Online Appendix to *Attention and the Economic Response to Immigration Enforcement*

Greg C. Wright
University of California, Merced

July 8, 2026

This Online Appendix collects the additional tables, figures, and diagnostics referenced in the main paper. It is provided for review and is not part of the typeset article. Sections and exhibits are numbered with an "S" prefix. Acronyms and variable definitions follow the main paper. References to "the main paper" point to *Attention and the Economic Response to Immigration Enforcement*.

S1 Event inventory and summary statistics

Table S1 reports event-level summary statistics by enforcement wave. The compiled single-worksite inventory contains 55 ICE-led events satisfying the event-date, geography, and geocoding screens described in the main paper. Table S3 lists the 51 events in the main analysis sample and the four single-worksite events that lack sufficient catchment POI coverage to fit the per-event distributed-lag DiD. Inclusion in the main sample requires at least 10 high-FB and 10 low-FB CBGs with sufficient POI coverage in the catchment. The four catchment-coverage exclusions are data-availability-driven rather than attention-correlated.

Table S1: Event-level summary statistics, by enforcement wave

	Full	2025–2026	2018–2019
Arrests reported at worksite	70 (128)	46 (89)	215 (218)
Spanish-ICE news spike	0.0298 (0.0242)	0.0337 (0.0235)	0.0054 (0.0100)
English-GDELT news spike	0.0146 (0.0145)	0.0163 (0.0148)	0.0036 (0.0027)
Late-window estimate (log pts)	-0.008 (0.047)	-0.009 (0.047)	-0.003 (0.046)
High-FB CBGs in catchment	118 (167)	130 (175)	43 (69)
Low-FB CBGs in catchment	134 (165)	143 (172)	84 (100)
Events	51	44	7

Notes: Cells report the cross-event mean with the standard deviation in parentheses, for the full 51-event sample and separately by enforcement wave. News spikes are the post-event peak above the pre-event baseline in the share of stories mentioning the term (see the main paper). The late-window estimate is the per-event weeks-4–8 differential in log foot traffic. High-FB and low-FB CBGs are the top- and bottom-decile Latin-American-foreign-born block groups in each event's catchment. Arrests are missing for two events.

Table S2: Sample-flow from full event inventory to estimation subsamples

Definition	Events
Single-worksite ICE-led events satisfying inventory screens	55
− Single-worksite ICE-led but no estimable late-window estimate (insufficient catchment POI coverage)[a]	−4
Main analysis sample (single-worksite ICE-led with estimable late-window estimate)	51
Estimation subsamples:	
Events with reported arrest counts (severity test)	49
Events in 2024–2026 subsample (Google Trends national)	44
Events with DMA-level Google Trends coverage	41
Events with at least one in-state Mediacloud Spanish-language source	31
Events with estimable per-event civic estimate	12

Notes: "Events" is the unit. The 51-event panel is the main analysis sample for the per-event distributed-lag DiD and the headline cross-event horse race. Smaller numbers below reflect availability of specific second-stage covariates rather than re-estimation. The Mediacloud Spanish-language national series and the English-GDELT series are defined for all 51 events. Trends, DMA Trends, and in-state Spanish sources cover smaller subsamples. The civic estimate requires sufficient civic-NAICS POI density in the high-FB CBGs of an event's catchment.

Table S3: Single-worksite ICE-led event inventory

Date	State	Event / worksite	Arrests
Main sample ($n = 51$)			
2018-04-05	TN	Southeastern Provision	97
2018-05-09	IA	MPC Enterprises (Midwest Precast Concrete)	32
2018-06-05	OH	Corso's Flower & Garden Center (Sandusky + Castalia)	114
2018-06-19	OH	Fresh Mark	146
2018-08-28	TX	Load Trail LLC	159
2019-04-03	TX	CVE Technology Group	280
2019-08-07	MS	Mississippi poultry operation (7 plants, 5 companies)	680
2025-01-23	NJ	Newark	
2025-02-03	PA	Philadelphia	7
2025-02-12	TX	Los Fresnos	2
2025-02-18	NY	Tupper Lake Pine Mill	9
2025-02-26	MS	3 J Underground LLC	7
2025-02-26	MS	Gulf Coast Prestress Partners Ltd	18
2025-02-26	NJ	North Bergen	16
2025-02-27	PA	Jumbo Meat Market	4
2025-03-27	CA	Powder & Protective Coatings	5
2025-04-02	WA	Mt. Baker Roofing	37
2025-04-22	VT	Berkshire	8
2025-04-27	CO	Colorado Springs	100
2025-05-02	NY	Lynn-Ette & Sons Farms	14
2025-05-13	FL	Wildwood	33
2025-05-20	WA	Eagle Beverage	17
2025-05-21	AL		
2025-05-27	LA	Mirabeau Water Garden stormwater project	20
2025-05-29	FL	Perla	100
2025-05-30	CA	Los Angeles	36
2025-05-30	CA	Buono Forchetta	3
2025-06-04	TX	Brownsville	25
2025-06-05	PA	Wyoming Valley Pallets Inc.	4
2025-06-06	CA	Los Angeles	44

continued on next page

(Table S3 continued)

Date	State	Event / worksite	Arrests
2025-06-10	CA	Oxnard	35
2025-06-10	NE	Glenn Valley Foods	76
2025-06-11	PA	Five10 Flats restoration jobsite	17
2025-06-17	LA	Delta Downs Racetrack	84
2025-07-02	TX	Groomer's Seafood	12
2025-07-08	NJ	Alba Wine and Spirits	20
2025-07-10	CA	Glass House Farms	361
2025-07-10	TX	Taco Ole restaurant	18
2025-07-16	PA	Super Gigante	14
2025-08-20	NJ	Edison	29
2025-08-23	CT	Optimo Car Wash	7
2025-09-02	MO	Golden Apple Buffet	12
2025-09-04	GA	HL-GA Battery Company	475
2025-09-04	NY	Nutrition Bar Confectioners	57
2025-10-09	OH	Pancho's Tacos	5
2025-10-15	CT	Hamden	7
2025-10-19	ID	La Catedral Arena	105
2025-10-30	OR	Woodburn	35
2025-10-30	LA	Barrois Welding Services shipyard	25
2025-11-04	MA	Allston Car Wash	9
2025-11-14	AL	Nori Thai and Sushi	18
Excluded: insufficient catchment POI coverage ($n = 4$)			
2018-08-08	NE	O'Neill NE agricultural operation (3 sites)	
2025-06-04	NM	Outlook Dairy Farms	11
2026-01-15	MN	D.R. Horton	
2026-01-15	TX	El Paso	38

