

Online Appendix for *The Value of Dignity Appeals: Evidence from a Social Media Experiment*

A1. Ethical Principles Statement

We follow common research and ethical standards, including the "Principles and Guidance for Human Subjects Research" outlined by the American Political Science Association.

We conducted our experiment on the Facebook platform, and respondents were identified through Facebook's advertising algorithm. Our sample is limited to Facebook users, but while our respondents are not representative of the entire United States, we tried to make it adequately diverse by conducting the experiment in the 48 contiguous states.

Given the platform, we could not provide informed consent to each respondent for our specific intervention, but in the Facebook terms and conditions, they let users know that features of their feed can be part of experiments. Our intervention was benign, mimicking the ads users see from organizations in their newsfeeds. Users can skip over or engage with the ad, and they are provided no monetary incentive to engage with the ad. Thus, we believe respondents did not feel overly pressured to participate. We also did not deceive respondents in our experiment, and provided only factual information about the different vulnerable groups. We constructed our dataset so respondents were anonymous, with their only identifying information being their home state. We constructed our final dataset so respondents were anonymous, with their only identifying information being their home state. The ads were only displayed during the experiment, so Facebook users can no longer see interactions with the posts.

We do not believe that the ads were sensitive in the contexts where they were asked. We considered that they could have an unintentionally stigmatizing impact. However, we tried to make sure the images in the ad were not inadvertently stigmatizing to any particular group, and we discussed our research project with an advocacy group for the incarcerated who saw the potential value and yet little risk of harm in the research for improving attitudes towards prisoners and other vulnerable groups.

A Canadian-based research organization funded the experiment, and some researchers who designed and implemented the experiment are affiliated with the organization. We do not, however,

believe that there are financial or other conflicts of interest because it did not evaluate a specific program or theory of change of that organization.

The researchers who implemented the experiment gained exempt review from the review board of the researcher's university.

A2. Notes on Facebook platform and deviation from Data Collection Plan

This experiment uses Facebook's A/B test infrastructure to randomly assign users to see one of the five treatment arms. We do this without the Campaign Budget Optimization option. The Campaign Budget Optimization allows Facebook to use a targeting algorithm to push advertisements to those most likely to engage with it. Circumventing this feature can lead to greater internal validity by avoiding issues with confounders that arise from optimization algorithms. Before algorithm updates in 2018, the A/B test functionality was not purely random. Changes to the algorithm have improved this issue (Orazi and Johnston 2020).

Based on piloting, we anticipated a sample size of 72,000–120,000 for the three-arm homelessness study and 48,000 for the two-arm incarceration study and that the advertising campaign would last just one day. In reality, the Facebook algorithm struggled to find an audience for our experiment in some markets, and we modified the ad buy to extend past one day. We note that Ad under-performance is a feature of Facebook ads that is not specific to our study and relates to overall competition in the ad marketplace in Facebook's business model. We chose to extend the time the ad was running based on pre-specified sample size goals developed through power analyses conducted with the `power.prop.test` function in R, and did so without looking at the winning ads from the campaigns. We elected to continue the campaign for up to seven days. **These changes were included in an updated pre-registration of our experiment.**

To describe how we generated individual-level data from aggregate results, we share an example: in a pilot ad, the control condition reached 8069 users, and the dignity condition reached 7989 users; 1058 users in the control condition engaged with the ad (e.g., Liked), while 1262 users in the dignity condition engaged with the ad. In this case, we generated a dataset in which 8069 users were assigned a 0 for control status and 7989 users were assigned a 1 for treatment status. Then, of those 8069 users in the control, 1058 received a 1 for the outcome (engagement), and the remainder received a 0. We repeated the same assignment exercise for each treatment condition. We subsequently merged these

data with our state-level datasets.

A3. Summary Statistics

Table A1 shows the overall engagement rates with the Facebook ads, pooled across the states and treatment conditions in the study. Engagement rates are 7.7%, with just over 7000 Facebook users engaging with the post, out of the nearly 90,000 (90994) users who were exposed to our ad.

Table A1 also shows the proportion of Facebook users in our study who were assigned to see each of the different ads. Users were approximately evenly assigned to each of the treatment conditions. Finally, Table A1 contains summary statistics of our main variables.

