Re-examining Moral Hazard in Risky Driving:

New Evidence from Behavioral Data in Auto Insurance

Yizhou Jin*

July 12, 2026

Abstract

This article measures ex-ante moral hazard in driving using insurance telematics data on U.S. rideshare drivers. We focus on drivers' handheld phone use ("HPU"), and find that it more than doubles the odds of a collision, yet drivers barely curb it when precipitation sharply amplifies its riskiness or when the platform raises their collision deductible. A one-time text-message warning produces a persistent 19.1% HPU reduction, but the response is only partially tailored to the incentives indicated by the message. Heavy phone users are the least responsive to everyday incentives yet the most moved by the intervention. In aggregate, moral hazard in HPU operates only weakly in everyday driving.

Keywords: Moral Hazard, Auto Insurance, Distracted Driving, Telematics, Road Safety
JEL Codes: D82, G22, K13, I12, R41

*Jin: University of Toronto, contact: yizhou.jin@utoronto.ca Previously circulated under the title "Re-examining Moral Hazard under Inattention." Thomas Yu provided excellent research assistance. I thank Xavier D'Haultfoeuille, Tatyana Deryugina, Ben Handel, Jonathan Kolstad, Scott Kominers, Ariel Pakes, Steve Tadelis, Tom Baker, Kevin He, and Jing Li for valuable comments and helpful suggestions, and the anonymous referees whose detailed feedback has significantly reshaped and benefited the paper.

1 Motivation and Introduction

Car crashes impose a staggering burden on society. The U.S. Department of Transportation estimates that, in 2019 alone, crashes in the United States killed 36,500 people and resulted in an economic cost of $340 billion. Including quality-of-life losses, the total societal harm from these crashes was $1.37 trillion (Blincoe et al. 2023).

This article investigates how risky driving behaviors respond to incentives. A central question is whether assigning financial responsibility to drivers deters accidents by incentivizing safer driving—the moral-hazard logic underlying the U.S. at-fault negligence regime (Shavell 1987), in contrast to higher-coverage Canadian and European systems that de-emphasize deterrence in favor of victim compensation (Wagner 2006).[1]

Such policy differences reflect unresolved theoretical ambiguity over the role of moral hazard in risky driving. When the cost of an accident rises, a rational driver may compensate by taking more care (Peltzman 1975; Wilde 1982). However, accidents are rare disasters that are multifactorial; drivers must also multitask across competing demands such as navigation. This profile invites systematic departures from the risk-compensation benchmark. Small risks can be elevated, either overweighted or over-perceived, as evidenced by consumers' tendency to over-insure modest risks (Barseghyan et al. 2013; Sydnor 2010). Yet they are just as readily neglected, whether under-perceived or unattended: most drivers rate themselves safer than average (Svenson 1981), people discount the odds of losses they have not lived through (Hertwig et al. 2004), and many forgo flood and catastrophe insurance (Gallagher 2014; Kunreuther and Pauly 2004) or skip low-cost preventive care (Baicker et al. 2015; Brot-Goldberg et al. 2017).

The strength of moral hazard in risky driving is ultimately an empirical question, but existing studies have been constrained by a lack of direct driving-behavior data. The literature has largely relied on indirect inference—testing for a positive correlation between coverage and claims (Chiappori and Salanie 2000), or exploiting exogenous coverage assignment and dynamic incentives such as experience-rating and monitoring (Abbring et al. 2003; Dionne 2013; Dionne and Liu 2021; Dionne et al. 2013; Israel 2004; Jeziorski et al. 2017; Jin and Vasserman 2021; Reimers and Shiller 2019; Weisburd 2015)—but cleanly separating ex-ante (driving behavior) from ex-post (claiming behavior) moral hazard, a

[1] "Reduction of risk through deterrence of harm is the true purpose of liability today, but compensation and avoidance of strife were also important historically" (Shavell 2003); see Calabresi (2008), James Jr. (1948), and Shavell (2018). Even in U.S. no-fault states, at-fault rules typically apply to severe lawsuits, and active efforts continue to repeal or limit no-fault systems (e.g. Florida Senate Bill 54, 2021; Michigan's 2019 insurance reform). Canadian and European systems mandate far higher liability coverage—ten to one hundred times higher than the U.S.—in which lawmakers "did not believe in the idea of prevention" (Wagner 2006).

distinction critical to welfare analysis, has been difficult with insurance data alone (Einav and Finkelstein 2018).[2] A parallel literature infers ex-ante moral hazard from aggregate fatalities, with mixed results. Cohen and Dehejia (2004) attribute a 6% rise in auto fatalities to no-fault reforms that reduce drivers' exposure to accident liabilities, whereas Cohen and Einav (2003) find no risk compensation at all, as mandatory seat-belt laws cut occupant fatalities without any offsetting rise in pedestrian deaths. Such fatality counts, though largely immune from under-reporting, are coarse, and their change can reflect injury severity, miles driven, or the risk composition of those miles as much as any change in driving behavior.

Recent advances in smartphone sensors and machine learning have made risky driving behaviors directly observable at scale. In this article, we obtain such data through a national auto insurer that underwrites the commercial coverage of a U.S. rideshare platform. We focus on drivers' handheld phone use (hereafter "HPU") because it can be identified with high precision from abnormal phone movements alone. Telematics data have appeared in recent insurance research, drawn from in-car hardware installed for drivers who opt into usage-based-insurance programs and summarized over days or months (Geyer et al. 2020; Jin and Vasserman 2021; Soleymanian et al. 2019). Our data are raw second-level smartphone-sensor streams, collected operationally to administer driver dispatch and navigation, and they pinpoint a specific behavior rather than a composite risk score.

We draw on three panel datasets: a five-second-interval sample focusing on drivers involved in collision accidents, in order to establish HPU's accident impact; a stratified trip-level sample from the first quarter of 2020, used for the precipitation analysis and a text-message experiment; and a stratified driver-day-level sample spanning a 2021 collision-deductible increase.

Our main results are as follows. First, in a discrete-time hazard model with driver fixed effects, we find that HPU multiplies the odds of a collision by 2.4 at mean conditions and by 2.5 during precipitation, based on second-level observations of accident-involved drivers and controlling for speed and time-of-day. This estimate is comparable to, if below, prior crash-risk estimates (Dingus et al. 2016; Klauer et al. 2014; McEvoy et al. 2005).

Despite this, HPU responds only weakly to precipitation shocks and the deductible increase. To establish this, we exploit both across- and within-trip variation in precipitation,

[2]Several of the studies above use the dynamics of experience-rated contracts to separate moral hazard from selection, but this leaves the ex-post margin untouched: the outcome is still a filed claim, so an estimate can conflate a change in driving with a change in whether an incident is reported. Strategies that instead target the reporting margin, such as restricting attention to severe or high-value claims that are hard to leave unreported, or contrasting claim categories with different reporting incentives, have not been systematically evaluated for how well they isolate ex-ante behavior.

absorbing location, date, hour, and driver fixed effects alongside a rich set of trip-level controls. We examine the one-time deductible increase in a difference-in-differences design, with treatment intensity given by each driver's predetermined expected out-of-pocket increase, measured from the insurer's county-level pricing. Across the two analyses, we calibrate the change in HPU's expected accident cost from our hazard-model results for precipitation and the insurer's own pricing for the deductible. This yields a moral-hazard elasticity, β: the percentage change in HPU in response to a one percent increase in its expected accident cost. The precipitation estimates give $|\hat{\beta}| \approx 0.005$–$0.032$; the deductible estimates are significant only on the extensive margin and only with across-state variation in expected out-of-pocket exposure, with $|\hat{\beta}| \approx 0.011$. Our ex-ante moral-hazard elasticity, measured on the risky behavior itself, is much more precisely estimated than, and sits at the lower end of, the claim- and fatality-based estimates of the prior literature, which range from indistinguishable from zero to $|\beta| \approx 0.1$–1.

This is not because HPU cannot move or our measure fails to detect it. We analyze a randomized experiment run on the platform, in which heavy phone users were sent a single text message stating that the platform had detected handheld phone use and reminding them that passenger reports of unsafe driving can lead to suspension. We find that it produces a sharp and persistent reduction in HPU of 19.1%. Yet the response is only partially targeted: HPU falls even on pickup segments, where no passenger is present to lodge the threatened complaint, though it falls more when a passenger is aboard. The drop is also specific to phone use, leaving other behaviors prone to safety complaints, such as harsh acceleration and cornering, untouched.

Interestingly, everyday incentives and the intervention sort drivers in opposite directions: the former disproportionately affects already-light phone users, while the latter primarily curbs HPU among heavy users. That these same heavy users cut HPU proportionally more in response to a costless message points to neglect (inattention or under-perception) of risk rather than an inability to self-regulate. Because they generate most of the exposure, everyday moral hazard in HPU is weak in aggregate.

The article proceeds as follows. Section 2 introduces our data. Section 3 establishes HPU's marginal contribution to accident risk in a discrete-time hazard model. Section 4 tests two implications of moral hazard using precipitation shocks and a collision-deductible increase. Section 5 reports the text-message experiment and analyzes heterogeneous treatment effects. We conclude by synthesizing the broader empirical literature on moral hazard in driving, revisiting competing theories, and discussing the policy implications and future research directions.

2 Background and Data

We focus on a U.S. ridesharing platform that dispatches drivers through its mobile application to pick up and transport passengers. The same application that coordinates a trip also records the sensor streams from which we measure behavior. The platform provides its own commercial auto insurance for dispatched driving, and transfers most of the risk to insurers. During our research window, the platform engaged our data-providing insurer for telematics event classification. The insurer received raw sensor data streams and used a proprietary algorithm to classify second-level telematics events: HPU, speed, harsh acceleration, cornering, and braking. To support safety interventions, the insurer retained telematics data for a stratified sample of drivers, summarized at the trip level before 2021 and at the driver-day level afterward.

2.1 Three Measurements

This subsection introduces the three measurement dimensions that anchor our analysis: HPU behavior, precipitation, and the rideshare-driver insurance environment.

HPU Behavioral Data The HPU measure is constructed from drivers' smartphone sensor records. The platform's application continuously records tri-axial accelerometer and gyroscope readings, speed, and magnetometer-based orientation from the driver's phone (Wahlström et al. 2017). These streams together permit inference of how the vehicle is moving and how the phone is moving relative to it—the empirical basis for distinguishing handheld interaction from passive transport in a cradle, cup-holder, or pocket. According to the insurer, the HPU classifier achieves over 99% per-second precision and industry-leading recall against ground-truth labels from FleetCam video footage, with the precision-leaning calibration reflecting a deliberate choice to limit liability from false-positive interventions; Online Appendix E details the underlying machine-learning pipeline. The present paper takes the telematics data as given but corroborates its empirical relevance in two ways: by establishing HPU's effect on accident risk (Section 3) and by demonstrating that the measure tracks sharp behavioral responses to a real-time intervention (Section 5).

HPU vs. Other Risky Behaviors We focus on HPU rather than other risky driving behaviors for two reasons. First, HPU can be identified with substantially greater precision because its detection is largely unaffected by environmental factors. HPU is primarily associated with short and consecutive spurts of phone gyration that indicate human interactions with the phone (Engelbrecht et al. 2015). By contrast, behaviors inferred from vehicle motion, such as harsh cornering, produce more diffuse and smoother sensor devi-

ations that are harder to attribute unambiguously. Moreover, behaviors such as driving under the influence ("DUI") and speeding require accurate measurement of instantaneous environmental constraints such as weather and posted speed limits.

Second, HPU has been extensively studied in laboratory settings and, given its increasing prevalence over the past two decades, has been a frequent target of safety campaigns and policy interventions (Llerena et al. 2015; Megat-Johari et al. 2023; Strayer and Johnston 2001; Wilson and Stimpson 2010). A parallel public-health literature has run randomized field trials showing that feedback and incentives can curb handheld use (Delgado et al. 2024; Ebert et al. 2024), which we discuss in Section 5.

Precipitation The platform continuously pings the DarkSky API to retrieve minute-by-minute precipitation type and intensity at the calendar-minute resolution, which DarkSky generates from contemporaneous radar imaging; this allows the feed to capture short-duration intermittence, such as showers and the start and end of squalls, that hourly station observations would smooth out. The platform stores these minute-level values at geohash6 spatial resolution—approximately $1.2\,\text{km} \times 0.6\,\text{km}$ rectangles—across all urban geohashes that the platform serves, with a near-perfect match rate to our trip-level data. For the interval-level panel, the platform matches precipitation to each within-trip calendar minute and geohash6, giving sub-trip variation in precipitation exposure; for the coarser trip-level panel, it matches only at the trip's start and end points and their calendar minutes.

Insurance and Trip Definition Commercial coverage applies throughout each *trip*, beginning when the driver accepts a request and ending at passenger drop-off. For a set of analyses in Section 5, our data splits each trip into two *segments*—a pickup segment, while the driver is en route to the passenger(s), and a passenger-aboard segment.[3]

The policy covers third-party bodily-injury and property-damage liability (BI/PD), uninsured/underinsured motorist coverage (UI/UIM), and, in a few no-fault states, personal injury protection and medical payments (PIP/MEDPAY); every limit meets its state mandate and historically sits well above it (BI/PD carries a $1 million limit). Collision damage to the driver's own vehicle is covered as well, subject to a deductible the platform raised from $1,000 to $2,500 on March 1, 2021—the natural experiment that anchors Section 4.2, where we measure the expected out-of-pocket increase this hike created and drivers' behavioral response to it.

[3]This split is important for the insurers because the pickup segment incurs no passenger liability risk.

2.2 Three Datasets

We assemble three datasets from the insurer's telematics data. The first is a high-resolution accident-driver panel that records telematics data at five-second intervals in the moments leading up to a collision; it powers the hazard model in Section 3. The second is a stratified trip-level panel from early 2020 that oversamples higher-HPU drivers. In addition to HPU records, it includes driver and trip characteristics, matched precipitation, and—for the experiment—each trip's randomized treatment status; it supports the precipitation analysis in Section 4.1 and the text-message experiment in Section 5. The third is a driver-day panel covering the March 2021 collision-deductible change; it powers the deductible analysis in Section 4.2.