Table S4: Source-type audit for the single-worksite inventory

Primary cited source type	Inventory	Main	Med./low	Low spike
Prior compilation or tracker	23	22	11	14
State or local news	15	15	9	7
Agency or public record	10	9	8	4
National news	3	3	0	1
Business or address verification	2	2	2	1
Trade/compliance tracker	2	0	2	2
Total	55	51	32	29

Notes: The table classifies the primary citation recorded in the event inventory. "Inventory events" are the 55 single-worksite ICE-led events satisfying the screens described in the main paper. "Main sample" counts the 51 events with sufficient catchment POI coverage to estimate the event-level distributed-lag DiD. "Low news spike" denotes events with an English-GDELT post-event spike at or below 0.01. The table is an audit of the public-source inventory rather than a claim that the inventory is an administrative universe of all obscure raids.

S2 Additional result diagnostics

S2.1 News measures, inference, and local attention

The cross-language and cross-source validity of the headline measure can be quantified directly. The Mediacloud Spanish-ICE event spike closely tracks both the English-GDELT ICE-raid spike across the full event panel and the US national Google Trends ICE-raid spike where Trends data are available. The English-GDELT spike also tracks Spanish-language search spikes for "redada" and "deportación" in the 2024–2026 subsample. Event-level ICE news cycles propagate consistently through English news supply, Spanish news supply, and English and Spanish search demand on essentially the same schedule.

The choice of search term within the Mediacloud Spanish corpus matters. The headline uses the noun "ICE," which has by far the cleanest event-specific signal. Alternative phrasings such as "agentes de ICE" and "redada migratoria" give the same negative sign but weaker statistical power. "Deportación" is essentially null, consistent with the term tracking the Spanish-language political-policy news cycle rather than event-specific raid coverage.

Because the headline selects one measure from this family, a correction for multiple testing is applied across the five news-volume measures. Table S5 reports Holm and Romano-Wolf family-wise adjusted p-values. The Spanish-ICE coefficient remains significant after both corrections (Holm $p = 0.003$, Romano-Wolf $p = 0.007$), as does the English-GDELT measure, while the weaker Spanish query variants do not. The Romano-Wolf adjustment, which accounts for the strong positive correlation across the measures, is smaller than the conservative Holm bound.

Table S5: Multiple-testing correction across attention measures

Attention measure	Coef	t	p	Holm p	RW p
Spanish "ICE" (headline)	-0.0184	-3.65	$<$0.001	0.003	0.007
English-GDELT national	-0.0154	-3.06	0.004	0.014	0.024
Spanish "agentes de ICE"	-0.0095	-1.74	0.088	0.177	0.212
Spanish "redada migratoria"	-0.0104	-1.97	0.054	0.162	0.177
Spanish "deportación"	+0.0019	+0.35	0.726	0.726	0.735

Notes: Each row is a separate precision-weighted regression of the per-event late-window estimate on the standardized attention measure, with Knapp–Hartung inference and inverse-variance weights (same estimator as the main paper's precision-weighted regression); $N = 51$. The family is the set of news-volume measures discussed in this appendix. "Holm p" is the Holm step-down family-wise adjusted p-value, which controls the family-wise error rate under arbitrary dependence. "RW p" is the Romano–Wolf step-down adjusted p-value from a recentered pairs bootstrap (2,000 resamples of events, fixed seed), which accounts for the strong positive correlation across measures and is therefore less conservative than Holm. The headline Spanish "ICE" coefficient remains significant after both corrections.

The predictive content is concentrated in sustained rather than transient coverage. Table S6 regresses the per-event late-window estimate on each of several coverage-timing measures in turn. The weekly peak, an eight-week area-under-the-curve measure, and national search each predict the response, while a single-day peak and counts of days on which coverage crosses a threshold are weaker, both in statistical significance and in the share of cross-event variance explained.

Table S6: Attention shape: which coverage-timing measure predicts the response

Coverage-timing measure	N	Coef. per SD	R^2
Panel A: Sustained / aggregate attention			
National Google Trends spike	44	-0.022^{***}	0.22
		(0.007)	
Spanish-ICE weekly peak	51	-0.022^{***}	0.23
		(0.006)	
English-GDELT weekly peak	51	-0.019^{***}	0.17
		(0.006)	
English-GDELT eight-week AUC	51	-0.019^{***}	0.17
		(0.006)	
Panel B: Transient / threshold measures			
English-GDELT one-day peak	51	-0.009	0.03
		(0.007)	
Days above baseline + 1 SD	51	-0.011^{*}	0.06
		(0.006)	
Days above 2× baseline	51	-0.009	0.04
		(0.006)	

Notes: Each row is a separate cross-event OLS regression of the per-event late-window foot-traffic estimate on one standardized coverage-timing measure, so the coefficient is per standard deviation of that measure and a more negative value means a larger predicted foot-traffic decline. R^2 is the share of cross-event variance the single measure explains. Measures are built from the English-language GDELT ICE news-volume series unless noted: the weekly peak, the post-event eight-week area under the coverage curve (AUC), the single-day peak, and counts of days on which coverage exceeds its pre-event baseline by more than one standard deviation or twice the baseline. The national Google Trends spike is on the 44-event 2024–2026 subsample; the other measures use the full 51-event sample. The English-GDELT weekly peak is the same series as the headline English-GDELT news spike used elsewhere in the paper, labeled here by its timing operationalization; it is shown once. Sustained and aggregate measures (Panel A) predict the response, while single-day and threshold-crossing measures (Panel B) are weaker. Stars: $^{*}\,p < 0.10$, $^{**}\,p < 0.05$, $^{***}\,p < 0.01$.