Table A1. Summary Statistics

Statistic	N	Mean	St. Dev.	Min	Max
Impressions	90,994	445.625	116.973	194	839
Reach	90,994	402.532	97.444	185	719
Sad	90,994	0.006	0.078	0	1
Like	90,994	0.056	0.229	0	1
Angry	90,994	0.001	0.023	0	1
Haha	90,994	0.002	0.042	0	1
Wow	90,994	0.0003	0.018	0	1
Love	90,994	0.001	0.030	0	1
Positive Reaction	90,994	0.056	0.231	0	1
Any Reaction	90,994	0.065	0.247	0	1
Dignity:Homeless Treatment	90,994	0.218	0.413	0	1
Dignity:Incarcerated Treatment	90,994	0.205	0.403	0	1
Econ:Homeless Treatment	90,994	0.195	0.396	0	1
Info:Homeless Treatment	90,994	0.198	0.398	0	1
Info:Incarcerated Treatment	90,994	0.185	0.388	0	1
Engagement	90,994	0.077	0.267	0	1

A4. Blameworthiness of different groups

To determine the perceived blameworthiness of different stigmatized/vulnerable groups, before the fielding of our main study, we surveyed a sample of 600 American residents on the mTurk platform and asked to what extent they felt individuals from each group were to blame for their circumstances.⁷ Figure A1 shows the distribution of responses to a question conducted on the mTurk

7. Of course, mTurk does not yield a representative sample, but as Berinsky, Huber, and Lenz 2012 notes, this platform yields a more representative sample than in-person convenience samples. We do not believe that imbalances in the mTurk

platform several months before the fielding of the Facebook experiment that asked respondents about the blameworthiness that different groups hold for their situations. We asked about disabled people, people addicted to drugs, people experiencing homelessness, immigrants, the incarcerated, and people experiencing poverty. While responses to the perceived blameworthiness of homeless people was mixed, the perception of blameworthiness for people experiencing incarceration was much more consistent, as most respondents felt as it incarcerated people were to be blamed for their incarceration.

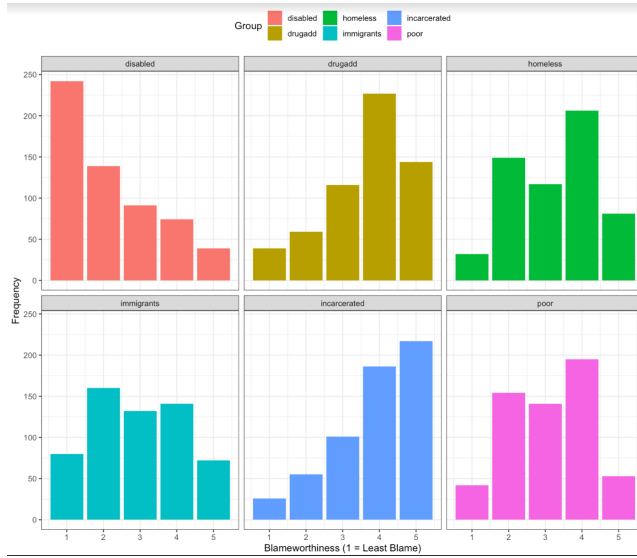


Figure A1. Perceived blameworthiness of different vulnerable or stigmatized social groups. Lower values indicate the least blameworthiness; higher values the most.

A5. Model Specifications

First, we evaluate whether the dignity frame can increase positive engagement with posts highlighting support for vulnerable groups compared to providing information by looking at the effect of the dignity prime compared to the pure information control using an OLS linear regression that controls for the ad domain. For this analysis, we drop respondents from the homeless-economic treatment group.

We use the following estimator:

survey would systematically affect perceptions of blameworthiness.

$$Y_i = \beta_0 + \beta_1 Z_i + \beta_2 D_i + \eta_i$$

Where i is the individual, Z is the treatment indicator, D is the domain (homelessness or incarceration) indicator, and Y is the outcome.

We incorporate potential differences across states by conducting a set of analyses that use a multi-level model where we allow the intercepts to vary at the state level to account for the multi-site, block-randomized approach of our experiment. We use a Random intercept, constant treatment coefficient (RICC) model.

Our model is:

$$Y_i = \alpha_j + \beta_1 Z_{ij} + \beta_2 D_{ij} + \eta_{ij}$$

$$\alpha_j = \alpha + \zeta_j$$

Where i is the individual, j is the state, Z is the treatment indicator, D is the domain indicator, and Y is the outcome.

Second, we assess whether the effectiveness of the dignity frame varies by the vulnerable group featured by interacting the domain with the treatment and dropping those respondents in the homeless-economic treatment group. We use the following estimator:

$$Y_i = \beta_0 + \beta_1 Z_i + \beta_2 D_i + \beta_3 Z_i * D_i + \eta_i$$

Where i is the individual, Z is the treatment indicator, D is the domain indicator, and Y is the outcome.