Dataset 1: Interval-level Accident-Trip Panel. We start constructing the dataset by focusing on a subset of accident trips, labeled based on driver or passenger reports, in which a vehicular collision is clearly identifiable based on telematics data, predominately accelerometer signatures, based on a proprietary algorithm that detects extreme deceleration followed by prolonged stopping. The critical reference point that anchors our analysis is "crash initiation": the moment when the collision deceleration begins, which can be at or before the time of the actual impact depending on whether and how the driver reacts before the crash. This also helps us avoid contaminating the analysis window with post-crash deceleration that could trigger false HPU detection.[4] For each driver, we also observe 10 randomly selected pre-accident trips by the same driver (twelve drivers with fewer than ten recorded prior trips contribute all of theirs), yielding 48,661 trips in the final panel.

The panel allows us to address the rarity of accidents by providing fine temporal resolution within each trip. In addition to HPU, we observe average speed at five-second intervals, minute-by-minute precipitation data, as well as whether the interval spans nighttime driving.[5] Panel A of Table 1 summarizes this dataset.

Dataset 2: Driver-Trip-level Panel. Our second dataset comprises trip-level records from January 21–March 30, 2020, the window that encompasses our text-message experiment and the precipitation analysis. It is drawn from the platform's full driver universe by a two-step stratification procedure over the look-back period of January 1–21, 2020: drivers are first restricted to regularly active drivers, then sorted by their average HPU intensity, the share of total drive distance with HPU detected. Our sample overweighs the right tail:

[4]Our data does not include trips that resulted in a fatality, or those that were flagged for assaults or other passenger-initiated incidents since these trips are subject to a more sensitive reporting and investigation pipeline that the insurer does not have visibility into.

[5]We are unable to merge in more information from our other datasets because, due to the sensitivity of accidents, this panel was created and merged by the platform internally and anonymized independently, with location and timestamps redacted and driver identifiers scrambled.

the top 5% HPU drivers are retained with certainty while the low-HPU mass is sampled at progressively smaller rates. Panel B of Table 1 reports unweighted summary statistics for this sample ($N = 1{,}821{,}908$ trips, $15{,}251$ drivers). For the text-message experiment, each trip additionally carries the driver's randomized treatment status and its timing relative to the message.

The average trip in the subsample is 8.5 miles and 21.6 minutes. The average driver is 37.7 years old and 21% are female. HPU intensity averages 66.5 m/km. Re-weighting by the inverse of sampling probabilities gives the fleet-representative statistics, where the mean HPU is 8.6 m/km (Online Appendix Table D3).

Dataset 3: Driver-Day Policy Panel. Our third dataset supports the collision-deductible analysis in Section 4.2. After 2020, the structure of the telematics data available to us changed: we observe driver activity aggregated to the *driver-day* level (daily HPU, driving time, and mileage). Like Dataset 2, the panel is constructed by stratifying on look-back HPU intensity—measured over the driver's recorded activity from January 1 to February 13, 2021—to oversample high-HPU and active drivers, and it spans February 14–March 9, 2021, ending with the wind-down of the insurer's contract with the platform.

For each driver, we observe demographics and a modal trip-start location during the look-back period, which we use to assign the driver's county, at which the deductible analysis's expected-cost dose is measured (Section 4.2). Section 4.2 details additional sample restrictions. The final dataset used in the deductible-analysis comprises $12{,}881$ drivers and $143{,}116$ driver-days. Panel C of Table 1 summarizes the dataset. As in Dataset 2, the top-HPU oversampling lifts mean HPU in this panel to 70.0 m/km.

3 Riskiness of HPU

Handheld cellphone use while driving is outlawed in 33 U.S. states, and texting while driving is banned in all states except Montana (Governors Highway Safety Association 2025). Enforcement, however, is weak because "pre-crash distractions often leave no evidence for crash investigators to observe and drivers are often reluctant to admit to having been distracted right before a crash" (Goodwin et al. 2015). As a result, distraction is recorded as a factor in only 8% of fatal crashes, and cellphone use in a far smaller share (National Highway Traffic Safety Administration 2024)—and even this understates the true incidence: in fatal crashes with independent evidence of phone use, national crash records coded the crash as cellphone-involved only about half the time (National

Table 1: Summary Statistics by Dataset

Variable	Mean	Std. Dev.	Pctl(10)	Median	Pctl(90)
Panel A: Accident-Driver Panel (Dataset 1) — $N = 12{,}965{,}344$ five-second intervals, 4,430 accident drivers					
HPU ($= 1$ if active)	0.039	0.194	0.000	0.000	0.000
Speed (mph)	27.5	29.4	0.8	17.0	75.4
Has precip. ($= 1$)	0.336	0.472	0.000	0.000	1.000
Precip. (in/hr | precip.)	0.186	0.281	0.010	0.060	0.511
Trip-mean precip. (in/hr | precip.)	0.186	0.074	0.104	0.178	0.274
Is night ($= 1$)	0.394	0.489	0.000	0.000	1.000
Accident ($= 1$)	0.00034	0.01848	0.00000	0.00000	0.00000
Panel B: Trip-level Panel (Dataset 2) — Jan–Mar 2020, $N = 1{,}821{,}908$ trips, 15,251 drivers					
Age (years)	37.7	11.5	24.8	35.0	55.0
Is female	0.21	0.41	0.00	0.00	1.00
Tenure (years)	1.5	1.2	0.2	1.3	3.2
Past rides completed	213.2	287.0	0.0	111.0	564.0
Distance (miles)	8.5	8.3	2.0	6.1	17.8
Duration (minutes)	21.6	12.9	9.7	18.6	36.8
Pct. drive-time with passenger	0.71	0.16	0.48	0.74	0.90
Speed (mph)	21.2	10.0	9.8	19.4	35.3
Is weekend	0.26	0.44	0.00	0.00	1.00
Is night	0.27	0.45	0.00	0.00	1.00
Has precip.	0.31	0.46	0.00	0.00	1.00
Precip. intensity (in/hr | precip.)	0.309	0.898	0.016	0.055	0.736
HPU (m/km)	66.5	130.1	0.0	9.0	207.2
Harsh acceleration per 1k miles	26.97	96.42	0.00	0.00	79.54
Harsh braking per 1k miles	53.85	126.27	0.00	0.00	172.99
Harsh cornering per 1k miles	6.96	34.36	0.00	0.00	0.00
Panel C: Driver-Day Policy Panel (Dataset 3) — Feb–Mar 2021, $N = 143{,}116$ driver-days, 12,881 drivers					
Age (years)	46.6	12.1	31.0	46.1	63.1
Is female	0.17	0.37	0.00	0.00	1.00
Tenure (years)	2.0	1.5	0.1	1.9	3.9
HPU (m/km)	70.0	136.9	0.0	9.5	218.0
Pct. drive-time with passenger	0.63	0.12	0.50	0.65	0.77
Driving hours per day	3.61	2.56	0.74	3.10	7.25
Driving miles per day	90.6	67.3	16.6	76.6	183.8
Expected OOP increase, \$1,000→\$2,500 (\$/yr)	62.1	15.5	42.3	61.3	91.3

Notes: Samples described in Section 2.2. *Panel A*: accident-trip case-control panel at 5-second intervals (Section 3); HPU is a 5-second indicator, and HPU m/km and trip-mean precipitation are trip-level collapses. *Panel B*: stratified trip-level Jan–Mar 2020 sample (driver attributes summarized over drivers, other rows over trips). *Panel C*: driver-day deductible panel (Section 4.2), dispatch phases pooled; the expected out-of-pocket increase is the deductible analysis's dose, ΔEOOP_c (Section 4.2), a predetermined driver-level characteristic summarized once per driver. Panels B and C oversample high-HPU drivers, so their unweighted HPU means exceed the fleet level.

Safety Council 2013).[6] This verification problem renders police-report-based risk inference unreliable (Bhargava and Pathania 2013; Levitt and Porter 2001). In crash investigation and liability adjudication, this gap is increasingly filled by telematics evidence (Bortles et al. 2017), which we use in this section to quantify HPU risk at scale.

We use a discrete-time hazard model to estimate the association between HPU and accident risk (Singer and Willett 2003). Dataset 1 partitions each trip into 5-second intervals, from departure to crash initiation for accident trips, and from departure to destination for control trips.[7] Each interval is an observation in which the outcome is 1 if the accident occurs in that interval and 0 otherwise. HPU and speed are measured at the interval level, with $\text{HPU}_{jt} = 1$ if any phone use is detected in interval t of trip j, and Speed_{jt} measuring the within-interval mean speed. Precipitation Precip_{jt} updates every calendar minute from the sub-kilometer minutely radar feed described in Section 2.1; Night_j is a trip-level indicator for night driving, determined from the trip's start timestamp and location and constant within the trip.

Sampling follows a case-control design: for each of the 4,430 drivers who experienced an accident, the accident trip and 10 randomly selected pre-accident trips by the same driver are included, yielding 48,661 trips and 12,965,344 interval-level observations. We denote the discrete-time hazard $h_{jt} = P(\text{accident in interval } t \mid \text{no accident before } t)$ and adopt the following canonical specification:

$$\log\left(\frac{h_{jt}}{1-h_{jt}}\right) = \alpha_{i(j)} + \text{HPU}_{jt}\left(\delta_0 + \delta' z_{jt}\right) + g(z_{jt}), \tag{1}$$

where $z_{jt} = (\text{Speed}_{jt}, \text{Precip}_{jt}, \text{Night}_j)'$ collects the driving conditions, so HPU's effect on the log-odds is $\delta_0 + \delta' z_{jt}$. The term $g(z_{jt})$ is a flexible control for conditions, including their main effects and pairwise interactions; $\alpha_{i(j)}$ is a driver fixed effect.

Identification Our approach is standard in the naturalistic driving study (NDS) literature, which uses case-control designs to estimate the association between HPU and crash risk (Dingus et al. 2016; Klauer et al. 2014; Redelmeier and Tibshirani 1997), relying on the consistency of slope coefficients in logistic regressions even when accident outcomes are over-sampled (Prentice and Pyke 1979). We advance the literature in a few ways. First, our measure of HPU is algorithmically generated from cheap and abundant sensor data rather than manual video coding, allowing us to scale to a larger sample of 12,965,344

[6] By comparison, in regions where police reports incorporate camera and phone records such as the ones used in this article, the share of fatalities attributed to phone use is substantially higher: 29.6% in Shanghai in 2014 (Chinadaily.com.cn).

[7] The partition is done backwards from destination or crash initiation to departure, and the intervals covering the departure, typically less than 5 seconds, are dropped.

5-second intervals by 4,430 accident drivers (compared to the 905 crash and 19,732 baseline 6-second intervals in Dingus et al. (2016)). Moreover, we observe the full interval-level time series of the same driver, allowing us to include driver fixed effects to better control for unobservable factors, such as drivers' habitual phone use propensity and baseline driver risk types, than the previous random-effect models.

We also report results that add trip fixed effects, but do not treat them as our main specification. Trip fixed effects absorb additional unobservables, but they identify the effect only from within the accident trip: the uniformly sampled control trips no longer contribute, and the crash interval is compared against its own run-up seconds. In those seconds HPU and the unobserved interval-level hazard would plausibly rise together toward the crash, so HPU is no longer conditionally independent of the hazard given the covariates, reflecting the upward bias discussed above. Some interval-level conditions that co-move with HPU within a trip, such as route complexity, traffic, or navigation demands, inevitably remain outside both fixed-effects structures; the anonymized telematics feed fixes the covariate set we observe, leaving no richer controls to absorb them.

Main results Estimates from our main specification are reported in Table 2 column (7). The HPU odds ratio at mean driving conditions is 2.37, implying that intervals with active phone use more than double the odds of accident occurrence.[8] This is similar to, if on the smaller side of, prior estimates: Dingus et al. (2016) report an overall handheld-interaction OR of about 3.6, Klauer et al. (2014) significant handheld-task ORs of 2.5 and above (with visual-manual sub-tasks such as texting and dialing larger still), and McEvoy et al. (2005) a fourfold increase in crash risk (OR 4.1, 95% CI 2.2–7.7).

With a positive HPU $\times$ Precip interaction ($\hat{\delta}_{\text{HPU}\times\text{Precip}} = 0.4068$, $p < 0.05$), HPU's expected marginal accident cost rises by 34.71% from dry conditions to the typical rainy intensity[9], driven both by the higher baseline crash likelihood that rain brings and by an even larger HPU odds ratio.[10] Speed and precipitation enter with expected signs; HPU's own risk multiplier is 38.79% higher at night than in daytime, and essentially flat in speed (2.74% per additional 10 mph).

The remaining columns move as expected: the pooled specifications (columns 1–3)

[8]This odds ratio is evaluated at sample-mean driving conditions, $\exp\big(\hat{\delta}_{\text{HPU}} + \hat{\delta}_{\text{HPU}\times\text{Speed}}\,\overline{\text{Speed}} + \hat{\delta}_{\text{HPU}\times\text{Precip}}\,\overline{\text{Precip}} + \hat{\delta}_{\text{HPU}\times\text{Night}}\,\overline{\text{Night}}\big)$, evaluated over all driving at $\overline{\text{Night}} = 0.394$ (Table 1) and an unconditional $\overline{\text{Precip}} = 0.0623$ in/hr. Speed is centered and has mean zero, so its term drops out.

[9]We use the precipitation mean in the Dataset 2 trip panel (0.2034 in/hr), the rain level at which Section 4.1 measures the behavioral response so that our calibration of the moral-hazard elasticity is apples-to-apples in Section 4.3 (Table A2).

[10]Even in the absence of an interaction effect between HPU and precipitation, precipitation lifts baseline hazard, which would scale HPU's marginal cost multiplicatively, as shown in prior studies of elevated crash risk during active precipitation (Stevens et al. 2019).

give an odds ratio of about 2.20, adding driver fixed effects nudges it up (columns 4 and 7), and the trip-fixed-effects columns (5 and 8) are larger, around 4, echoing our identification discussion above.