The generated-regressor concern does not drive the result. In inverse-variance-weighted WLS, where each event is weighted by the inverse of its squared standard error from the distributed-lag DiD, the Spanish-ICE coefficient is smaller than the OLS estimate but remains statistically significant. Adding arrests leaves the result intact. An event-level bootstrap on the OLS specification also excludes zero.

The post-event attention measure also does not appear to be standing in for anticipatory media attention. For each news-volume series an immediate pre-event spike is constructed, defined as the peak over $\tau \in [-4, -1]$ minus the same $\tau \in [-8, -5]$ baseline used in the headline post-event spike. The immediate pre-event Spanish-ICE spike is uncorrelated with both the late-window estimate and the average pre-event foot-traffic coefficient. Adding it to the headline horse race leaves the post-event Spanish-ICE coefficient unchanged, and the English-GDELT measure gives the same pattern.

An event-local Spanish-news series is also constructed by restricting the Mediacloud query to outlets tagged as published in the event's home state. The local-Spanish spike moves in the same direction as the national headline, but its predictive power is weaker and it adds little once national Spanish-language attention is included. This mirrors the same-platform national-vs.-DMA contrast within Google Trends. The predictive content of local-language local-news attention lies primarily in the part that tracks the national signal.

Table S7: Early-window robustness

	Weeks 0–2	Weeks 1–3	Weeks 4–8
Panel A. Event-estimate distribution			
Mean estimate	-0.001	-0.004	-0.008
Median estimate	0.000	-0.003	-0.003
Share negative	0.51	0.55	0.57
Panel B. Cross-event effect of news coverage			
Pearson r, Spanish-ICE spike	-0.263	-0.375	-0.477
OLS coef., Spanish-ICE spike	-0.323^{*}	-0.524^{***}	-0.919^{***}
	(0.174)	(0.191)	(0.256)
OLS coef., Spanish-ICE spike, with arrests	-0.388^{**}	-0.592^{***}	-0.999^{***}
	(0.181)	(0.202)	(0.265)
Panel C. Corroborating English-news measure			
Pearson r, English-GDELT spike	-0.227	-0.333	-0.416
OLS coef., English-GDELT spike	-0.466	-0.774^{**}	-1.338^{***}
	(0.302)	(0.333)	(0.438)
N	51	51	51

Notes. Each column summarizes the same per-event distributed-lag DiD path over a different post-event window. Week 0 is the calendar week containing the enforcement event and can include pre-event days. The weeks 1–3 window excludes the event week. The weeks 4–8 column is the headline persistent-response window used in the main horse race. OLS rows report single-regressor HC1 standard errors in parentheses, except for the "with arrests" row, which also includes $\log(1 + \text{Arrests}_e)$ as a control. The arrests-control regressions use 49 events. Stars denote $^{*}\, p < 0.10$, $^{**}\, p < 0.05$, $^{***}\, p < 0.01$.

Table S8 reports the between-event standard deviation of true effects at each event time, with Q-profile confidence intervals, alongside the per-event-time effect of news coverage. The corresponding figure appears in the paper's appendix.

Table S8: Between-event dispersion of the response over event time

	$\hat{\tau}$ (REML)	Q-profile 95% CI	I^2	Effect of news coverage
Panel A: Pre-period average ($\tau \in [-8, -2]$)				
Mean over pre-event weeks	0.0162	—	53.5%	-0.0021
Panel B: Post-period average ($\tau \in [0, 8]$)				
Mean over post-event weeks	0.0243	—	63.2%	-0.0139
Panel C: Selected event times				
$\tau = -8$	0.0197	[0.0110, 0.0299]	57.9%	0.0005 (0.0047)
$\tau = -4$	0.0179	[0.0108, 0.0279]	57.0%	-0.0008 (0.0041)
$\tau = -2$	0.0118	[0.0059, 0.0175]	52.5%	-0.0022 (0.0029)
$\tau = +0$	0.0131	[0.0081, 0.0235]	52.4%	-0.0028 (0.0033)
$\tau = +2$	0.0227	[0.0158, 0.0355]	67.6%	-0.0103 (0.0047)
$\tau = +4$	0.0249	[0.0185, 0.0416]	63.3%	-0.0183 (0.0047)
$\tau = +6$	0.0294	[0.0220, 0.0476]	63.0%	-0.0192 (0.0054)
$\tau = +8$	0.0346	[0.0272, 0.0565]	71.2%	-0.0172 (0.0067)

Notes: Each row summarizes a random-effects meta-analysis of the 51 per-event event-study coefficients γ_τ^e at the given event time, using each event's CBG-clustered standard error. $\hat{\tau}$ is the estimated between-event standard deviation of true effects net of sampling noise (REML); the confidence interval is the Q-profile interval. The effect of news coverage is a precision-weighted regression of γ_τ^e on the standardized Spanish-language coverage spike with Knapp–Hartung standard errors (in parentheses). The reference week $\tau = -1$ is zero for every event by construction and is omitted. Because γ_τ^e is a deviation from the reference week, dispersion mechanically accumulates with distance from $\tau = -1$ in both directions; the relevant comparison is at matched distance, where post-event dispersion exceeds pre-event at 7 of 7 distances (mean ratio 1.38). Per-τ estimates are independent across events but not across τ; inference is pointwise.

S2.2 Clean-event robustness

A contamination metric is computed for each event. The metric is the number of other in-sample events whose dates fall within ±8 weeks of event e and whose CBG catchments overlap with e's. Of the 51 main-sample events, 34 have zero contaminating neighbors. Of those 34, 24 fall in the 2024–2026 subsample for which the pre-event state-Trends baseline is internally consistent (see the main paper). On this subsample the saturation-test correlation between the late-window estimate and the pre-event state-Trends baseline strengthens slightly relative to the full 2024–2026 Trends sample, in the same direction and with the same statistical strength. The leave-one-out diagnostic for the headline Spanish-ICE horse race keeps the coefficient close to the full-sample point estimate when any single event is dropped. The corresponding English-GDELT leave-one-out diagnostic leads to the same conclusion.