We again incorporate state-level heterogeneity using the following multi-level model specification:

$$Y_i = \alpha_j + \beta_1 Z_{ij} + \beta_2 D_{ij} + \beta_3 Z_{ij} * D_{ij} + \eta_{ij}$$

$$\alpha_j = \alpha + \zeta_j$$

Where i is the individual, j is the state, Z is the treatment indicator, D is the domain indicator,

and Y is the outcome.

Third, we test whether the dignity frame is also effective compared to another common frame – the economic cost of ignoring a vulnerable group. This analysis only looks at those who received the homelessness domain. We use an ordinary least squares regression to estimate this quantity using the following estimator:

$$Y_i = \beta_0 + \beta_1 Z_i + \eta_i$$

Where i is the individual, Z is the treatment indicator, and Y is the outcome.

We then incorporate state-level heterogeneity using the following specification:

$$Y_i = \alpha_j + \beta_1 Z_{ij} + \eta_{ij}$$

$$\alpha_j = \alpha + \zeta_j$$

Where i is the individual, j is the state, Z is the treatment indicator, and Y is the outcome.

We correct for multiple comparisons by controlling the family-wise error rate using the Benjamini–Hochberg (B–H) procedure as seen in section A7.1.

A6. Table of Results

Table A2. Results from an ordinary least squares regression that compares the effect of the dignity frame to the economic and information ones.

	Positive Reaction			Any Reaction			Any Engagement		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Intercept	0.0605*** (0.0015)	0.0585*** (0.0017)	0.0684*** (0.0018)	0.0732*** (0.0016)	0.0729*** (0.0019)	0.0757*** (0.0019)	0.0848*** (0.0018)	0.0847*** (0.0021)	0.0874*** (0.0020)
Dignity Frame (Base = Information)	0.0061*** (0.0017)	0.0100*** (0.0025)		0.0023 (0.0018)	0.0027 (0.0027)		0.0025 (0.0020)	0.0027 (0.0029)	
Vulnerable Group Incarcerated	–0.0157*** (0.0017)	–0.0115*** (0.0024)		–0.0200*** (0.0018)	–0.0195*** (0.0026)		–0.0196*** (0.0020)	–0.0194*** (0.0028)	
Dignity Frame: Incarcerated		–0.0080* (0.0034)			–0.0009 (0.0036)			–0.0005 (0.0039)	
Information Frame (Base = Dignity)			–0.0100*** (0.0025)			–0.0027 (0.0027)			–0.0027 (0.0029)
Economic Frame (Base = Dignity)			–0.0106*** (0.0025)			–0.0083** (0.0027)			–0.0083** (0.0029)
Domain	Both	Both	Homeless Only	Both	Both	Homeless Only	Both	Both	Homeless Only
R ²	0.0013	0.0014	0.0004	0.0017	0.0017	0.0002	0.0014	0.0014	0.0002
Adj. R ²	0.0013	0.0014	0.0004	0.0016	0.0016	0.0001	0.0014	0.0013	0.0001
Num. obs.	73287	73287	55513	73287	73287	55513	73287	73287	55513
RMSE	0.2299	0.2299	0.2408	0.2457	0.2457	0.2587	0.2658	0.2658	0.2772

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; \cdot $p < 0.1$

A7. Robustness Checks

A7.1 Multiple Hypothesis Test Corrections

We apply the Benjamini-Hochberg (B-H) correction to control for the false discovery rate. We assess three families of hypotheses – hypotheses related to the dignity ad versus the control ad, those related to the effect of the dignity ad versus the control ad varying by group, and those related to the dignity ad versus the economic ad. For each of the three families of hypotheses, we correct for testing hypotheses for three outcomes – any engagement with the ad, any reaction to the ad, and a positive reaction to the ad.

We apply the B-H correction to the specifications presented in the main text of the paper. We test hypotheses for three outcome variables, so we correct for three comparisons. All results remain consistent with the MH corrections via the BH procedure except for the interaction term between the dignity ad and the domain.

A7.2 Multi-level Models

Table A3 presents our robustness analysis that estimates the average treatment effects of the dignity frame across different specifications. We use the same dependent variables (positive reaction, any reaction, and any engagement) as in the remainder of the paper. These results provide a robustness check to the ordinary least squares results by treating our field experiment as a 48-state multi-site study. In these models, we allow for the intercepts to vary at the state-level to account for the multi-site, block-randomized approach of our experiment. We will use a Random intercept, constant treatment coefficient (RICC) model.