Table 2: Discrete-Time Crash Hazard Model Estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
HPU	0.8000***	0.7922***	0.7901***	0.9545***	1.3270***	0.5637***	0.7075***	1.2829***
	(0.0498)	(0.0497)	(0.0497)	(0.0625)	(0.0937)	(0.0838)	(0.0949)	(0.0983)
Speed (mph)		0.0110***	0.0114***	0.0128***	0.0158***	0.0081***	0.0086***	0.0157***
		(0.0005)	(0.0005)	(0.0005)	(0.0009)	(0.0008)	(0.0008)	(0.0009)
Precipitation			0.7281***	0.7528***	0.7356***	0.6754***	0.6923***	0.6833***
			(0.0552)	(0.0573)	(0.0818)	(0.0745)	(0.0770)	(0.0923)
Night	0.5236***	0.4055***	0.4225***	0.5234***		0.3466***	0.4333***	
	(0.0302)	(0.0303)	(0.0305)	(0.0354)		(0.0354)	(0.0398)	
HPU × Speed						0.0021	0.0027	0.0023
						(0.0017)	(0.0018)	(0.0024)
HPU × Precip						0.3751**	0.4068**	0.3406
						(0.1487)	(0.1669)	(0.2295)
Speed × Precip						0.0017	0.0021	−0.0018
						(0.0019)	(0.0021)	(0.0026)
HPU × Night						0.2996***	0.3278***	
						(0.1139)	(0.1210)	
Speed × Night						0.0050***	0.0068***	
						(0.0010)	(0.0011)	
Precip × Night						0.0084	0.0224	
						(0.1162)	(0.1228)	
Interactions	No	No	No	No	No	Yes	Yes	Yes
Driver FE	No	No	No	Yes	No	No	Yes	No
Trip FE	No	No	No	No	Yes	No	No	Yes
Observations (intervals)	12,965,344	12,965,344	12,965,344	12,965,344	840,152	12,965,344	12,965,344	840,152

Notes: $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$. Driver-clustered standard errors in parentheses (all columns; Abadie et al. (2023)). The dependent variable is an indicator for crash initiation in the 5-second interval; entries are discrete-time hazard logit coefficients (log-odds), with HPU and the event indicator coded as the within-interval maximum and speed and precipitation as within-interval means. Precipitation enters in inches per hour, so its coefficients (and the HPU × Precip, Speed × Precip, and Precip × Night interactions) are per inch/hr. **Column (7)**, driver fixed effects with interactions, is our preferred specification (shaded). Columns (5) and (8) add trip fixed effects and are identified only from accident-trip seconds ($N = 840{,}152$), which biases the HPU coefficient upward; we read them as upper bounds. Night and its interactions are trip-level and absorbed by the trip fixed effects in these columns. Columns (3)–(4) and (6)–(7) additionally control for weekend and season (not shown; likewise absorbed by the trip fixed effects in columns (5) and (8)). Fixed effects are estimated by `fixest::feglm`. Online Appendix Table D1 reports sensitivity to coarser interval widths.

Robustness Our robustness checks focus on alternative interval widths and sampling procedures. Online Appendix Table D1 varies the interval width from 5 to 30 seconds, holding the column (7) specification fixed. The HPU main-effect odds ratio remains stable: 2.03 at the canonical 5-second width (Table 2), and 1.81 at 30 seconds. Online Appendix Table D2 examines the case-control sampling design by comparing the main specification

against three alternatives: accident trips only with trip FE (no control trips), a reduced set of 2 controls per driver, and inverse-probability-of-sampling-weighted (IPSW) estimation. The HPU odds ratio at mean conditions is stable across the three case-control variants (main 2.37, 2-control 2.50, IPSW 2.40). The similarity of the main and IPSW columns confirms Prentice and Pyke (1979)'s result that case-control slope coefficients are consistent without reweighting.

4 Observational Evidence

Section 3 establishes that HPU substantially increases accident risk. We now ask whether drivers respond to changes in this risk, i.e., the empirical content of ex-ante moral hazard. We use "moral hazard" here to mean the responsiveness of HPU behavior to variation in its expected accident cost, distinct from the more commonly studied "ex-post" moral hazard around insurance claims and utilization. A standard self-protection framework (Appendix A) summarizes this responsiveness in a single elasticity, β, defined as the percent change in HPU per 1% change in HPU's expected per-unit accident cost. We estimate β by exploiting two sources of plausibly exogenous variation in this cost: precipitation shocks that amplify HPU's marginal accident impact, and a major deductible hike that raised driver-borne out-of-pocket cost on every collision claim. The answer, in short, is muted in both designs: a small, sign-consistent response to precipitation, and a deductible response that is only significant on the extensive margin when comparing drivers across states with different expected out-of-pocket exposures.

4.1 Behavioral Response to Precipitation Shocks

Section 3 established that precipitation sharply raises the marginal accident cost of HPU: by 34.71% from dry to mean-rainy conditions. If drivers internalize this heightened cost, they should reduce HPU during precipitation.

Dataset 2 is well-suited to test this. 31% of all trips experience some precipitation. Comparing very similar trips by the same driver, differences in precipitation or how it changes within the trip are plausibly exogenous to a driver's underlying demand for phone use. We regress HPU b_{ij} (driver i, trip j) on a flexible function of the trip's precipitation profile, controlling for driver and trip-level observables:

$$g(\mathbb{E}[\, b_{ij} \mid \cdot\,]) = f(R_{ij},\, Inch_{ij},\, \Delta Inch_{ij};\, \boldsymbol{\kappa}) + x'_{ij}\psi + \phi_{\ell(ij)} + \tau_{t(ij)} + \eta_{h(ij)} + \alpha_i, \qquad (2)$$

Here b_{ij} is the regressed HPU outcome and $g(\cdot)$ the estimator's link. The extensive-margin linear probability model takes $b_{ij} = \mathbf{1}\{\text{HPU}_{ij} > 0\}$, any handheld phone use on the trip, with g the identity; the second, a Poisson quasi-maximum-likelihood (QMLE) estimator, takes b_{ij} to be HPU intensity, the fraction of the trip's distance driven under handheld phone use, with g the log link, fit as a mileage-offset Poisson on HPU-miles[11] so that longer trips carry proportionately more weight. R_{ij} is an indicator that either the trip's starting or its ending precipitation is positive, and $f(\cdot; \boldsymbol{\kappa})$ takes one of two forms reported in Table 3: a simpler "avg-intensity" form $R_{ij}\,(\kappa_0 + \kappa\,\overline{Inch}_{ij})$ using the trip's average rain intensity, and a "start+within" form $R_{ij}\,(\kappa_0 + \kappa_1\,Inch^0_{ij} + \kappa_2\,\Delta Inch_{ij})$ separating the trip-start rain intensity $Inch^0_{ij}$ from its within-trip change $\Delta Inch_{ij}$, where the within-trip coefficient can further isolate potential selection on unobservable differences in drivers' need to use the phone when initiating trips under different precipitation conditions.

For each trip the platform supplies a harmonized vector of 15 trip-level covariates, x_{ij}, in three groups: *driver and trip economics* (driver earnings and the pickup-phase share of the trip), *origin demand* (population density and trip-start ride-request volume at the origin, measured at a finer geohash-6 geographic resolution and not absorbed by the location effects), and *in-app interactions* (number of pickups, drop-offs, reroutes, cancellations, added rides, and app crashes, rider calls and texts and platform notifications). These controls, especially the last group, plausibly control for demand-driven need to use the phone other than a moral-hazard response. By absorbing this platform- and passenger-generated demand for phone interaction, the specification isolates the driver's own private demand for handheld use. Online Appendix Table D6 reports coefficients for the main controls across the paper's canonical specifications.

$\phi_{\ell(ij)}$, $\tau_{t(ij)}$, and $\eta_{h(ij)}$ are additive location-pair, date, and hour-of-day fixed effects, where $\ell(ij)$, $t(ij)$, and $h(ij)$ denote the origin–destination location pair (each endpoint a geohash-5 cell), date, and hour of day of trip ij; and α_i is a driver fixed effect. We saturate the model progressively across the columns of Table 3, adding the trip-level controls and then driver fixed effects on top of the location-pair, date, and hour effects, and the precipitation response stays small and moral-hazard-signed across these saturations, which is inconsistent with large residual selection on weather-correlated driver or trip characteristics. Reporting both margins follows Chen and Roth (2024): trip-level HPU intensity is a non-negative rate with many zeros, and the pair avoids the unit-dependence of a separate log-of-positives regression. The Poisson estimator itself requires no count structure: its consistency requires only that the conditional mean is correctly specified,

[11] Because HPU intensity is a per-distance rate, its coefficients are invariant to the choice of distance unit: miles and kilometers cancel in $\log(\text{HPU-miles}) - \log(\text{miles})$.

so it applies to continuous non-negative outcomes such as ours, with driver-clustered standard errors valid under arbitrary dispersion; it is the estimator Chen and Roth (2024) recommend for proportional effects on zero-heavy outcomes.

Our preferred estimates are reported in columns (3) and (7) of Table 3 with full controls and fixed effects. HPU reduces by a small amount in response to both across- and within-trip precipitation increases, corresponding to $\hat{\kappa}_1$ and $\hat{\kappa}_2$ estimates. Evaluated at the average rainy intensity, the column-(7) intensity slopes $\hat{\kappa}_1$ and $\hat{\kappa}_2$ imply HPU reductions of 0.458% (SE 0.146%) and 0.842% (SE 0.324%); the discrete effect of rain onset is captured separately by the rain indicator.[12] On the extensive margin, the probability of any HPU falls by 0.189% (SE 0.06%) and 0.271% (SE 0.103%) (column 3).[13] The trip-average specifications (columns 4 and 8) collapse the two κ terms into one and give estimates almost identical to the across-trip $\hat{\kappa}_1$. The remaining columns move as expected: the effect is negligible and imprecise before the trip-level controls enter (columns 1 and 5) and sharpens once the controls (columns 2 and 6) and driver fixed effects do. Online Appendix Table D6 reports the control coefficients for columns (4) and (8).

Heterogeneity and sampling Splitting the sample by baseline phone use shows the response to precipitation is concentrated among the lighter users, while the heavy users barely respond; we revisit this gradient in Section 5, where a single exhibit (Table 6) sets it beside the deductible and experimental evidence. The same gradient shapes how weighting matters: reweighting to fleet-representative HPU (Online Appendix Table D4) up-weights low-HPU drivers, raising the within-driver response to about 6.2 times the headline—yet even this elevated estimate implies a small $\hat{\beta} \approx -0.034$. We therefore report the unweighted sample as our headline, reading it as the response among the HPU-active drivers where the behavior and most of its externality concentrate rather than as a fleet average (Solon et al. 2015).

Robustness We defer discussion of why the estimated moral-hazard response is small to the end of the section, where it is considered jointly with the deductible evidence; here we address only explanations specific to the precipitation estimate: measurement error in precipitation or HPU, or a competing channel through which rain raises the demand for phone use. First, to verify that the pattern is not an artifact of trip-level aggregation, Online Appendix Table D5 replicates the analysis on the accident-driver interval-level

[12] $\overline{Inch} = 0.19$ is the rainy-trip mean of the rain-intensity variable, corresponding to a typical rainy intensity of 0.2034 in/hr. For a given $\hat{\kappa}$, the percentage change is $\%\Delta\text{HPU} = [\exp(\hat{\kappa}\,\overline{Inch}) - 1]$, and the standard error is $\text{SE}(\hat{\kappa})\,\overline{Inch}\exp(\hat{\kappa}\,\overline{Inch})$.

[13] Similarly, for a given $\hat{\kappa}$, the extensive-margin percentage change is $\%\Delta\text{HPU} = \hat{\kappa}\,\bar{X}/\bar{p}$, where $\bar{p}$ is the dependent variable mean of 0.666, and the standard error is $\text{SE}(\hat{\kappa})\,\bar{X}/\bar{p}$.

Table 3: Precipitation Effects on HPU: Extensive Margin and Poisson QMLE on $E[\text{HPU}]$

	Extensive: P(HPU >0) (LPM)				QMLE: $\log E[\text{HPU}]$ (Poisson)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Rain indicator	-0.0004	-0.0001	-0.0001	-0.0001	-0.0154^{**}	-0.0135^{**}	-0.0024	-0.0023
	(0.0015)	(0.0015)	(0.0011)	(0.0011)	(0.0064)	(0.0063)	(0.0037)	(0.0037)
Rain $\times$ start intensity	-0.0040	-0.0063^{**}	-0.0068^{***}		-0.0059	-0.0187	-0.0248^{***}	
	(0.0028)	(0.0028)	(0.0021)		(0.0119)	(0.0120)	(0.0079)	
Rain $\times$ Δ intensity	-0.0100^{**}	-0.0108^{**}	-0.0098^{**}		-0.0247	-0.0367	-0.0457^{***}	
	(0.0045)	(0.0044)	(0.0037)		(0.0237)	(0.0229)	(0.0176)	
Rain $\times$ avg intensity				-0.0065^{***}				-0.0237^{***}
				(0.0021)				(0.0078)
$\%\Delta$ HPU at mean rain	-0.110	-0.174^{**}	-0.189^{***}	-0.181^{***}	-0.109	-0.346	-0.458^{***}	-0.435^{***}
	(0.079)	(0.077)	(0.060)	(0.059)	(0.221)	(0.222)	(0.146)	(0.143)
Location + Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Trip-level controls (15)	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Driver FE	No	No	Yes	Yes	No	No	Yes	Yes
Observations	1,689,022	1,689,022	1,687,611	1,687,611	1,677,339	1,677,339	1,675,350	1,675,350

Notes: $^{*}p<0.10$; $^{**}p<0.05$; $^{***}p<0.01$. Standard errors (in parentheses) are two-way clustered by driver and date (Abadie et al. 2023; Cameron et al. 2011): the driver is the repeatedly-sampled unit and rain is a spatially and serially correlated day-level shock. Cols (1)–(4): LPM on $\mathbf{1}\{\text{HPU}>0\}$ (baseline 0.666), coefficients in probability units per unit of rain intensity; cols (5)–(8): Poisson QMLE on HPU intensity with a log-mileage exposure offset (Chen and Roth 2024), coefficients in log-points per unit of rain intensity. Cols (1)–(3) and (5)–(7) decompose the rain intensity into its trip-start level and within-trip change (Δ); cols (4) and (8) use trip-average intensity. **Columns (3) and (7)**, the driver-fixed-effects specifications with the trip-start/within-trip decomposition, are our preferred specifications (shaded). *Trip-level controls* (*15*) is the harmonized set defined in Section 4.1; *Location + Time FE* is the additive location-pair (origin–destination geohash), date, and hour-of-day effects. $\%\Delta$ *HPU at mean rain* evaluates the trip-start (or, in cols 4 and 8, avg-intensity) coefficient at the rainy-trip mean intensity ($\overline{Inch} \approx 0.19$, a typical rainy rate of ≈ 0.203 in/hr) with $\Delta = 0$: LPM, $100\,\hat{\kappa}\,\overline{Inch}$/base; Poisson, $100(\exp(\hat{\kappa}\,\overline{Inch})-1)$; SE delta-method-propagated.

panel (Dataset 1), which matches the minutely radar precipitation feed to five-second driving intervals. With driver fixed effects, precipitation lowers the probability of HPU in a given five-second interval only slightly: evaluated at the same rainy intensity as the trip-level analysis (0.2034 in/hr), the implied $\%\Delta$HPU is -0.621% (SE 0.144%) in the linear specification, statistically distinguishable from zero and in line with the trip-level magnitudes; the conditional-logit counterpart is of similar size but imprecise (Online Appendix Table D5). The fine-grained panel corroborates the trip-level response.