As the sample is restricted to events with progressively flatter pre-trends, the precision-weighted regression coefficient stays between -0.018 and -0.022 log points per standard deviation and remains negative and statistically significant. This holds as the pre-trend cutoff tightens from the full sample to the events that fail to reject flat trends at $p \geq 0.25$, a restriction that discards 60 percent of events. The confidence interval widens as the sample shrinks while the point estimate is stable. Only at the two strictest cutoffs, which retain about 14 events, does the coefficient attenuate to roughly -0.014 and lose significance, consistent with a loss of power rather than a change in the underlying relationship. The effect of news coverage is therefore not an artifact of the events whose pre-trends are least credible.

Table S9 addresses the timing of the event within its week. The main specification anchors $\tau = 0$ at the Monday of the week containing the event, so the event week mixes pre- and post-event days when an event does not fall on a Monday. Re-anchoring $\tau = 0$ at the first Monday on or after the event date, with the reference week moved to $\tau = -2$ so the partial event week never enters the reference, leaves the heterogeneity decomposition and the effect of news coverage essentially unchanged, and the per-event estimates correlate at 0.975 with the baseline. Dropping the partial event week entirely reproduces the baseline exactly, because within a single-event panel event time and calendar week are collinear, so the partial week's observations are absorbed entirely by its own coefficient and cannot enter the late-window estimate.

Table S9: Robustness of the late-window response to the event-time definition

	Original	Monday after	Donut
Panel A: Late-window response and heterogeneity			
Events	51	51	51
RE pooled mean	-0.0049	-0.0063	-0.0049
(RE SE)	(0.0049)	(0.0048)	(0.0049)
95% prediction interval	[-0.057, 0.047]	[-0.055, 0.043]	[-0.057, 0.047]
τ (REML, log points)	0.0287	0.0259	0.0287
I^2	69.7%	65.5%	69.7%
Panel B: Precision-weighted regression (per SD Spanish-ICE spike)			
Spanish-ICE spike	-0.0184	-0.0159	-0.0184
(KH SE)	(0.0051)	(0.0051)	(0.0051)
KH p	0.0006	0.0028	0.0006
True variance explained	30.1%	25.0%	30.1%
N	51	51	51

Notes: Each column re-estimates the per-event distributed-lag DiD, the random-effects heterogeneity decomposition, and the precision-weighted regression under a different event-time definition. *Original* anchors $\tau = 0$ at the Monday of the week containing the event, so the event week mixes pre- and post-event days for non-Monday events. *Monday after* anchors $\tau = 0$ at the first Monday on or after the event date, so $\tau = 0$ is the first full post-event week; the partial event week enters as its own coefficient at $\tau = -1$ and the reference week moves to $\tau = -2$. *Donut* keeps the original anchor and drops the partial event week from the sample (reference $\tau = -1$). The late window is $\tau \in [4, 8]$ throughout; per-event standard errors are the CBG-clustered standard error of the late-window mean (linear combination), not the mean of weekly standard errors. Panel B reports precision-weighted regressions with Knapp–Hartung inference; spikes are standardized within each column's estimation sample.

Table S10 reports the effect of news coverage estimated separately by enforcement wave. Within the 2025–2026 wave alone, which contains 44 of the 51 events, the effect is at least as strong as in the full sample, so the cross-event relationship is not an artifact of the structural differences between the two enforcement waves. The 2018–2019 wave has only seven events, too few for inference, but its slope has the same sign.

Table S10: Effect of news coverage by enforcement wave

Sample	N	Spanish-ICE coef (per SD)	p	I^2
Full sample	51	−0.0184	0.001	69.7%
2025–2026 wave	44	−0.0199	0.001	73.1%
2018–2019 wave	7	−0.0149	n/a	n/a

Notes: Precision-weighted regression of the per-event late-window estimate on the standardized Mediacloud Spanish-ICE spike, estimated separately by enforcement wave, with Knapp–Hartung inference. The spike is standardized within each sample. The 2018–2019 wave has only seven events, too few for inference, so its slope is reported descriptively and I^2 and p are omitted.

The effect is also unlikely to be an artifact of mismeasured operational scale. Table S11 shows that the attention coefficient is essentially unchanged when arrests, operation type, and sector are added as controls, that attention dominates arrests in a per-standard-deviation horse race, and that the small arrest coefficient remains small relative to the attention coefficient even under generous corrections for measurement error in arrests. Arrests are the only continuous operational-scale measure the shock inventory records, so it is controlled for directly rather than by building a scale index from a single variable. Operation type and sector separately capture composition.

Table S11: Operational scale and the effect of news coverage

	Attn. coef	(KH SE)	p	True var. expl.	N
Panel A: Attention coefficient as scale controls are added					
Spanish-ICE spike	-0.0184	(0.0051)	0.0006	30.1%	51
+ log arrests	-0.0192	(0.0053)	0.0007	25.1%	49
+ arrests, type, sector	-0.0190	(0.0074)	0.0141	15.1%	49
English-GDELT spike	-0.0154	(0.0050)	0.0036	18.5%	51
+ log arrests	-0.0161	(0.0052)	0.0035	16.9%	49
Panel B: Per-SD horse race, attention vs. log arrests					
News coverage spike (per SD)	-0.0192	(0.0053)	0.0007	25.1%	49
Log arrests (per SD)	-0.0040	(0.0055)	0.4675	–	–
Panel C: De-attenuated arrests coefficients					
Reliability ρ	Arrests coef.	Ratio			
1.0 (no error)	-0.0030	0.16			
0.7	-0.0042	0.23			
0.5	-0.0059	0.32			
0.4	-0.0074	0.40			

Notes: Precision-weighted regressions of the per-event late-window estimate on standardized regressors, with Knapp–Hartung adjusted inference and inverse-variance weights from the clustered linear-combination standard errors (same estimator as the main paper's precision-weighted regression). Arrests are the only continuous operational-scale measure recorded in the shock inventory: fines are unrecorded for every event, and operation end-dates are recorded for only 1 of 51 events (and reflect case resolution rather than operation length), so arrests are controlled for directly rather than building a scale index from one variable. Operation type/composition controls are worksite operation, sector agriculture, sector food, sector manufacturing, sector construction. Publicity is a visibility measure and is deliberately excluded from the scale block. Panel C reports the implied true single-regressor arrests coefficient b_{arr}/ρ under classical measurement error with reliability ρ; even at $\rho = 0.5$ the implied coefficient is small relative to the attention coefficient. Attention coefficients are per standard deviation of the spike; the arrests specifications have $N = 49$.