A8. Comparison to Economic Appeals

To test whether the dignity frame is only spurring more positive engagement because it provides more information about the problem, we compare its effect to another commonly used frame. Specifically, we compare responses to the dignity frame relative to a frame that focuses on the economic costs of not supporting the vulnerable. We do this only in the case of the homeless and using OLS analysis. The results presented in Figure A2 show that the dignity frame increases positive engagement with the ad compared to both the economic cost frame and the information-only frame. The dignity frame seems particularly powerful compared to the economic cost framing, increasing all forms of

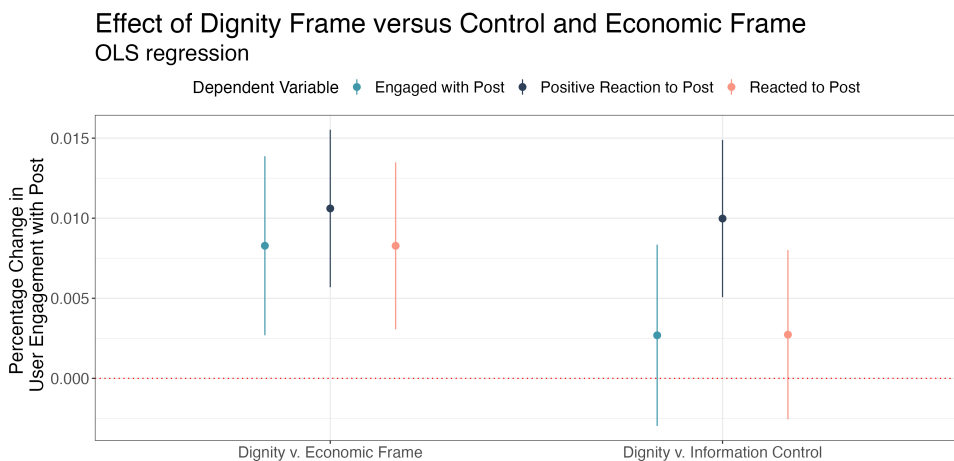
Table A3. Results from a multi-level model regression that compares the effects of the dignity frame to the economic and information frames.

	Dependent variable:								
	Positive Reaction			Any Reaction			Any Engagement		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Information Frame (Base = Dignity)	-0.007*** (0.002)			0.0001 (0.003)			0.0003 (0.003)		
Economic Frame (Base = Dignity)	-0.007*** (0.002)			-0.005* (0.003)			-0.005 (0.003)		
Dignity Frame (Base = Information)		0.005*** (0.002)	0.007*** (0.002)		0.001 (0.002)	-0.0001 (0.003)		0.001 (0.002)	-0.0003 (0.003)
Vulnerable Group Incarcerated		-0.013*** (0.002)	-0.011*** (0.002)		-0.017*** (0.002)	-0.018*** (0.003)		-0.017*** (0.002)	-0.018*** (0.003)
Dignity Frame:Incarcerated			-0.005 (0.003)			0.002 (0.004)			0.003 (0.004)
Constant	0.070*** (0.005)	0.064*** (0.005)	0.063*** (0.005)	0.077*** (0.005)	0.077*** (0.005)	0.078*** (0.005)	0.089*** (0.006)	0.089*** (0.006)	0.090*** (0.006)
Observations	55,513	73,287	73,287	55,513	73,287	73,287	55,513	73,287	73,287
Log Likelihood	625.783	4,313.987	4,310.226	-3,345.258	-558.535	-563.008	-7,204.337	-6,345.473	-6,349.778
Akaike Inf. Crit.	-1,241.567	-8,617.975	-8,608.452	6,700.516	1,127.070	1,138.016	14,418.670	12,700.950	12,711.560
Bayesian Inf. Crit.	-1,196.945	-8,571.964	-8,553.239	6,745.138	1,173.081	1,193.229	14,463.300	12,746.960	12,766.770

Note:

*p<0.1; **p<0.05; ***p<0.01

engagement with the ad. These results help allay concerns that our results are merely an artifact of enhanced messaging, irrespective of content.

**Figure A2.** Results from an ordinary least squares regression that compares the effect of the dignity frame to the information and economic frames on engagement with the post for just the ad featuring a homeless person. Effects also presented in Table A2.