Two further explanations concern the HPU measure itself and the demand for phone use. Precipitation might degrade the HPU classifier, but as we described in Section 2.1, its detection logic and insensitivity to environmental factors make this unlikely. Precipitation might also raise the private benefit of HPU, for instance the value of communication for delay management or of in-trip rerouting; our main specification controls for in-app communication directly (per-mile calls and texts from the rider, platform notifications, and added rides; Online Appendix Table D6). For such a demand channel to mask a

moral-hazard response it would have to offset it almost exactly and simultaneously across the extensive and Poisson margins, the trip-start and within-trip precipitation profiles, and the saturations of Table 3, which we regard as implausible.

4.2 Behavioral Response to a Deductible Hike: A Natural Experiment

As discussed in Section 2.1, the platform raised its commercial collision deductible from \$1,000 to \$2,500 on March 1, 2021, a one-time policy change that increased the dollar amount drivers pay out-of-pocket on every collision claim. Moral hazard implies that HPU should fall, with the response scaling with the driver's expected exposure to collision claims.

Although every driver faces the same \$1,500 deductible step, its expected out-of-pocket consequence varies with the local claim environment. We measure this exposure directly from the insurer's pricing system, which provides the 2021 actuarially-fair premium at the \$500 and \$1,000 deductibles, $P_c(500)$ and $P_c(1000)$, for every county in our estimation sample, which represents the company's expected collision loss given those deductibles. Because the insurer's expected cost at deductible d is $\lambda\, E[(s-d)^+]$ for claim frequency λ and loss draw s, the two deductible points identify the county claim-cost structure: under an exponential loss distribution the inversion is analytic, and the hike's dose is the driver's annual expected out-of-pocket increase,

$$\Delta\mathrm{EOOP}_c \;=\; \mathrm{EOOP}_c(2500) - \mathrm{EOOP}_c(1000) \;=\; P_c(1000)\Big(1 - \big(P_c(1000)/P_c(500)\big)^3\Big),$$

averaging \$62 per year across drivers.[14] Online Appendix Figure D1 maps the dose.

We follow the standard dose-response difference-in-differences design (Callaway et al. 2024; Card 1992), fit on a balanced driver-day panel by the same two estimators as the precipitation panel:

$$g(\mathbb{E}[\,b_{id} \mid \cdot\,]) \;=\; \alpha_i + \tau_{s(i),\,d} + \theta\,\mathrm{Post}_d{\cdot}\Delta\mathrm{EOOP}_{c(i)} + x'_{id}\psi, \tag{3}$$

where $g(\cdot)$ is the estimator's link and b_{id} the HPU outcome, both taking the same two forms as in the precipitation panel (Equation 2): an extensive-margin linear probability model on $\mathbf{1}\{\mathrm{HPU}_{id} > 0\}$ with g the identity, and a Poisson QMLE on HPU intensity with g the log link and a log-mileage exposure offset. α_i is a driver fixed effect (absorbing time-invariant driver characteristics, including the level of the dose itself), $\tau_{s(i),\,d}$ a time effect, and x_{id} the

[14] Writing $r_c = P_c(1000)/P_c(500) = e^{-500\zeta_c}$ pins the exponential rate ζ_c and, from the level, the frequency λ_c; expected out-of-pocket cost at deductible d is $\lambda_c(1-e^{-\zeta_c d})/\zeta_c$. The construction uses only the two pricing points—no claims or severity data. Appendix C gives the derivation; Online Appendix Table D7 shows the results are robust to a Pareto severity family and to the 2023 pricing vintage.

day's passenger-carrying share; the extensive-margin LPM additionally conditions on log mileage, while the Poisson carries mileage as its exposure offset. The Post $\times$ ΔEOOP slope θ is the dose-response parameter—the log-rate change in HPU intensity per dollar per year of expected out-of-pocket increase, post-policy (Table 4 scales the dose in \$100/yr units).

We report two fixed-effect saturations as main specifications, differing only in the time effect. With driver fixed effects and a plain date effect τ_d (columns 2 and 5 of Table 4), θ is identified from both across- and within-state differences in the predetermined dose. With a state-by-date effect $\tau_{s(i),d}$ (columns 3 and 6), which additionally absorbs every state-specific time shock over the window, θ is identified from within-state variation alone. Standard errors are clustered on the county of the driver's modal location, the level at which the dose is constructed (a county-grain measure; 63% of its variance lies between states); driver-clustered SEs, which are smaller on the Poisson margin, are reported for comparison in the note to Table 4.

The event window is bracketed by two sets of concurrent shocks: Winter Storm Uri struck in mid-February, just before the window opens, with its recovery running into the pre-period, and several states changed COVID-era policies in early March, inside the post-period. To avoid contaminating the dose-response with these shocks, both specifications use a window running from 10 days before March 1, 2021 through the end of the panel on March 9, and drop twelve affected states (TX, OK, LA, AR, MS, AL, TN, WY, MT, AZ, FL, SC; the note to Online Appendix Figure D1 details the selection criteria). We drop these twelve states because their shocks sit inside the window—the storm states' recovery occupies the pre-period, the policy states' rollouts the post-period—so no window trim can rescue them: re-adding them fails the design's pre-trend checks (Online Appendix Table D7).

Results. Table 4 reports our main results. Two outcome panels—an extensive-margin LPM (columns 1–3) and a Poisson QMLE on HPU intensity with a log-mileage exposure offset (columns 4–6)—across three fixed-effects saturations: date effects only (columns 1 and 4); +driver fixed effects (columns 2 and 5), which identify θ from cross-driver differences in the predetermined dose, including across-state differences; and +state-by-date fixed effects (columns 3 and 6), which identify it from within-state differences alone. Within-state variation in expected out-of-pocket exposure produces a precise null: per dollar per year of expected-cost increase, $\%\Delta\text{HPU} = 0.0321\%$ (SE 0.0311) on the Poisson intensity margin and 0.0032% (SE 0.0126) on the extensive margin. Adding across-state variation produces a small negative response that is statistically significant only on the extensive margin: with driver and date effects (columns 2 and 5) the extensive margin

Table 4: HPU Response to the March 2021 Collision-Deductible Increase

	Extensive: $\Pr(\text{HPU} > 0)$ (LPM)			QMLE: $\log E[\text{HPU}]$ (Poisson)		
	(1)	(2)	(3)	(4)	(5)	(6)
ΔEOOP (\$100/yr)	-0.01916	—	—	-0.01618	—	—
	(0.02385)			(0.27239)		
Post $\times$ ΔEOOP	-0.016794^{***}	-0.012874^{**}	$+0.002818$	-0.06598	-0.05611	$+0.03159$
	(0.006274)	(0.005971)	(0.011094)	(0.05421)	(0.05379)	(0.03012)
$\%\Delta$ HPU per \$100/yr (k)	-1.91^{***}	-1.46^{**}	$+0.32$	-6.38	-5.46	$+3.21$
	(0.71)	(0.68)	(1.26)	(5.08)	(5.09)	(3.11)
Driver-day controls	Yes	Yes	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	-	Yes	Yes	-
State $\times$ Date FE	No	No	Yes	No	No	Yes
Driver FE	No	Yes	Yes	No	Yes	Yes
Pre-trend F (p)	3.36 (0.00)	1.72 (0.08)	1.01 (0.43)	0.71 (0.70)	0.63 (0.77)	0.71 (0.70)
N (driver-day $\times$ phase obs.)	284,370 ($N_{\text{drv}} = 12{,}881$, $N_{\text{cty}} = 466$, $N_{\text{st}} = 39$)					

Notes: Standard errors clustered on the driver's modal county in parentheses; $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. Observations are driver-day $\pm$10-day panel around March 1, 2021 with the twelve excluded states removed (Section 4.2). Columns 1–3 are an extensive-margin LPM on $\mathbf{1}\{\text{HPU} > 0\}$ (baseline 0.881); columns 4–6 a Poisson QMLE on HPU intensity with a log-mileage offset. Columns 2 and 5 (Driver FE) identify the dose-response from cross-driver variation, including across-state; columns 3 and 6 (State$\times$Date FE) from within-state variation alone; the four identified columns are shaded. All columns include the driver-day controls (passenger-phase share; the LPM also conditions on log mileage). ΔEOOP is the county expected annual out-of-pocket increase caused by the hike, derived from the insurer's expected collision costs at the \$500 and \$1,000 deductibles under an exponential loss family (Section 4.2), in \$100-per-year units, so the interaction coefficient over the baseline rate (LPM) and directly (Poisson) reads as the $\%\Delta$HPU per \$100 per year of expected-cost increase (the text quotes the same semi-elasticity per dollar per year, $k/100$). $\%\Delta$HPU is the semi-elasticity k: $100\,\hat{\theta}/\bar{p}$ (LPM) and $100(e^{\hat{\theta}} - 1)$ (Poisson), delta-method SEs; multiplying k by the sample-average dose (\$62/yr $= 0.62$ hundred) gives the average effect of the hike. *Pre-trend F* jointly tests the pre-period event-study $\times$ dose coefficients (county-clustered). Window, sample, severity-family, pricing-vintage, clustering, and labor-supply robustness in Online Appendix Table D7.

falls by 0.0146% per dollar per year (SE 0.0068)—a -0.91% decline at the sample-average \$62-per-year increase—while the Poisson estimate (-0.0546% per dollar, SE 0.0509) is indistinguishable from zero. The remaining date-only columns (1 and 4) omit the driver fixed effect; their cross-sectional dose level is small and statistically indistinguishable from zero, and the driver fixed effect absorbs it in the identified columns.

Heterogeneity. The across-state reduction masks the same gradient as the rain evidence: it is carried by the lighter phone users and fades in the heavy tail. As with precipitation, we revisit the decomposition and set it beside the experimental evidence in Table 6 (Section 5).

Robustness. Online Appendix D.3.2 (Online Appendix Table D7) subjects both specifications to a battery of variations: county-, driver-, and state-clustered inference; re-adding the storm states and narrowing the window; rebuilding the dose under a Pareto severity family and under the 2023 pricing vintage; and an alternative labor-supply outcome. Across these variations the within-state estimates remain small and statistically indistinguishable

from zero, and the across-state extensive-margin reduction recurs at a similar magnitude across most variants. Re-adding the twelve storm states fails the design's pre-trend checks (Online Appendix Table D7). The labor-supply outcome shows no offsetting change in driving hours conditional on driving, so the HPU-intensity measure is not distorted by a mileage-denominator response.

4.3 How big is β?

Dividing each reduced-form $\%\Delta$HPU by a calibrated change in HPU's expected marginal accident cost ($\%\Delta$MC) turns it into a unit-free moral-hazard elasticity $\hat{\beta}$ (Equation A3). The precipitation $\%\Delta$MC is already estimated from our hazard model (Table 2), where HPU's marginal accident cost increases by 34.71% from dry to Dataset 2's typical rainy conditions, so numerator and denominator share the same rain contrast. The deductible $\%\Delta$MC is measured from the same pricing points as the dose: the driver's expected out-of-pocket increase over the base expected out-of-pocket cost at the \$1,000 deductible, averaged over drivers, gives $\%\Delta\text{MC} = 84.2\%$, so the deductible $\hat{\beta}$ is an elasticity with respect to out-of-pocket cost. Appendix C gives the derivation and the sensitivity band spanning the severity family and pricing vintage.

The precipitation response implies an ex-ante moral-hazard elasticity of $\hat{\beta} \approx -0.005$, with all four precipitation rows statistically significant and spanning $|\hat{\beta}| \approx 0.005$–$0.032$ (Appendix Table A2). The deductible hike implies a semi-elasticity of -0.0146% per dollar per year of expected out-of-pocket cost on the extensive margin across states ($|\hat{\beta}| \approx 0.011$), and a response indistinguishable from zero within state or inclusive of the intensive margin. The claims-based moral-hazard literature's estimates range from precise nulls (Abbring et al. 2003; Chiappori and Salanie 2000) to $|\beta| \approx 0.1$–1 (Jeziorski et al. 2017; Weisburd 2015); Appendix Table A1 places our precipitation and deductible estimates beside that literature study by study.

Discussion Our elasticities measure ex-ante moral hazard alone, a different object from what the claims-based moral-hazard literature identifies: the numerator is the driving behavior itself, not the claims through which prior work must infer it. The difference cuts both ways. Claim filing and resolution respond to insurance and tort incentives—especially after an accident, when the stakes are salient—and miles driven is rarely observed or adjusted in claims-based studies, so estimates that detect moral hazard mix the ex-ante driving response with ex-post filing and exposure. Claims are also rare, so a response of our size is observationally a zero in the tests that detect none. A small ex-ante, purely behavioral, component is consistent with both halves of the literature—and it is the one

deterrence rests on: laws deter by changing how people drive, not how claims are filed.

Why is the behavioral response itself so small? Previewing our conclusion, we read it as neglect. Two features of our setting make neglect especially plausible here. First, both cost shifts we study—precipitation and the deductible hike—are risk changes drivers face *passively*. Some insurance settings share this feature, such as an automatic experience-rating premium increase; others hinge on an *active* choice—opting into monitoring, or selecting a lower coverage level—which is far more salient, so our estimates may understate the response in settings where drivers engage the price directly.