The headline precision-weighted regression is also robust to how its inference is conducted. Table S12 reports a randomization-inference p-value that permutes the news coverage spike across events ($p = 0.0006$), leave-one-out estimates that remain negative and significant in all 51 fits, and the heterogeneity standard deviation under three estimators (DerSimonian–Laird, REML, and Paule–Mandel), all between 0.025 and 0.031 log points. The effect is not an artifact of the normal approximation underlying Knapp–Hartung inference, of a single influential event, or of the τ^2 estimator.

Table S12: Inference robustness for the precision-weighted regression

Panel A: Attention slope, alternative inference	
Slope (per SD)	-0.0184
Knapp–Hartung p	0.0006
Randomization-inference p	0.0006
True variance explained	30.1%
Panel B: Leave-one-out and influence	
LOO slope range	[-0.0202, -0.0161]
LOO fits significant at 5%	51 of 51
Most influential event (drop)	slope $\rightarrow -0.0161$
Panel C: Heterogeneity SD τ by estimator	
τ (DerSimonian–Laird)	0.0254
τ (REML)	0.0287
τ (Paule–Mandel)	0.0314
I^2	69.7%

Notes: $N = 51$. Panel A compares the Knapp–Hartung p-value for the precision-weighted regression slope (per-event late-window estimate on the standardized Spanish-ICE spike, inverse-variance weighted) with a randomization-inference p-value that permutes the spike across events 10,000 times and recomputes the slope; the permutation test makes no normality or large-sample assumption. Panel B drops each event in turn: the slope range, the number of leave-one-out fits that remain significant at 5%, and the slope when the single most influential event is dropped. Panel C reports the heterogeneity standard deviation τ under three estimators. The slope is negative and significant under permutation inference and in every leave-one-out fit, and τ is stable across estimators.

The effect also does not reflect how events entered the inventory. Table S13 classifies each event by the type of its source citation. Events documented through news outlets carry higher news coverage spikes than those documented through official or other channels, but discovery source does not predict the per-event estimate, and the effect of news coverage is, if anything, stronger under discovery-source fixed effects.

Table S13: Discovery-source robustness

Panel A: News coverage spike and estimate by source type			
Source type	Events	Mean spike	Mean β_{late}
news/web	32	0.0340	-0.0094
official	9	0.0349	+0.0050
other	10	0.0116	-0.0156
Source → spike, joint F		5.25	$p = 0.01$
Source → β, joint F		1.63	$p = 0.21$
Panel B: Effect of news coverage with source fixed effects			
	Slope	(KH SE)	p
Baseline	-0.0184	(0.0051)	0.0006
+ discovery-source FE	-0.0228	(0.0051)	0.0001

Notes: $N = 51$. Each event is classified by the type of its inventory source citation (official government, news/web, advocacy/social, or other; categories smaller than 5 events are pooled into "other"). Panel A reports the mean news coverage spike and mean late-window estimate by source type and the joint F from regressing the standardized spike and the estimate on source dummies. Panel B is the precision-weighted attention regression with and without discovery-source fixed effects. The source citation is a proxy for the discovery channel, not a measured discovery process; this is a selection probe, not an identifying control.

S3 Instrument robustness and balance

Table S14 reports the full OLS-versus-IV comparison in per-unit-of-spike units, including the overidentified specification with both disaster instruments and the U.S.-disaster instrument alone.

Table S14: OLS and IV estimates of the effect of ICE media coverage on foot traffic

	(1) OLS	(2) IV: Foreign disasters	(3) IV: US disasters (cross-state)	(4) Both IVs (over-id)
Panel A: Mediacloud Spanish-ICE spike (endogenous)				
Coverage coefficient	-0.919^{***}	-1.059^{***}	-0.911	-1.052^{***}
	(0.256)	(0.338)	(0.796)	(0.340)
t-statistic	-3.59	-3.13	-1.14	-3.09
First-stage F on instruments	—	52.66	28.44	25.78
Sargan over-id (p)	—	—	—	0.830
Panel B: English-GDELT spike (endogenous, corroborating)				
Coverage coefficient	-1.338^{***}	-1.699^{***}	-2.272	-1.682^{***}
	(0.438)	(0.565)	(1.779)	(0.563)
t-statistic	-3.05	-3.01	-1.28	-2.99
First-stage F on instruments	—	53.12	12.85	26.07
Sargan over-id (p)	—	—	—	0.748
N	51	51	51	51

Notes: Each panel reports the coefficient on the post-event news coverage spike in a regression of per-event β^e_{late} on attention, with no other covariates. Column (1) is OLS. Columns (2)–(4) are two-stage least squares (2SLS) specifications using predetermined news-shock instruments. The foreign-disaster instrument is the number of days in the post-event window $\tau \in [0, 8]$ that fall within the first seven days of a GDACS red-alert sudden disaster outside the United States, Latin America, and the Caribbean. The US-disaster instrument counts major natural disasters in a different US state from the ICE event. Column (4) uses both instruments jointly. First-stage statistics are robust Wald F tests for the excluded instruments. Heteroskedasticity-robust standard errors in parentheses. Stars: $^{*}\, p < 0.10$, $^{**}\, p < 0.05$, $^{***}\, p < 0.01$.

