	-2.414^{***}	-2.652^{***}
	(0.270)	(0.432)
Observations	3,271	3,055
Log Likelihood	$-1,999.611$	$-1,358.801$
Akaike Inf. Crit.	4,019.222	2,761.602
Note:	*p<0.1; **p<0.05; ***p<0.01	