Second, and specific to the deductible change, drivers may not have internalized it as a within-trip risk signal, even though its dollar magnitude is large. A well-developed literature on health-insurance choice documents that consumers frequently misunderstand contract details and under-respond to changes in their financial exposure (Handel and Schwartzstein 2018). In our case, the deductible change reached drivers by email; many plausibly never opened or fully processed it, so the treatment is better read as the availability of a higher deductible than as a change every driver consciously registered.

We do not read these features as a design flaw: that drivers barely register these passively-imposed, low-salience changes is itself informative about how moral hazard operates in everyday driving, and our friction-inclusive estimate—rather than ones measured in salient and isolated choice environments—is the policy-relevant one for the liability and tort regulations discussed at the outset (Section 1).

Alternatively, the small response may have nothing to do with underperception or inattention. Our HPU measure may be too coarse to detect a small change, or handheld use may simply be worth too much to give up even when drivers perceive an increase in its cost. The next section's experiment tests these alternatives: a one-time text message alters salience alone, at no cost to the driver, and HPU falls sharply and persistently, ruling out the measurement and preference readings.

5 Experimental Evidence

A natural question following Section 4 is whether the small moral-hazard response is in fact an artifact of HPU being unable (or too costly) to move, or being too coarsely measured in our setting to register movement. To address this, we analyze a simple one-time text-message intervention in February 2020.

5.1 Experimental Design

Treatment. Treated drivers received a one-time text message that reads:
"*Our app shows you may be holding your phone while driving. Passenger reports of unsafe driving, like handheld phone use, can lead to suspension.*" (Online Appendix Figure D2)

The intervention is designed to increase the salience of the accident and suspension risks associated with HPU, similar to salience and feedback experiments in driving safety (Delgado et al. 2024; Ebert et al. 2024). Because the message changes no prices or payoffs—only the salience of risks drivers already face—we refer to it as a *salience cue* when contrasting it with the everyday cost variation of Section 4. The message links to an FAQ page that provides legal context, explains how telematics data are collected, and describes the suspension process, including procedures for verifying passenger safety complaints.

Among drivers in Dataset 2, the intervention targeted $2{,}503$ regularly-active, high-HPU drivers—those with more than 20 hours per week of drive time and average HPU intensity in the top quintile—measured between January 28 and February 10, 2020.

On February 11, 2020 the intervention was implemented during a 30-minute window (5:00–5:30 pm EST), chosen for operational convenience. Among the targeted drivers who logged on and started a shift during this window, a randomly selected 50% were sent the text message immediately upon log-on, with the rest held out as controls.

The message was delivered to 343 treated drivers—45.7% of the targeted drivers who logged on during the window, and 13.7% of all targeted drivers. Our main difference-in-differences compares the treated drivers to the full set of $2{,}160$ untreated targeted drivers; the 407 within-window randomized controls form the strict experimental control used as a robustness check (Online Appendix Table D11). This sample contains $464{,}045$ trips and $849{,}488$ trip-segments.

Balance Online Appendix Table D12 reports pre-treatment balance against both the randomized control group (drivers who logged on in the experimental window and were randomized to control) and the full untreated control group used in the difference-in-differences. The pre-determined characteristics—age, gender, and the platform-recorded histories of accidents, unsafe rides, safety tickets, rides completed, and earnings—are well balanced against both control groups. Two characteristics differ. Treated drivers enter with lower pre-period HPU frequency than either control group (roughly a fifth below the control means, significant at the 5% level or better), and with about 0.2 years less platform tenure ($p < 0.05$ against both groups). Every other characteristic, including the harsh-event rates, is statistically indistinguishable at the 5% level. Our main specifications thus include driver fixed effects to absorb time-invariant level differences.

5.2 Average Treatment Effects

We estimate the treatment effect with the same specification as our observational tests (Equation 2), now as a difference-in-differences that compares the treated drivers against the full set of untreated targeted drivers before and after the intervention:

$$g(\mathbb{E}[\, b_{it} \mid \cdot\,]) = \theta\, D_{it} + x'_{it}\psi + \alpha_i + \phi_{\ell(it)} + \tau_{t(it)} + \eta_{h(it)}, \tag{4}$$

The outcome b_{it}, link $g(\cdot)$, and controls x_{it} follow Equation 2—both the extensive-margin LPM and the intensity Poisson QMLE—and α_i, $\phi_{\ell(it)}$, $\tau_{t(it)}$, $\eta_{h(it)}$ are driver, geohash-5 location-pair, date, and hour-of-day fixed effects, added progressively across the columns of Table 5 (the location-pair effects, as in Equation 2, enter the fullest column). The new object is the treatment indicator $D_{it} = \mathrm{Treat}_i \cdot \mathrm{Post}_t$, one for a treated driver after the intervention; the driver and date effects absorb treatment status and the common post-intervention shift, so θ is the average treatment effect on the treated. Standard errors are clustered at the driver level.[15] Restricting the comparison to only the within-window randomized control group yields a consistent estimate (Online Appendix Table D11). Table 5 reports the average treatment effect.

The treatment effect is large. Table 5 reports a 19.1% reduction in handheld phone use (column 3), significant at the 1% level and stable across the progression of controls and fixed effects (columns 1–4). Adding the geohash-5 location-pair fixed effects of Equation 2 (columns 4 and 8) leaves both the point estimate and the pre-trend F essentially unchanged, so we take the specification without them—columns 3 and 7—as our main spec.

It is also persistent. Figure 1 plots the weekly-binned event-study coefficients $\{\theta_k\}$, and the drop appears in the first post-treatment week and holds. The grey dotted line marks the onset of the pandemic-declaration week (event day 28); the data extract continues through March 30, 2020, but the estimation sample for Table 5 ends at event day 33 (March 16, 2020), as the first shelter-in-place orders were being announced: matched-pool trip volume collapses thereafter (roughly $2{,}400$ trips per day, against $8{,}700$ in the pre-period).

This result bears on the two alternative readings of the observational nulls of Section 4. HPU is movable: a one-time salience cue produces a persistent, easily measurable response, which weighs against the "HPU cannot be moved" reading (whether a hard behavioral

[15] Driver fixed effects make this a within-driver (block-design) comparison, so θ up-weights drivers observed in both periods. Because the treatment does not move driving activity—hours driven are a precise null (Online Appendix Table D8), and daily trip counts and the probability of driving at all are likewise flat (Online Appendix D.4), this weighting is not confounded by treatment-induced compositional change.

Table 5: Experiment: Effect on HPU Intensity (Poisson QMLE)

	Average effect (trip)				Passenger targeting (trip segment)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post × Treated	−0.2098**	−0.2095**	−0.2114***	−0.1956***	−0.1252*	−0.1242*	−0.1328***	−0.1224**
	(0.0967)	(0.0960)	(0.0542)	(0.0525)	(0.0697)	(0.0697)	(0.0471)	(0.0547)
× Passenger					−0.2000**	−0.1994**	−0.1865*	−0.1982*
					(0.0913)	(0.0911)	(0.0978)	(0.1044)
Trip-segment controls	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Driver + Date + Hour FE	No	No	Yes	Yes	No	No	Yes	Yes
Location-pair FE	No	No	No	Yes	No	No	No	Yes
Pre-trend F (p)	1.46 (0.09)	1.46 (0.09)	1.06 (0.39)	1.06 (0.39)	1.11 (0.33)	1.14 (0.31)	1.17 (0.27)	1.17 (0.27)
Observations	464,045 trips				849,488 segments			

Notes: $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$. Driver-clustered SEs. Dependent variable: handheld-phone miles per mile (Poisson QMLE, log-mileage offset; semi-elasticities). A trip is the aggregation of its trip segments (pickup-driving P2 and passenger-aboard P3). Columns (1)–(4) are the TRIP-level average effect, adding controls, then driver+date+hour fixed effects (column 3, the canonical specification), then geohash-5 location-pair fixed effects (column 4, matching the precipitation analysis). Columns (5)–(8) are at the trip-segment level, with the outcome measured as handheld-phone seconds per driving-second (Poisson QMLE, log-driving-time offset), and add the Post×Treated×Passenger interaction (first row: the pickup-segment (P2) effect; ×Passenger row: the passenger-segment (P3) differential), mirroring the same fixed-effects progression. Shaded columns are the main specifications (columns 3 and 7). The pre-trend row reports the joint Wald F (p-value in parentheses) on the pre-period treatment leads (with weekday in place of date fixed effects, collinear with event time). Poisson estimation samples fall slightly below the panel (column 3: 463,872 trips; column 7: 847,885 segments) and, with location-pair fixed effects, further exclude pairs with no handheld use in the sample (column 4: 387,983 trips; column 8: 724,632 segments); such cells are fit perfectly by the fixed effect and carry no identifying variation. Poisson QMLE on handheld-phone miles with a log-mileage offset (Chen and Roth 2024); coefficients are semi-elasticities.

constraint or a private benefit so high that no realistic incentive could overcome it). Our measure also registers the change: the post-intervention reduction is picked up by the same sensor pipeline used in the observational regressions, so coarse measurement cannot account for the observational nulls.

5.3 Treatment Effect Heterogeneity

Why is the response to everyday incentives so muted when a one-time cue moves HPU so sharply? Our heterogeneity analyses are suggestive about both the channel behind the response and the drivers who account for it. First, the response is only partially targeted to the suspension threat the message names. Second, the cue moves the heaviest phone users most, while the everyday incentives of Section 4 move only the lighter, low-use majority; the contrast points to a neglect channel.

Does the response target the message's specific threat? The message warns specifically of *passenger-complaint-driven suspension*, so a precise reaction would concentrate on passenger-bearing driving. We split each trip into its passenger-aboard and pickup (no-passenger) segments, defined in Section 2, yielding $849{,}488$ segments across the $464{,}045$ trips; the

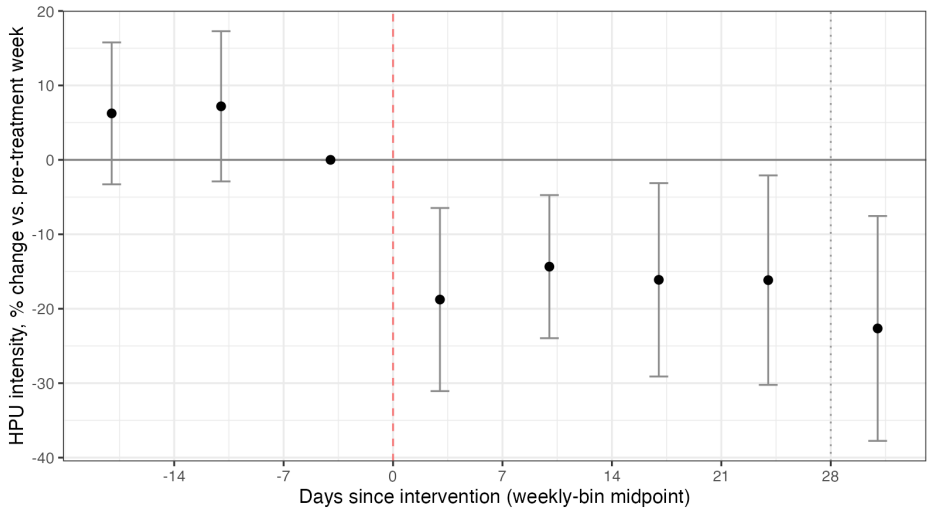

Figure 1: Direct Intervention: Weekly Treatment Effect on HPU

Notes: Weekly treatment coefficients (θ_k, the event-time generalization of θ in Equation 4, event days grouped into 7-day bins) from the Poisson QMLE event study on HPU intensity (handheld phone use intensity with a log-mileage exposure offset; the same trip-level Poisson estimator as Table 5, with weekday rather than calendar-date fixed effects, since date effects would absorb event time), so each coefficient is a semi-elasticity ($\approx$ percent change in HPU intensity relative to the pre-treatment week). Standard errors are clustered at the driver level; bars are 95% confidence intervals. A joint test of the 2 pre-period bins does not reject parallel pre-trends ($F = 1.10$, $p = 0.333$). The red dashed line marks the intervention date; the grey dotted line marks the onset of the pandemic-declaration week (event day 28). The pre-period reference is the 7-day bin ending at day -1. Table 5 reports formal DiD estimates; Online Appendix Table D8 reports robustness across outcomes and specifications; the day-by-day event study is in Online Appendix Figure D3.

segment regressions carry the same driver-day controls as the trip panel.[16] The targeting test is the Post $\times$ Treated $\times$ Passenger interaction in columns (5)–(8) of Table 5; as in the average-effect columns, column (7) (driver + date + hour) is the main specification and adding location-pair fixed effects in column (8) leaves the differential essentially unchanged. HPU falls significantly on the *pickup* segments, where no passenger is aboard and the threatened complaint cannot arise, by about -0.13 (SE 0.05; $p < 0.01$). The cut is somewhat larger on passenger-bearing segments, a differential of about -0.19 (SE 0.10), but this differential is imprecise (significant only at the 10% level), so we do not lean on it. The response is thus only partially targeted: drivers curtail phone use even on pickup segments, where the named complaint channel cannot operate.

The response is also specific to HPU. Online Appendix Table D8 returns precise nulls on harsh acceleration, braking, and cornering, despite the message warning broadly about "unsafe driving." Drivers' response again tracks the behavior the message names rather than the suspension incentive it also invokes.

[16] Covariates undefined within a segment are dropped: the passenger share (which is the split itself) and trip mileage (absorbed by the log-driving-time offset).

Two features of the intervention discipline its interpretation. First, the enforcement threat the message names never materialized: no driver in our sample was suspended during the research window, so the response operates through the warning alone. Nor can we test the crash margin directly: Dataset 2 carries no accident labels, and at observed crash rates a sample of this size would be underpowered to detect the implied change in accidents; the experiment establishes that HPU is movable and measurable, not the downstream crash reduction. Second, drivers could in principle reduce detected rather than actual handheld use, by mounting the phone, switching to hands-free operation, or shifting to a second device. The daily event study (Online Appendix Figure D3) makes gradual evasion an unlikely account, since the reduction appears immediately upon receipt. And each of these substitutes (a mount, hands-free operation, or in the worst case a separate phone that does not also handle navigation and platform communication) is itself a safer alternative to handheld manipulation, so the measured decline is a safety improvement under any of these channels.