Table S15: Competing-news IV robustness: foreign disasters and controls

	(1) Foreign	(2) + sev./era	(3) + event ctrls	(4) Both IVs
Attention coef.	-1.059^{***}	-1.123^{***}	-0.993^{**}	-1.052^{***}
	(0.338)	(0.341)	(0.475)	(0.340)
t-statistic	-3.13	-3.29	-2.09	-3.09
First-stage F	52.66	77.83	24.80	25.78
Sargan over-id (p)	—	—	—	0.830
Foreign-disaster IV	Yes	Yes	Yes	Yes
US-disaster IV	No	No	No	Yes
$\log(1 + \text{arrests}) + 2025+$	No	Yes	Yes	No
Publicity + sector ctrls	No	No	Yes	No
N	51	49	49	51

Notes: The dependent variable is per-event β^{e}_{late} and the endogenous regressor is the Mediacloud Spanish-ICE news coverage spike. Columns (1)–(3) use the foreign-disaster instrument alone. Column (4) uses the foreign-disaster and US-disaster instruments jointly. Severity/era controls are $\log(1 + \text{arrests})$ and an indicator for events in 2025 or later. Event controls add the event-publicity flag and broad sector indicators for agriculture/dairy, restaurant/food, manufacturing/industrial, car wash, and construction. First-stage statistics are robust Wald F tests for the excluded instruments. Heteroskedasticity-robust standard errors in parentheses. Stars: $^{*}\,p < 0.10$, $^{**}\,p < 0.05$, $^{***}\,p < 0.01$.

Table S15 shows that the foreign-disaster IV result is not an artifact of obvious event-level controls. Adding $\log(1+\text{arrests})$ and a second-Trump-era indicator leaves the estimate close to the headline foreign-disaster estimate. Adding event-publicity and broad-sector controls attenuates the estimate but preserves the sign and statistical strength. The overidentified specification with foreign disasters and out-of-state U.S. disasters also leads to the same conclusion.

Table S16 unpacks the joint balance test reported in the paper's appendix. The rejection is driven almost entirely by the mechanical correlation between foreign-disaster timing and the cross-state U.S.-disaster measure, which is the second instrument rather than a confounder of foot traffic. The predetermined design-quality covariates, the pre-event coefficient and the backward placebo, are each individually unrelated to the instrument. The foreign-disaster second-stage estimate is stable when these covariates are added as controls, moving from -1.06 to -1.04, so the balance result reflects a controllable timing correlation rather than a channel that drives the estimate.

Table S16: IV balance: what drives the joint test, and conditional-IV survival

Panel A: Loadings of predetermined covariates on the instrument (per SD)				
	Coef		t	p
Log arrests	+0.237		+1.26	0.208
Pre-event β	-0.129		-1.14	0.255
Backward placebo β	-0.081		-0.47	0.637
Cross-state disaster days	+0.338		+3.05	0.002
Joint F (all four)	4.35			0.006
Panel B: Foreign-disaster IV second stage, conditioning on the imbalance covariates				
	Coef	(SE)	First-stage F	N
No controls (baseline)	-1.059	(0.338)	52.7	51
+ imbalance covariates	-1.056	(0.324)	62.0	49
+ imbalance, era, publicity	-1.040	(0.359)	92.4	49

Notes: Panel A regresses the standardized foreign-disaster instrument on the four predetermined covariates that enter the joint balance test reported in the paper's appendix, with the same era, publicity, and broad-sector controls; coefficients are per standard deviation of each covariate. The joint F reproduces the appendix value and identifies which covariate carries the rejection. Panel B re-fits the foreign-disaster 2SLS coverage-on-foot-traffic estimate, adding the predetermined covariates as controls (the cross-state disaster measure is the second instrument, not a control). The second-stage coefficient is stable across specifications, so the balance result reflects a controllable mechanical correlation rather than a violation that drives the estimate. The conditional specifications drop two events with missing arrest counts ($N = 49$). Heteroskedasticity-robust inference throughout.

S4 Exposure-measure robustness

The headline high-FB indicator uses share of Latin-American foreign-born population (ACS B05006) at the CBG level. This is a broad measure because it counts naturalized US citizens alongside non-citizens, all Latin-American national-origin groups alongside the highest-enforcement-risk Mexican and Central American populations, and all adults alongside the US-citizen children of foreign-born parents. To test whether the chilling response is concentrated in narrower subgroups most directly exposed to deportation risk, the per-event distributed-lag DiD is refit with three sharper exposure indicators. These are the share of Hispanic non-citizen adults, the share of residents born in Mexico, El Salvador, Guatemala, Honduras, or Nicaragua, and the share of US-citizen children under 18 with at least one foreign-born parent.

Table S17: Exposure-measure robustness, per-event estimate on Spanish-ICE spike by high-FB indicator

Exposure indicator	N	mean estimate	sd estimate	r vs base	coef	t	R^2
Latin-Amer. FB (baseline, B05006)	51	−0.011	0.044	+1.000	−0.92	−3.59	0.227
Hispanic non-citizen (B05003I)	51	−0.007	0.046	+0.440	−0.17	−0.41	0.006
Mex. + Cent. Amer. FB (B05006)	51	−0.006	0.043	+0.716	−0.58	−1.98	0.089
Mixed-status kids (B05009)	49	−0.001	0.040	+0.551	−0.29	−1.05	0.027

Notes: For each exposure indicator, the per-event distributed-lag DiD (main paper) is refit, defining high-FB CBGs as the top-decile and low-FB CBGs as the bottom-decile within each event's catchment by the indicated variable. "Mean estimate" and "sd estimate" are the cross-event distribution of the resulting late-window estimates. "r vs base" is the cross-event correlation of the refined-indicator estimates with the baseline estimates. The last three columns report the single-regressor OLS of the late-window estimate on the Mediacloud Spanish-ICE spike with HC1 standard errors.