A9. Comment Analysis

Following the study's conclusion, we manually copied and pasted all (475) user comments made on our 5 Facebook ads, recording the state and ad each comment was associated with. We subsequently hired two research assistants to classify the text along several dimensions independently: Blinded to treatment assignment, they classified each comment in terms of whether the tone was respectful/positive (Coded as one if there is any caring, empathetic, or positive tone concerning the category of the person mentioned, including simply recognizing the category) or disrespectful/negative towards the vulnerable group (Coded as one if there is any angry, blaming or negative tone concerning the category of person mentioned, including to relegate the importance of this category to a secondary position), and whether the comment intimated that members of the vulnerable group were blameworthy (Coded as one if the comment highlighted that the members of this group – incarcerated people or homeless people – deserved to be in this position because of actions they took, lack of moral code, etc.).

Inter-coder reliability was reasonably high (Pearson's $R > 0.5$), given the ambiguity of many of the statements. We coded the respective variables as 0.5 when just one coder indicated a positive match and 1.0 when both did. Table A4 displays the summary statistics for coding each of the 475 comments. Nearly all the comments were relevant to the ad (98% of all comments were deemed relevant by the coders). Only one-third of comments contained positive or respectful content, compared to over half (52.8% of comments) that contained negative or disrespectful attitudes. Over a third of comments (37.2%) also invoked themes of blameworthiness when discussing the ad on their newsfeeds.

In turn, we replicated some of our main analyses, now using the coding of comments as the key outcome variables, albeit with a much smaller sample. As we discuss in the main text, the findings from this additional analysis are very consistent with those from the main analysis in the sense that the dignity treatment tends to lead to positive/affirming comments only in the case of the homelessness ads, but has a somewhat negative effect (on average) concerning the ads focused on incarceration.

We present the findings in table A5. As in the main analyses, the models that generated these figures regressed three dependent variables on the experimental treatment, comparing how the dignity frame performed across all vulnerable groups, comparing how the dignity frame performed differentially for the homeless and incarcerated groups, and how the dignity frame performed compared

Table A4. Summary Statistics for Coding of Content in Expressed in Facebook Comments

Statistic	N	Mean	St. Dev.	Min	Max
Relevant	475	0.983	0.111	0.000	1.000
Positive Respect	475	0.335	0.415	0.000	1.000
Negative Disrespect	475	0.528	0.431	0.000	1.000
Blameworthy	475	0.372	0.424	0.000	1.000

to both the economics appeal and the information-only ad for the homeless group. Our outcome variables are: whether the comments mentioned blameworthiness, were negative or disrespectful, or were positive or respectful. We show that across both groups, the dignity frame did not result in statistically different comment content. However, when we compare how dignity performed by group, we can see that the results differ substantially.

Table A5. Results from an ordinary least squares regression that compares the effect of the dignity frame to the information and economic frames on the content of post comments.

	Positive or Respect			Negative or Disrespect			Blameworthy		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Intercept	0.5547*** (0.0369)	0.5000*** (0.0441)	0.5000*** (0.0441)	0.2608*** (0.0378)	0.3082*** (0.0453)	0.2083*** (0.0469)	0.1392*** (0.0382)	0.1918*** (0.0457)	0.1000* (0.0410)
Dignity Frame (Base = Information)	0.0288 (0.0374)	0.1500* (0.0657)	0.1500* (0.0657)	0.0052 (0.0383)	-0.0999 (0.0674)		0.0249 (0.0387)	-0.0918 (0.0681)	
Vulnerable Group Incarcerated	-0.3666*** (0.0400)	-0.2800*** (0.0556)	-0.2800*** (0.0556)	0.4229*** (0.0410)	0.3478*** (0.0570)		0.3716*** (0.0414)	0.2882*** (0.0576)	
Dignity Frame:Incarcerated		-0.1782* (0.0797)	-0.1782* (0.0797)		0.1546 (0.0818)			0.1715* (0.0826)	
Information Frame (Base = Dignity)						0.0999 (0.0634)			0.0918* (0.0554)
Economic Frame (Base = Dignity)						0.1537* (0.0670)			0.0293 (0.0585)
Domain	Both	Both	Homeless Only	Both	Both	Homeless Only	Both	Both	Homeless Only
R ²	0.1685	0.1785	0.1785	0.2066	0.2134	0.0283	0.1667	0.1753	0.0153
Adj. R ²	0.1645	0.1725	0.1725	0.2027	0.2077	0.0180	0.1626	0.1693	0.0048
Num. obs.	417	417	417	417	417	191	417	417	191

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; $p < 0.1$