Table 6: Who responds: the salience cue versus everyday incentives

	Precipitation		Deductible		Experiment	
	$\Pr(\text{HPU}>0)$	$\log E[\text{HPU}]$	$\Pr(\text{HPU}>0)$	$\log E[\text{HPU}]$	$\Pr(\text{HPU}>0)$	$\log E[\text{HPU}]$
Lighter users						
HPU $\leq$100 m/km	-0.23^{***}	-0.66^{***}	-1.29^{***}	-6.28	-4.6^{*}	-17.8^{***}
	(0.07)	(0.23)	(0.47)	(4.39)	(2.5)	(5.2)
Bottom 80%	-0.24^{***}	-0.65^{***}	-1.17^{**}	-6.29	-4.2	-17.5^{***}
	(0.08)	(0.24)	(0.47)	(4.27)	(2.6)	(5.3)
Heavier users						
HPU >100 m/km	-0.03	-0.23	0.68	-0.67	-5.5^{*}	-20.6^{***}
	(0.06)	(0.22)	(0.81)	(3.24)	(3.1)	(6.9)
Top 20%	-0.03	-0.29	0.29	-0.02	-6.5^{**}	-20.9^{***}
	(0.05)	(0.20)	(0.91)	(3.35)	(3.1)	(6.7)

Notes: $^{*}p<0.10$, $^{**}p<0.05$, $^{***}p<0.01$. Each cell is the $\%\Delta$HPU from re-estimating the setting's own specification on the driver subsample named in the row: the precipitation regression (Equation 2, avg-intensity form, at the mean rainy-trip intensity), the across-state deductible dose-response (Equation 3, driver and date fixed effects, at the mean ΔEOOP dose), and the text-message difference-in-differences (Equation 4). $\Pr(\text{HPU}>0)$ columns are the linear-probability $\%\Delta$; $\log E[\text{HPU}]$ columns the Poisson QMLE $\%\Delta$ in HPU intensity, as in Table 3. Standard errors (in parentheses) are clustered as in each parent analysis (precipitation: driver$\times$date; deductible: county; experiment: driver). Drivers are split on *predetermined* HPU intensity—the driver's pre-period mean HPU intensity—two ways: an intensity threshold at 100 m/km (handheld use on $\approx$10% of miles driven) and a percentile split at the 80th percentile.

Who responds: the salience cue versus everyday incentives. Table 6 splits each setting's response by drivers' predetermined HPU intensity and exposes a sharp contrast between the everyday incentives of Section 4 and the experimental cue. The two natural incentives

load on the *lighter* users: for both the precipitation shock and the deductible hike, the any-use reduction is significant among the lighter 80% of drivers (and among those with handheld use on under 10% of their miles) and is indistinguishable from zero in the heavy tail. The experimental cue runs the opposite way: its intensity reduction is larger among the *heaviest* users in point estimate, and unlike the incentive responses it does not fade there. The average cue effect carries the same tilt—the volume-weighting Poisson semi-elasticity (about 19.1%) exceeds the equally-weighted extensive reduction because the cut is proportionally largest among the heaviest users, who account for most aggregate exposure. The same inattention reconciles both gradients: drivers who already self-regulate against everyday risk have little left for a message to change, while heavy users, who find the risk less salient, respond to the message rather than to price. This intervention-side gradient echoes telematics-feedback evidence that monitoring improves the worst-performing drivers most (Soleymanian et al. 2019). This also bears on why heavy users barely respond to incentives. One account is a belief about one's own efficacy, since a driver who thinks attention does little to change his crash risk would both under-prevent and under-respond to its price (Spinnewijn 2015). The cue makes that account incomplete on its own, because a costless message that changes no payoff still moves heavy users sharply, which fixed pessimism about control would not predict; we read the contrast as suggestive of inattention rather than an inability to self-regulate, while stressing that the two are not cleanly separable here and we do not claim to identify the mechanism. The gradients are stronger on some margins than others—the precipitation contrast is sharpest on the extensive margin, the experiment on the intensity margin—and the deductible gradient likewise concentrates on the extensive margin (its intensity cells are imprecise under Section 4.2's county-clustered inference).

Comparison with similar interventions and takeaway. Our intervention is not designed to calibrate a specific moral-hazard elasticity under engineered incentives; our results align with recent field experiments, which have further established that the size of a financial stake does not meaningfully alter HPU beyond the salient signal alone.[17] Our 19.1% reduction sits within the range these trials report.

What we add, by pairing the intervention with the observational tests in the same setting, is a field baseline. The same phone use that a costless message moves sharply

[17]Two recent RCTs put the effect close to ours and locate the lever. Ebert et al. (2024) find a 20.5% HPU reduction from a salience-and-feedback gamification arm but none from a phone mount, commitment, or habit tips, and adding cash prizes did not improve on feedback alone; Delgado et al. (2024) find a 14.5–21.2% reduction only when feedback is paired with an incentive—largest for a loss-framed stake, with doubling the amount adding nothing—concluding that incentives' "effectiveness may depend less on the amount of money and more on how it is delivered."

barely responds to the everyday variation in its expected cost studied in Section 4; HPU is thus readily movable, so the muted observational responses are not a behavioral floor but the other side of the "salience" effect this literature identifies—behavior tracks a salient, well-delivered signal rather than the size of the cost (Delgado et al. 2024; Ebert et al. 2024)—and the everyday cost, present but not salient, goes largely neglected. That said, this salience-and-neglect reading is not a single structural mechanism but a family of them: neither our design nor the existing evidence identifies which generates the neglect—a misperception of the risk (under-perception or overconfidence, whether informational or preferential), or the risk simply not entering the choice (inattention or a limited consideration set, which may be rational—through the pull of a default, decision or adjustment costs, or memory limits—or not). We take up these directions in the conclusion.

6 Conclusion

Do drivers become safer when they bear a higher accident cost? This question sits, as the introduction laid out, at the center of a legal and regulatory debate over liability and tort (Calabresi 2008; Shavell 1987; Wagner 2006). On the premise that drivers do, economists have proposed usage-based insurance pricing that would "bring home to the automobile user the costs he imposes in a manner that will appropriately influence his decisions" (Vickrey 1968), and have cautioned that safety mandates may be substantially offset, as drivers who bear less of the risk drive more intensely (Peltzman 1975). However, this premise has been conceptually challenged by competing theories of decision-making under uncertainty, in which risk can be elevated (Barseghyan et al. 2013; Sydnor 2010) or neglected (Hertwig et al. 2004; Kunreuther and Pauly 2004). In everyday driving, our observational evidence, and its contrast with the intervention, points to the neglect pole. Heavy phone users, who account for most phone use, are readily moved by a salient cue yet unresponsive to natural variation in the cost of risk.

More broadly, whether a risk incentive is salient offers a useful lens for parsing the empirical literature (Table A1). The deterrence effect is strongest in feedback-and-incentive trials (Delgado et al. 2024; Ebert et al. 2024) and monitoring programs with active feedback (Jin and Vasserman 2021; Soleymanian et al. 2019), which produce robust reductions of 15–30% in the monitored behavior, similar to our experimental intervention (Panel B). Traditional auto-insurance and tort incentives are weaker and harder to measure, confounded by selection and claim-reporting margins (Panel A): studies that detect an ex-ante response split between clean causal estimates that isolate it, putting $|\beta| \approx 0.1$–1 (Jeziorski

et al. 2017; Weisburd 2015), and aggregate-fatality designs that register large regime effects but cannot separate behavior from selection, miles, or claim filing (Panel C) (Cohen and Dehejia 2004); dynamic and static insurance-data tests, by contrast, detect none (Abbring et al. 2003; Chiappori and Salanie 2000). Our collision-deductible analysis aligns with these null tests: the within-state response is a precise null, and only across-state, on the extensive margin, is there a small significant reduction ($|\hat{\beta}| \approx 0.011$; 0.0146% per dollar per year of out-of-pocket cost). Lastly, our hazard-model and precipitation analyses demonstrate that common environmental factors can sharply raise driving risk without becoming salient to the riskiest drivers, yielding a vanishingly small response.

Many policy-relevant questions are still outstanding. Whether repeated or continuous cueing sustains the reduction, and at what frequency before habituation sets in, is the margin on which any behavior-contingent program turns, and it may be the most practical way to learn why everyday incentives move drivers so little. Which structural mechanism underlies the neglect, neither our design nor the existing field evidence resolves: a misperception of the risk, or its simply not entering the choice set at all, whether through a rational default and adjustment costs or through inattention; and even eliciting risk perceptions may change the behavior it measures (Zwane et al. 2011). Our drivers, moreover, bear only a narrow slice of accident cost, chiefly the collision deductible; private motorists, who face premium experience-rating and higher liability exposure, may respond more, which consumer-grade telematics now makes tractable to test. Because deductibles and weather shift the marginal cost of driving rather than whether a driver is covered at all, whether behavior responds more sharply on the extensive margin, when coverage is lost or dropped, remains open, as does whether other risky behaviors such as speeding, harsh acceleration, or drowsy driving are governed differently; mapping how drivers' perceived riskiness of each departs from its statistical truth is a natural next step. Where behavior does respond, our intervention, together with a companion paper (Jin and Yu 2021), shows how sensor data can bring contracts closer to the first-best of Holmström (1979).

References

Abadie, Alberto, Susan Athey, Guido W. Imbens, and Jeffrey M. Wooldridge (2023). "When Should You Adjust Standard Errors for Clustering?" In: *The Quarterly Journal of Economics* 138.1, pp. 1–35.

Abbring, Jaap H, Pierre-André Chiappori, and Jean Pinquet (2003). "Moral hazard and dynamic insurance data". In: *Journal of the European Economic Association* 1.4, pp. 767–820.

Baicker, Katherine, Sendhil Mullainathan, and Joshua Schwartzstein (2015). "Behavioral hazard in health insurance". In: *Quarterly Journal of Economics* 130.4, pp. 1623–1667.

Barseghyan, Levon, Francesca Molinari, Ted O'Donoghue, and Joshua C Teitelbaum (2013). "Distinguishing probability weighting from risk misperceptions in field data". In: *American Economic Review* 103.3, pp. 580–585.

— (2018). "Estimating risk preferences in the field". In: *Journal of Economic Literature* 56.2, pp. 501–64.

Bhargava, Saurabh and Vikram Pathania (2013). "Driving under the (cellular) influence: The link between cell phone use and vehicle crashes". In: *American Economic Journal: Economic Policy* 5.3, pp. 92–125.

Blincoe, Lawrence, Ted R. Miller, Jing-Shiarn Wang, David Swedler, Tristan Coughlin, Bruce Lawrence, Feng Guo, Sheila Klauer, and Thomas Dingus (Feb. 2023). *The Economic and Societal Impact of Motor Vehicle Crashes, 2019 (Revised)*. Tech. rep. DOT HS 813 403. Washington, D.C.: National Highway Traffic Safety Administration, U.S. Department of Transportation.

Bortles, William, Sean McDonough, Connor Smith, and Michael Stogsdill (2017). *An Introduction to the Forensic Acquisition of Passenger Vehicle Infotainment and Telematics Systems Data*. SAE Technical Paper 2017-01-1437. SAE International.

Brot-Goldberg, Zarek C, Amitabh Chandra, Benjamin R Handel, and Jonathan T Kolstad (2017). "What does a deductible do? The impact of cost-sharing on health care prices, quantities, and spending dynamics". In: *The Quarterly Journal of Economics* 132.3, pp. 1261–1318.

Calabresi, Guido (2008). *The cost of accidents: a legal and economic analysis*. Yale University Press.

Callaway, Brantly, Andrew Goodman-Bacon, and Pedro H. C. Sant'Anna (2024). "Difference-in-Differences with a Continuous Treatment". In: *arXiv working paper*. arXiv:2107.02637.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller (2011). "Robust Inference with Multiway Clustering". In: *Journal of Business & Economic Statistics* 29.2, pp. 238–249.

Card, David (1992). "Using regional variation in wages to measure the effects of the federal minimum wage". In: *Industrial and Labor Relations Review* 46.1, pp. 22–37.

Chen, Jiafeng and Jonathan Roth (2024). "Logs with zeros? Some problems and solutions". In: *The Quarterly Journal of Economics* 139.2, pp. 891–936.

Chiappori, Pierre-André and Bernard Salanie (2000). "Testing for asymmetric information in insurance markets". In: *Journal of Political Economy* 108.1, pp. 56–78.

Cohen, Alma and Rajeev Dehejia (2004). "The effect of automobile insurance and accident liability laws on traffic fatalities". In: *The journal of law and economics* 47.2, pp. 357–393.

Cohen, Alma and Liran Einav (2003). "The Effects of Mandatory Seat Belt Laws on Driving Behavior and Traffic Fatalities". In: *The Review of Economics and Statistics* 85.4, pp. 828–843.

Delgado, M. Kit, Jeffrey P. Ebert, Ruiying A. Xiong, Flaura K. Winston, Catherine C. McDonald, Roy M. Rosin, Kevin G. Volpp, Ian J. Barnett, Dylan S. Small, Douglas J. Wiebe, Dina Abdel-Rahman, Jessica E. Hemmons, Rafi Finegold, Benjamin Kotrc, Emma Radford, William J. Fisher, Kristen L. Gaba, William C. Everett, and Scott D. Halpern (2024). "Feedback and Financial Incentives for Reducing Cell Phone Use While Driving: A Randomized Clinical Trial". In: *JAMA Network Open* 7.7, e2420218.

Dingus, Thomas A, Feng Guo, Suzie Lee, Jonathan F Antin, Miguel Perez, Mindy Buchanan-King, and Jonathan Hankey (2016). "Driver crash risk factors and prevalence evaluation using naturalistic driving data". In: *Proceedings of the National Academy of Sciences* 113.10, pp. 2636–2641.

Dionne, Georges (2013). "The empirical measure of information problems with emphasis on insurance fraud and dynamic data". In: *Handbook of insurance*. Springer, pp. 423–448.