All three refinements produce the same negative sign as the baseline at smaller mean magnitudes and substantially weaker cross-event predictive power for the Spanish-ICE news coverage spike. The pattern suggests that the cross-event response operates at the broad community level rather than narrowly at the personal-deportation-risk level. This weakening is not mechanically driven by fewer POIs in the sharper top-decile cells. Top-decile POI density under the sharper indicators is within ±6 percent of the baseline.

S5 Venue-coverage and civic-mechanism checks

Advan's POI panel does not measurably capture two venue categories that are central to the sensitive-locations policy and to the broader immigration-fear literature. Religious organizations (NAICS-3 = 813) contain 4,025 POIs across the 26-state catchment universe, of which 3,817 (95 percent) are the Independent Community Bankers of America contamination documented in the POI data-quality audit in the Online Appendix. After stripping obvious mislabels, only a very small number of panel POIs correspond to real religious congregations. Healthcare venues are absent. NAICS-2 = 62 has zero POIs in the panel. The civic-versus-commercial decomposition therefore captures the schools and parks/museums response but is silent on churches and hospitals.

Table S18 reports the cross-event inference behind the civic-versus-commercial contrast in the main paper.

Table S18: Civic vs. commercial response, with inference

	N	Mean	(SE)	t
Civic estimate	12	-0.120^{**}	(0.045)	-2.67
Commercial estimate	12	-0.010	(0.009)	-1.03
Civic $-$ commercial	12	-0.111^{**}	(0.040)	-2.79
High-attention events	6	-0.060	(0.044)	
Low-attention events	6	-0.162	(0.063)	

Notes: Sample is the 12 events with both a civic (NAICS-3 712 and 611: parks, museums, educational services) and a commercial (NAICS-2 44, 45, 72: retail and food services) per-event estimate. Each row reports a one-sample t-test across events of the indicated per-event estimate; the differential is computed within event before averaging. Inference is at the cross-event level and does not down-weight noisier event estimates. The attention split uses the Mediacloud Spanish-ICE spike. Stars: $^{*}\,p < 0.10$, $^{**}\,p < 0.05$, $^{***}\,p < 0.01$.

I also test whether attention moderates the civic-vs.-commercial differential. With the English-GDELT spike, the within-event civic-minus-commercial differential is -0.153 in the low-spike group and -0.051 in the high-spike group. With the Mediacloud Spanish-ICE spike, the pattern is similar. The differential is -0.162 in the low-spike group and -0.060 in the high-spike group. The civic-vs.-commercial gap is wider in low-attention events than in high-attention events under both measures, the opposite of what an attention-moderated mechanism would predict.

S5.1 Spending

Table S19 reports the card-spending response by attention. Per-event spending estimates come from the exact-geocoded SafeGraph SpendPatterns subsample. The 30 main-sample events with both a spending estimate and a news coverage spike enter this comparison, and the estimates are much noisier than the foot-traffic estimates. The cross-event aggregate is not distinguishable from zero. Descriptively, above-median-attention events show a differential decline of about 3 percent while below-median-attention events show a small increase of about 4 percent, but the cross-event slope of the spending estimate on the news coverage spike is not statistically significant, so the spending evidence is suggestive rather than conclusive.

Table S19: Spending response by attention

	N	Mean β_{late}	(SE)	Implied % decline
Panel A: Cross-event aggregate				
All events	30	+0.0069	(0.0314)	−0.7
Panel B: By median Spanish-ICE attention				
Above-median attention	15	−0.0278	(0.0285)	+2.7
Below-median attention	15	+0.0415	(0.0556)	−4.2
Panel C: Cross-event slope of β_{late} on the standardized spike (per SD)				
Spanish-ICE spike	30	−0.0284	(0.0329)	
English-GDELT spike	30	−0.0285	(0.0339)	

Notes: Per-event late-window spending estimates β^{e}_{late} from the per-event distributed-lag DiD specification (main paper) fit on log(1 + weekly spending) over the exact-geocoded SafeGraph SpendPatterns subsample, the 30 main-sample events with both a spend estimate and a Spanish-ICE news coverage spike. A positive β_{late} means high-FB spending rose relative to low-FB; the implied percent decline is $100\,(1 - \exp(\text{mean } \beta))$, so a negative mean β prints as a positive decline. Panel A reports the cross-event mean and its standard error; the aggregate is not distinguishable from zero. Panel B splits events at the median Spanish-ICE spike. Panel C regresses β^{e}_{late} on the standardized news coverage spike with HC1 standard errors; coefficients are per standard deviation of the spike. Per-event spend estimates are much noisier than for foot traffic, so these are order-of-magnitude patterns rather than precise effects. Stars: $^{*}\, p < 0.10$, $^{**}\, p < 0.05$, $^{***}\, p < 0.01$.

S6 POI data-quality audits

S6.1 Advan and SafeGraph POI classification

This section discusses non-trivial point-of-interest classification errors in the underlying SafeGraph POI universe. These include NAICS misclassification across sectors, POIs identified as active businesses that are in fact vacant or demolished structures, and visit-count series that are mathematically inconsistent with the POI's physical footprint.

S6.2 Geometric POI flags

Two independent quality flags are constructed for all Advan POIs in the catchments of every event in the sample, covering approximately 1.54 million distinct POIs across the event catchments used in the audit. Building-containment is tested against Microsoft Building Footprints, and freeway-proximity is tested against TIGER Primary and Secondary Roads. A POI receives the freeway-traffic flag if its lat/lon falls outside any building footprint and is within 50 meters of an interstate centerline. A POI receives the remote flag if it falls outside any building, more than 100 meters from the nearest building, and more than 200 meters from any S1100 or S1200 road.

Across the 26-state panel, 7,881 POIs receive the freeway-traffic flag (0.5 percent) and 20,324 receive the remote flag (1.3 percent), for a combined drop rate of about 1.8 percent. The flagged set catches the expected error mode. In California, the five most-visited flagged POIs are the Port of Los Angeles, the Port of Long Beach, an Irwindale aggregates quarry, the Long Beach port-adjacent industrial complex, and the Port of Los Angeles San Pedro terminal. The flagged POIs are disproportionately in low-Latino-foreign-born CBGs rather than high-Latino-foreign-born CBGs, so removing them shrinks the comparison group rather than the treatment group. Re-estimating the per-event late-window effect on the cleaned panel shifts the cross-event mean by 0.0004 log-points. Three of 51 events flip sign, all with absolute estimates below 0.01 in either direction. The mean absolute change per event is 0.0056 log-points. The news-coverage finding does not depend on the small subset of POIs that fail the building-containment and freeway-proximity checks.

S6.3 POI verification and sector recoding

The geometric check is supplemented with hand-verification of 100 randomly sampled California POIs, stratified by Veridion firm-registry match quality. Of the 100, 96 correspond to a real business or commercial property at the listed address. The four exceptions are all residential single-family homes that Advan labels with a generic NAICS-category string, and all four fall in the no-Veridion-match tier where Advan errors should be concentrated. No POI in the sample was a demolished structure or a business that never existed.

A second verification of 500 California POIs stratified by NAICS-3 was designed to test whether classification noise could reshuffle the civic-versus-commercial buckets. The Advan-bucket $\times$ true-bucket cross-tabulation shows that the directional contamination that would erase the mechanism finding is essentially zero. Only 1 of 200 civic-labeled POIs is actually commercial. The audit did identify a consequential bug in NAICS-3 = 813. In California, 110 of the 118 Advan POIs in NAICS-3 = 813 are labeled "Independent Community Bankers of America" but correspond to individual community-bank branch addresses. A further three NAICS-813 records are sports stadiums. NAICS-3 = 813 is therefore dropped from the civic bin in the civic-versus-commercial decomposition. Including 813 attenuates the civic estimate toward zero by mixing in community-bank visitor activity whose own per-event response is closer to the commercial-bucket value.

S6.4 Structural name-level checks

The structural name-recode and residential-mislabel exercises were run on the earlier 21-state catchment universe (California plus the 20-state extension, 1,155,235 POIs), a strict subset of the 26-state estimation panel. Because these audits bound Advan's classification-error rates rather than filter the estimation sample, the narrower universe is sufficient. To move from spot-checks to panel-wide quantification, a verification exercise is conducted at the name-recurrence level. This universe contains only 129,773 unique `location_name` values. The top 1,500 names cover 978,660 POIs. Each of the top names is classified as a national chain with consistent NAICS, a generic NAICS-category-as-name label, a public-sector entity, a brand-at-host pattern in which a service brand sits inside a host retailer, or a trade-association-at-members pattern in which an association name is spread across member-business addresses.

The verification identifies 79 names with bulk-recodable structural miscoding, of which 44 require an actual NAICS-3 change. The 44 effective recoding rules cover 63,924 POIs (5.5 percent of the 21-state universe by POI count and 5.3 percent of the panel by visitor-week count). The largest pattern is money-transfer-agent POIs that live inside grocery and convenience-store hosts. For the civic-versus-commercial decomposition, however, the implication is limited. Only 148 POIs would leave the civic bucket under the recoding rules, and the audit identifies no name that would move POIs from civic to commercial or commercial to civic. Re-estimating the category decomposition after applying the full set of structural recoding rules moves the high-baseline civic estimate from -0.135 to -0.120, while the commercial estimate moves from -0.001 to $+0.001$. The order-of-magnitude civic-versus-commercial gap is preserved.

S6.5 Panel coverage

A separate concern is that businesses may be real but missing from the Advan POI universe altogether. If Advan systematically under-covers immigrant-serving businesses, the chilling regression would be attenuated downward in heavily Latino CBGs because the relevant POIs are not in the panel. This is tested using Veridion as an external firm registry. For each 5-digit ZIP in the panel, Advan POIs and Veridion firms are counted and the coverage ratio computed. Restricting to dense urban ZIPs and stratifying by foreign-born share quartile, the firm-weighted Advan/Veridion ratio is 0.28 in the lowest-FB quartile and 0.33 in the highest-FB quartile. The mild positive gradient does not support systematic undercoverage of immigrant neighborhoods.

Two Census establishment references corroborate the Veridion-based finding. At the county $\times$ NAICS-3 level, comparing Advan POIs to Census County Business Patterns, the establishment-weighted Advan/CBP coverage ratio rises monotonically from the lowest-FB quartile to the highest-FB quartile. At the ZIP level, comparing against Census ZIP Business Patterns, the Advan/ZBP coverage ratio is essentially flat across foreign-born-share quartiles. Three independent external references therefore point the same direction. The panel-composition channel is a real concern in principle, but it is not the empirical source of the result.

S6.6 Residential labels and visit-pattern mismatches

A further exercise targets residential single-family homes that Advan labels with a generic NAICS category string. A residence-mislabel flag is constructed, combining a generic NAICS-category-as-name label with a strict-residential street suffix. The flag identifies 28,204 POIs (2.4 percent of the 21-state panel). A stronger version that also requires no Veridion firm within 100 meters identifies 17,456 POIs (1.5 percent). Spot-checks confirm high precision. Only 157 of the flagged residences fall in the civic bucket, and dropping them would shrink the civic bucket by less than 1 percent. The remainder of the flag concentrates in the commercial and other buckets and does not shift the commercial estimate off its near-zero baseline.

A final check addresses visit-count series that are inconsistent with what the POI claims to be, using only the visit data itself. For each POI the Advan day-of-week and hour-of-day visit arrays are aggregated over all of 2025, each normalized to a shape profile, and a reference signature is built for every NAICS-3 with at least 30 well-measured POIs. A POI is flagged `flag_visit_pattern_mismatch` if both its day-of-week and hour-of-day profile correlate below 0.30 with its own sector's reference signature. The flag identifies 75,963 POIs (6.6 percent of the panel). It is the noisiest quality check. The flagged set mixes genuine anomalies with legitimate businesses that keep atypical hours. It is reported as a diagnostic rather than folded into the headline cleaning, because at this rate and with non-trivial false positives it is too blunt to use as a panel filter. Its value is that it catches a class of error that no external firm registry can detect, a real address with a real NAICS label but an implausible visit series.