Dionne, Georges and Ying Liu (2021). "Effects of Insurance Incentives on Road Safety: Evidence from a Natural Experiment in China". In: *Scandinavian Journal of Economics* 123.2, pp. 453–477.

Dionne, Georges, Pierre-Carl Michaud, and Maki Dahchour (2013). "Separating Moral Hazard from Adverse Selection and Learning in Automobile Insurance: Longitudinal Evidence from France". In: *Journal of the European Economic Association* 11.4, pp. 897–917.

Ebert, Jeffrey P., Ruiying A. Xiong, Neda Khan, Dina Abdel-Rahman, Aaron Leitner, William C. Everett, Kristen L. Gaba, William J. Fisher, Catherine C. McDonald, Flaura K. Winston, Roy M. Rosin, Kevin G. Volpp, Ian J. Barnett, Douglas J. Wiebe, Scott D. Halpern, and M. Kit Delgado (2024). "A randomized trial of behavioral interventions yielding sustained reductions in distracted driving". In: *Proceedings of the National Academy of Sciences* 121.32, e2320603121.

Ehrlich, Isaac and Gary S Becker (1972). "Market insurance, self-insurance, and self-protection". In: *Journal of Political Economy* 80.4, pp. 623–648.

Einav, Liran and Amy Finkelstein (2018). "Moral hazard in health insurance: What we know and how we know it". In: *Journal of the European Economic Association* 16.4, pp. 957–982.

Engelbrecht, Jarret, Marthinus J. Booysen, Gert-Jan Van Rooyen, and F. J. Bruwer (2015). "Survey of smartphone-based sensing in vehicles for intelligent transportation system applications". In: *IET Intelligent Transport Systems* 9.10, pp. 924–935.

Gallagher, Justin (2014). "Learning about an Infrequent Event: Evidence from Flood Insurance Take-Up in the United States". In: *American Economic Journal: Applied Economics* 6.3, pp. 206–233.

Geyer, Alois, Daniela Kremslehner, and Alexander Muermann (2020). "Asymmetric Information in Automobile Insurance: Evidence From Driving Behavior". In: *Journal of Risk and Insurance* 87.4, pp. 969–995.

Goodwin, Arthur H, Libby Thomas, Bevan Kirley, William Hall, Natalie P O'Brien, and Kate Hill (2015). *Countermeasures that work: A highway safety countermeasure guide for state highway safety offices, 2015*. Tech. rep. DOT HS 812 202. National Highway Traffic Safety Administration.

Governors Highway Safety Association (2025). *Distracted Driving: State Laws by Issue*. Laws last reviewed by State Highway Safety Offices June 2025. Governors Highway Safety Association. URL: https://www.ghsa.org/state-laws/issues/distracted%20driving (visited on 06/24/2026).

Handel, Benjamin and Joshua Schwartzstein (2018). "Frictions or mental gaps: What's behind the information we (don't) use and when do we care?" In: *Journal of Economic Perspectives* 32.1, pp. 155–178.

Hertwig, Ralph, Greg Barron, Elke U. Weber, and Ido Erev (2004). "Decisions from Experience and the Effect of Rare Events in Risky Choice". In: *Psychological Science* 15.8, pp. 534–539.

Holmström, Bengt (1979). "Moral hazard and observability". In: *The Bell Journal of Economics* 10.1, pp. 74–91.

Israel, Mark (2004). *Do We Drive More Safely When Accidents Are More Expensive? Identifying Moral Hazard from Experience Rating Schemes*. CSIO Working Paper 0043. Center for the Study of Industrial Organization, Northwestern University.

James Jr., Fleming (1948). "Accident Liability Reconsidered: The Impact of Liability Insurance". In: *Yale Law Journal* 57.4, pp. 549–570.

Jeziorski, Przemyslaw, Elena Krasnokutskaya, and Olivia Ceccarini (2017). "Adverse selection and moral hazard in a dynamic model of auto insurance". In: *Working Paper*.

Jin, Yizhou and Shoshana Vasserman (2021). *Buying Data from Consumers: The Impact of Monitoring in U.S. Auto Insurance*. NBER Working Paper 29096. National Bureau of Economic Research.

Jin, Yizhou and Thomas Yu (2021). "How to Prevent Traffic Accidents: Moral Hazard, Inattention, and Behavioral Data". In: *UC Berkeley Working Paper*.

Klauer, Sheila G., Feng Guo, Bruce G. Simons-Morton, Marie Claude Ouimet, Suzanne E. Lee, and Thomas A. Dingus (2014). "Distracted Driving and Risk of Road Crashes among Novice and Experienced Drivers". In: *New England Journal of Medicine* 370.1, pp. 54–59.

Kunreuther, Howard and Mark Pauly (2004). "Neglecting disaster: Why don't people insure against large losses?" In: *Journal of Risk and Uncertainty* 28.1, pp. 5–21.

Levitt, Steven D and Jack Porter (2001). "How dangerous are drinking drivers?" In: *Journal of political Economy* 109.6, pp. 1198–1237.

Llerena, Luis E, Kathy V Aronow, Jana Macleod, Michael Bard, Steven Salzman, Wendy Greene, Adil Haider, and Alex Schupper (2015). "An evidence-based review: distracted driver". In: *Journal of trauma and acute care surgery* 78.1, pp. 147–152.

McEvoy, Suzanne P., Mark R. Stevenson, Anne T. McCartt, Mark Woodward, Claire Haworth, Peter Palamara, and Rina Cercarelli (2005). "Role of mobile phones in motor vehicle crashes resulting in hospital attendance: a case-crossover study". In: *BMJ* 331.7514, p. 428.

Megat-Johari, Nusayba, Megat-Usamah Megat-Johari, Peter Savolainen, and Timothy J Gates (2023). "Evaluation of enforcement and messaging campaign focused on reducing cell phone-related distracted driving". In: *Transportation research record* 2677.1, pp. 1741–1752.

National Highway Traffic Safety Administration (Apr. 2024). *Distracted Driving in 2022*. Traffic Safety Facts Research Note DOT HS 813 559. U.S. Department of Transportation, National Center for Statistics and Analysis.

National Safety Council (May 2013). *Crashes Involving Cell Phones: Challenges of Collecting and Reporting Reliable Crash Data*. White Paper. Itasca, IL: National Safety Council.

Nevo, Aviv, John L Turner, and Jonathan W Williams (2016). "Usage-based pricing and demand for residential broadband". In: *Econometrica* 84.2, pp. 411–443.

Peltzman, Sam (1975). "The Effects of Automobile Safety Regulation". In: *Journal of Political Economy* 83.4, pp. 677–725.

Prentice, Ross L and Ronald Pyke (1979). "Logistic disease incidence models and case-control studies". In: *Biometrika* 66.3, pp. 403–411.

Redelmeier, Donald A. and Robert J. Tibshirani (1997). "Association between cellular-telephone calls and motor vehicle collisions". In: *New England Journal of Medicine* 336.7, pp. 453–458.

Reimers, Imke and Benjamin R. Shiller (2019). "The impacts of telematics on competition and consumer behavior in insurance". In: *The Journal of Law and Economics* 62.4, pp. 613–632.

Shavell, Steven (1987). *Economic Analysis of Accident Law*. Cambridge, MA: Harvard University Press.

— (2003). *Economic analysis of accident law*.

— (2018). "On liability and insurance". In: *Economics and Liability for Environmental Problems*. Routledge, pp. 151–163.

Singer, Judith D and John B Willett (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. Oxford university press.

Soleymanian, Miremad, Charles B. Weinberg, and Ting Zhu (2019). "Sensor Data and Behavioral Tracking: Does Usage-Based Auto Insurance Benefit Drivers?" In: *Marketing Science* 38.1, pp. 21–43.

Solon, Gary, Steven J. Haider, and Jeffrey M. Wooldridge (2015). "What Are We Weighting For?" In: *Journal of Human Resources* 50.2, pp. 301–316.

Spinnewijn, Johannes (2015). "Unemployed but Optimistic: Optimal Insurance Design with Biased Beliefs". In: *Journal of the European Economic Association* 13.1, pp. 130–167.

Stevens, Scott E, Carl J Schreck III, Shubhayu Saha, Jesse E Bell, and Kenneth E Kunkel (2019). "Precipitation and fatal motor vehicle crashes: continental analysis with high-resolution radar data". In: *Bulletin of the American Meteorological Society* 100.8, pp. 1453–1461.

Strayer, David L and William A Johnston (2001). "Driven to distraction: Dual-task studies of simulated driving and conversing on a cellular telephone". In: *Psychological science* 12.6, pp. 462–466.

Svenson, Ola (1981). "Are we all less risky and more skillful than our fellow drivers?" In: *Acta Psychologica* 47.2, pp. 143–148.

Sydnor, Justin (2010). "(Over) insuring modest risks". In: *American Economic Journal: Applied Economics* 2.4, pp. 177–99.

Vickrey, William (1968). "Automobile accidents, tort law, externalities, and insurance: An economist's critique". In: *Law and Contemporary Problems* 33.3, pp. 464–487.

Wagner, Gerhard (2006). "Tort law and liability insurance". In: *The Geneva Papers on Risk and Insurance-Issues and Practice* 31, pp. 277–292.

Wahlström, Johan, Isaac Skog, and Peter Händel (2017). "Smartphone-based vehicle telematics: A ten-year anniversary". In: *IEEE Transactions on Intelligent Transportation Systems* 18.10, pp. 2802–2825.

Weisburd, Sarit (2015). "Identifying moral hazard in car insurance contracts". In: *Review of Economics and Statistics* 97.2, pp. 301–313.

Wilde, Gerald J. S. (1982). "The Theory of Risk Homeostasis: Implications for Safety and Health". In: *Risk Analysis* 2.4, pp. 209–225.

Wilson, Fernando A and Jim P Stimpson (2010). "Trends in fatalities from distracted driving in the United States, 1999 to 2008". In: *American journal of public health* 100.11, pp. 2213–2219.

Zwane, Alix Peterson, Jonathan Zinman, Eric Van Dusen, William Pariente, Clair Null, Edward Miguel, Michael Kremer, Dean S. Karlan, Richard Hornbeck, Xavier Giné, Esther Duflo, Florencia Devoto, Bruno Crepon, and Abhijit Banerjee (2011). "Being surveyed can change later behavior and related parameter estimates". In: *Proceedings of the National Academy of Sciences* 108.5, pp. 1821–1826.

A Conceptual Framework

To fix ideas, we present a rational expectation benchmark for HPU behavior and formalize moral hazard as an elasticity to risk exposure that we quantify in Section 4. Our framework follows the self-protection tradition of Ehrlich and Becker (1972), in which agents choose effort to reduce the probability of loss, as distinct from self-insurance (reducing loss magnitude). The driver's choice of a privately beneficial but risk-increasing behavior against its accident cost similar to the driver-choice model of Peltzman (1975).

HPU, denoted as b, influences consumption h by providing deterministic private benefit $g(b;\beta)$, where β governs the HPU elasticity with respect to incentives, and by causing stochastic accidents at rate $\lambda(b,x)$, where x represents environmental factors related to the riskiness of HPU. When accidents occur, consumers suffer loss L based on their insurance y, which the transportation company sets exogenously.

We assume that (1) the arrival of accidents is Poisson; (2) accident loss L is deterministic in drivers' expectation conditional on y; (3) we adopt the Arrow-Pratt approximation $u(b \mid \Omega;\vartheta) = \mathbb{E}[h] - \frac{1}{2}\gamma\,\mathbb{V}[h]$, with γ as the absolute risk aversion coefficient. This is appropriate given that each unit of HPU carries minute incremental accident risk, and can be viewed as a "negligible third-derivative" approach (Barseghyan et al. 2018).

We can then tractably characterize utility as follows, as the private benefit of b minus the certainty equivalent (CE) of the risk of accident loss associated with b.

$$\begin{aligned} u(b|\Omega;\vartheta) =& \mathbb{E}[h] - \frac{\gamma}{2}\mathbb{V}[h] \\ =& g(b;\beta) - \overbrace{\underbrace{\lambda(b,x;p)}_{\text{risk occurrence}}\ \underbrace{L(y)\left[1+\frac{\gamma}{2}L(y)\right]}_{\text{loss after insurance}}}^{CE} \end{aligned} \tag{A1}$$

where $\Omega = \{x, y\}$ and $\vartheta = \{\beta, \gamma\}$

If we further assume that (4) $g(b)$ is increasing and concave so that there are diminishing returns to the private benefit of conducting b, we obtain a tractable first-order condition shown in Equation A2, in which we define the implicit price of HPU as the certainty equivalent of the additional risk exposure caused by each unit of b. We further specialize $g(b;\beta) := b^{1+1/\beta}/(1+1/\beta)$ for tractability and to nest the constant-elasticity benchmark commonly used in moral-hazard calibration (cf. how cellphone usage is modeled in Nevo et al. (2016)); the parameterization restricts the response of b to its implicit price to be constant-elasticity, with concavity requiring $\beta < 0$ (the empirically relevant case here). This leads to a simple characterization of b (Equation A3).

We have now expressed the ex-ante moral hazard effect as a log-linear "demand curve," in which consumers' risk-taking behavior b responds to the "implicit price" of b according to an elasticity term β that denotes how much they value b independently from the risk. Because HPU's marginal accident contribution $\lambda_b(b;x)$ can itself vary with b, Equation A3 is an implicit first-order condition rather than a closed-form demand curve; we read β as a *local* elasticity—the response of b to a marginal shift in its implicit price, holding the accident technology's dependence on b fixed—which is how we recover it empirically, as the ratio $\%\Delta\text{HPU}/\%\Delta\text{MC}$. Assuming individuals are risk-averse and that the private benefit of HPU is positive, we can deduce that b reduces in environments that heighten its marginal accident cost; and that more insurance coverage increases b.

$$b^*_{it} = \left[\underbrace{\xi(b,y;x,\gamma)}_{\text{implicit price of } b}\right]^{\beta}, \quad \xi(b,y;x,\gamma) := \frac{\partial CE(b,y;x,\gamma)}{\partial b} > 0 \tag{A2}$$

$$\log b^* = \underbrace{\beta}_{\text{HPU elasticity}} \cdot \left[\underbrace{\log \lambda_b(b;x)}_{\substack{\text{HPU's marginal} \\ \text{contribution on} \\ \text{accident rate}}} + \underbrace{\log(L(y))}_{\text{deductibles}} + \underbrace{\log(1+\frac{\gamma}{2}L(y))}_{\text{risk aversion}}\right] \tag{A3}$$

B Literature Comparison

Table A1 situates our $|\hat{\beta}|$ alongside existing estimates of behavioral response to risk in driving, organized by intervention design, with the literature first and this paper's analyses placed alongside their comparable studies (Panel A: auto-insurance moral-hazard estimates, with this paper's precipitation and deductible analyses; Panel B: engineered monitoring and salience interventions, with this paper's text-message experiment; Panel C: aggregate-fatality studies) and discussed in Section 4 and the Conclusion.

Table A1: Behavioral Response to Risk in Driving: Literature Comparison and Implied Elasticity

Study (data)	Design / identification (*dep. var.*)	Reported estimate (source exhibit)	Implied $\lvert\beta\rvert$ or k
Panel A: Moral hazard			
Chiappori & Salanié (2000); French auto, young drivers	*Claims:* Conditional coverage–claim (positive-correlation) test	$\hat{\rho} \approx 0$, insignificant	missing num. & denom.; ≈ 0

(*continued on next page*)

Table A1 (*continued*)

Study (data)	Design / identification (*dep. var.*)	Reported estimate (source exhibit)	Implied $\lvert\beta\rvert$ or k
Abbring, Chiappori & Pinquet (2003); French auto	*Claims:* Test for negative occurrence dependence under experience rating	"No evidence of moral hazard"	missing denom.; ≈ 0
Israel (2004); US (Illinois) auto, 1989–98	*Claims:* Experience-rating claim-cost changes ("insurance events" + 1997 surcharge DiD)	\$100 higher claim cost $\Rightarrow -0.4$pp (off 2.9% base) [Table 5]	$k \approx -14\%$ per \$100 (reported); β missing denom.
Abbring, Chiappori & Zavadil (2008); Dutch auto, 1995–2000	*Claims:* Dynamic claim-rate model; incentive-dependence tests	Significant: $\hat{\beta} \approx 0.4$ [Table 5]—a semi-elasticity of the claim rate to a utility incentive ΔV (no price coefficient to convert to cost); part ex-post	missing denom.; > 0, mixed w/ ex-post
Dionne, Michaud & Dahchour (2013); French auto panel, 1995–97	*Accidents:* Dynamic-data test (bonus-malus, lagged coverage) separating MH from selection/learning	MH among < 15-yr-experience drivers ($\hat{\phi}_{\text{MH1}} = -2.24$, $\hat{\phi}_{\text{MH2}} = +0.65$); none > 15 yr [Table 5]	missing num. (index coefs., no marginal effects); > 0
Weisburd (2015); Israeli auto, 2001–08	*Accidents:* Employer-set coverage IV on own accident cost C_A	\$100 lower $C_A \Rightarrow +1.7$pp ($= +10\%$ at mean 16.3%) [Table 3]	≈ 0.1–0.4 ($k \approx 10\%$ per \$100)
Jeziorski et al. (2017); Portuguese auto, 2002–07	*Accidents:* Structural dynamic coverage-choice / effort model	Comprehensive $\approx 2\times$ liability rate [Table 17]; -30% risk, class 1→13 [Table 16]	≈ 0.1–1
Dionne & Liu (2021); China auto, 2009–12	*Claims:* Two-city natural experiment (DiD); premiums priced on past traffic violations	Violation pricing $\Rightarrow$ significantly lower claim frequency; past-claims pricing no effect	missing denom.; > 0
This paper (precipitation)	*HPU:* Rideshare HPU; weather shocks amplify HPU's accident cost	Small, sign-consistent %ΔHPU (Section 4)	$\approx \mathbf{0.005}$–$\mathbf{0.032}$
This paper (deductible hike)	*HPU:* Rideshare HPU; collision deductible \$1,000→\$2,500	\$100/yr higher expected OOP $\Rightarrow -1.46\%$ HPU (across-state ext.; within-state null) [Table 4]	$\approx \mathbf{0.011}$ ($k = -1.46\%$ per \$100/yr)
Panel B: Financial-incentive or salience interventions[§]			
Bolderdijk et al. (2011)	*Speed:* Dutch young drivers; PAYD speed pricing + feedback	Speeding -0.9pp off an 18.6% base while active; rebounds on removal	$k \approx -0.8\%$ per EUR 100/yr (at max stake)
Reimers & Shiller (2019)	*Fatal acc.:* U.S. auto; PHYD market expansion	-1.6% fatal acc. per PHYD firm [Table VI]; $\approx 50\%$ per enrollee	missing denom.
Soleymanian et al. (2019)	*Hard-braking:* U.S. auto; voluntary UBI monitoring + premium discount	$\approx -21\%$ hard-braking over 6 mo. for a 12% premium discount	missing denom. (premium level unobserved)
Delgado et al. (2024)	*HPU:* U.S. motorists; feedback + weekly cash	-21.2% HPU at \$7.15/wk vs. -16.0% at \$14.29/wk	$k \approx -5.7$ and -2.2% per \$100/yr (two stakes)
Ebert et al. (2024)	*HPU:* U.S. motorists; gamified leaderboard ± cash	-20% HPU (gamif.); -28% (w/ cash, n.s. vs. gamif.)	missing denom. (cash increment n.s.)
Jin & Vasserman (2024)	*Claims:* U.S. auto; voluntary telematics monitoring, 22 states	$\approx 30\%$ safer while monitored [Table 2]	missing denom. (monitoring, not priced)
This paper (SMS experiment)	*HPU:* Rideshare HPU; randomized salient text-message warning	$\mathbf{-19.1}\%$ HPU, sustained (Section 5)	zero-price cue; k undefined
Panel C: Aggregate-fatality designs			

(*continued on next page*)

Table A1 (*continued*)

Study (data)	Design / identification (*dep. var.*)	Reported estimate (source exhibit)	Implied $\lvert\beta\rvert$ or k
Cohen & Einav (2003); U.S. state panel, 1983–97	*Fatalities:* Seat-belt usage IV'd by mandatory-law variation; nonoccupant-fatality offset test	No significant compensating behavior; usage cuts total fatalities	missing denom.†; ≈ 0
Cohen & Dehejia (2004); U.S. state panel, 1970–98	*Fatalities:* Staggered compulsory-insurance / no-fault; uninsured-ratio IV	No-fault adoption $\Rightarrow +6\%$ fatalities [Table 7]	$\approx 0.1^{\dagger}$, mixed w/ selection & miles

Notes: $\lvert\beta\rvert$ is the elasticity of risky-driving behavior with respect to its expected per-unit accident cost, $\beta = (\%\Delta \text{ risky behavior})/(\%\Delta \text{ expected marginal accident cost})$; k is a dollar semi-elasticity ($\%\Delta$ behavior per \$100 per year of stake), reported wherever a study attaches a dollar figure to the response. Cells state which ingredient a study lacks: a measured behavior change (numerator) or a measured cost change (denominator), with the identified sign appended (≈ 0 for a null test, > 0 for a response detected but not sized); *mixed w/* flags responses the source attributes partly to channels other than ex-ante moral hazard. The dependent variable is italicized in each design cell; claim-, accident-, and fatality-based outcomes bundle driving with claim filing, selection, and mileage. § Panel B incentives are priced off *observed behavior* rather than off the accident itself, so a Panel B k is the price of behavior, not of accident risk—a distinction to bear in mind when comparing with Panel A; consistent with salience rather than price doing the work, doubling the stake in Delgado et al. (2024) halved k rather than doubling the response, and adding cash to gamification in Ebert et al. (2024) added nothing significant. † Panel C fatality designs use a discrete legal regime as the "price," so the $\%\Delta$MC denominator is assumed (a stated 30–100% cost shift), not measured, and any $\lvert\beta\rvert$ is illustrative and conflates behavior with selection, miles, and claim filing. Headline estimates are read from each paper's full text; study-by-study derivations are available from the authors.

C Moral-Hazard Elasticity Calibration Details

This appendix documents the back-of-envelope arithmetic underlying each panel and row of Table A2, whose headline numbers Section 4.3 reports. The general formula is $\hat{\beta} = (\%\Delta\text{HPU})/(\%\Delta\text{MC of HPU})$, where MC is the expected marginal accident cost of HPU per unit HPU; because a positive cost shock reduces HPU, $\%\Delta\text{HPU} < 0$ and $\hat{\beta} < 0$. Both terms are evaluated at a comparable per-unit reference (per-rainy-second for precipitation; per year of expected out-of-pocket cost for the deductible). Both denominators are calibrated at their midpoints; at either denominator's floor the implied $\lvert\hat{\beta}\rvert$ stays at or below the $\lvert\beta\rvert \approx 0.1$–1 causal auto-insurance range (Appendix Table A1).

Precipitation denominator: joint $\%\Delta$MC. With a multiplicative hazard, HPU's marginal accident cost at a given rain level is the baseline hazard times HPU's proportional kick, $\text{MC}_{\text{HPU}} \approx \pi_{\text{baseline}} \cdot (\text{OR}_{\text{HPU}} - 1)$. Going from dry to the typical rainy intensity (0.2034 in/hr, the rainy-trip mean of the Section 4.1 rain-intensity variable and the level at which ev-

Table A2: Calibrated moral-hazard elasticity $\hat{\beta}$ across observational analyses

Identification	% Δ HPU (SE)	% Δ MC	$\hat{\beta}$ (SE)
Panel A: Precipitation (weather-amplified marginal cost)			
Within-driver variation			
Extensive margin	-0.189^{***} (0.060)	+34.71	-0.005^{***} (0.002)
QMLE	-0.458^{***} (0.146)	+34.71	-0.013^{***} (0.004)
Within-trip variation			
Extensive margin	-0.271^{***} (0.103)	+26.32	-0.010^{***} (0.004)
QMLE	-0.842^{***} (0.324)	+26.32	-0.032^{***} (0.012)
Panel B: Collision deductible \$1,000→\$2,500 (out-of-pocket cost)			
Within-state variation			
Extensive margin	0.20 (0.78)	+84.23	0.002 (0.009)
QMLE	1.98 (1.91)	+84.23	0.024 (0.023)
Across-state variation			
Extensive margin	-0.91^{**} (0.42)	+84.23	-0.011^{**} (0.005)
QMLE	−3.42 (3.23)	+84.23	−0.041 (0.038)

Notes: $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$ (two-sided, from the %ΔHPU clustered sampling SE). $\hat{\beta} = (\%\Delta\text{HPU})/(\%\Delta\text{MC})$; $\hat{\beta} < 0$ is moral-hazard-consistent. The $\hat{\beta}$ SE is the %ΔHPU SE divided by the %ΔMC midpoint (third column) and reflects numerator sampling variation only; denominator-band sensitivity is reported in the surrounding text. %ΔHPU numbers sourced from Tables 3 and 4; %ΔMC from the hazard model (Table 2, cols 7/8) for precipitation and, for the deductible, the pricing-implied percent change in expected out-of-pocket cost (driver-weighted; Section 4.2), whose sensitivity band spans the severity family and pricing vintage (see the Deductible-rows paragraph below).

ery numerator below is evaluated) moves both factors: the Table 2 column-7 precipitation main effect lifts $\log \pi_{\text{baseline}}$ by 0.14 log-points, and the HPU $\times$ Precip interaction lifts $\log(\text{OR}_{\text{HPU}} - 1)$ by 0.16, giving $\%\Delta\text{MC} = 34.71\%$. The within-driver rows of Table A2 use this denominator; the within-trip rows use its column-8 (trip-FE) counterpart, 26.32%. Using the interaction alone (8.63%) would understate MC by a factor of 4.0.

Precipitation numerators (Panel A). All four numerators are evaluated at the same rainy intensity. On the extensive margin (Table 3, column 3), the Rain $\times$ start-intensity coefficient times the rainy-trip regressor mean $\overline{Inch} = 0.19$ (corresponding to 0.2034 in/hr), over the baseline $\Pr(\text{HPU} > 0) = 0.666$, gives $\%\Delta\text{HPU} = -0.189\%$; the Rain $\times$ Δ-intensity coefficient gives the within-trip -0.271%. The Poisson QMLE (column 7) recovers $\Delta \log E[\text{HPU}]$ directly from the same two coefficients: -0.458% and -0.842%, both significant at the 1% level. Each start-intensity numerator is divided by the within-driver denominator and each Δ-intensity numerator by the within-trip denominator, giving $\hat{\beta} \approx -0.005$, -0.013, -0.010, and -0.032.

Deductible rows (Panel B). Both the denominator and the dose that scales each numerator derive from two county pricing points—the expected collision costs at the \$500 and \$1,000 deductibles, $P_c(500)$ and $P_c(1000)$—that build Section 4.2's dose. With $P_c(d) = \lambda_c\, E[(s-d)^+]$ for loss draw s and exponential severity, the ratio $r_c = P_c(1000)/P_c(500) = e^{-500\zeta_c}$ identifies the rate ζ_c, the level identifies the frequency λ_c, and $\mathrm{EOOP}_c(d) = \lambda_c(1 - e^{-\zeta_c d})/\zeta_c$. All four deductible rows share one denominator, the driver-weighted percent change in expected out-of-pocket cost, $\%\Delta\mathrm{MC} = 100 \times \overline{\Delta\mathrm{EOOP}}/\overline{\mathrm{EOOP}(1000)} = 84.2\%$. The numerators convert each regression coefficient $\hat{\theta}$ (per \$100/yr of dose) into $\%\Delta$HPU at the mean dose—$100(e^{\hat{\theta}\bar{x}} - 1)$ for the Poisson QMLE and $100\,\hat{\theta}\bar{x}/\bar{p}$ for the LPM, as in Online Appendix Table D7—so each $\hat{\beta}$ equals the text's dollar semi-elasticity $\hat{k}$ scaled by the base cost, $\hat{\beta} = \hat{k} \times \overline{\mathrm{EOOP}(1000)}/10^2$. The band spans the Pareto severity family ($x_m = \$100$: $\alpha_c = 1 - \log_2 r_c$, $\Delta\mathrm{EOOP}_c = P_c(1000)(1 - r_c^{\log_2 2.5})$) and the 2023 pricing vintage.