Received: 9 September 2021 | Revised: 8 March 2022 | Accepted: 13 March 2022
DOI: 10.1002/cb.2049

RESEARCH ARTICLE

WILEY

Do honesty-nudges really work? A large-scale field experiment in an insurance context

Jareef Bin Martuza[1] | Siv Rosendahl Skard[1] | Lavrans Løvlie[2] | Helge Thorbjørnsen[1,3]

[1]Department of Strategy and Management, Norwegian School of Economics, Bergen, Norway
[2]PwC, Oslo, Norway
[3]SNF, Center for Applied Research at NHH, Bergen, Norway

Correspondence
Jareef Bin Martuza, Department of Strategy and Management, Norwegian School of Economics, Helleveien 30, 5045 Bergen, Norway.
Email: jareefbinmartuza@gmail.com

Abstract

Businesses want their customers to self-report information honestly. One increasingly popular way to stimulate desired behavior is by using nudge interventions. But can customers be *nudged* to self-report information more honestly? This is currently a debate in the literature, where empirical results are inconclusive. Building on related literature on nudges, we add to this debate with a controlled field experiment ($N = 5704$). We used data from actual customers making real decisions when they file claims online to a large Nordic insurance provider. To the best of our knowledge, this is the first study to investigate the effects of honesty-nudges on self-reported information when filing insurance claims using a controlled field experiment. We designed and tested three honesty-nudges on insurance customers: (1) signing-at-the-beginning, (2) a descriptive social norm message, and (3) a solidarity message. Across five outcome measures, we found that the honesty-nudges, standalone or in any possible combination, do not have significant effects in reducing indicators of insurance claims fraud. But interestingly, customers in all treatment groups used significantly more characters to describe losses than customers in the control group. Also, in post hoc analyses, we found signs that the direction of nudge effects varies across customers' age and customer loyalty.

1 | INTRODUCTION

Insurance companies want their customers to *not* falsely report or exaggerate damages when they file claims. Fraudulent claims cost the insurance provider because those increase settlement costs. These can also cost honest customers as providers increase premiums to cover extra costs. In 2017, fraudulent claims cost European insurers and customers an estimated 13 billion Euros (Insurance Europe, 2019a). Finance Norway (Finanse Norge, 2019) estimates the total value of detected fraudulent non-life insurance claims in Norway at NOK 495.1 million (approximately 49 million Euros) in 2019. The total value of undetected yearly insurance fraud in Denmark is estimated to be 10 times that of detected-300 million Euros versus 32.75 million Euros[1] (Insurance Europe, 2019b). The Federal Bureau of Investigation (2020) estimates the cost of non-health insurance fraud in the United States at over 40 billion USD. Dealing with insurance claims fraud is difficult because providers rely on self-reported information and customers have economic incentives to report false or exaggerated damages.

[1]Estimates were in Euros than in Danish Kroners for uniformity in comparing figures across countries.

Willingness to commit insurance claims fraud may be influenced by accepting attitudes toward insurance fraud and negative perceptions of insurance companies (Tennyson, 1997). Insurance claims fraud can be categorized as (a) when an event of loss occurred, but the customer reported exaggerated damages, and (b) when no event of loss occurred, but the customer filed a false claim regardless (Tennyson, 2008). For example, if a customer lost one of their bags on a vacation that they bought travel insurance for, they may over-report the value of items in the lost bag (a). Alternatively, a customer may claim they lost a bag even though they did not lose any (b).

To reduce costs of insurance claims fraud, much effort has been put into detecting fraud (Morley et al., 2006). Detection measures often use data mining and deep learning techniques to flag suspicious cases (Kirlidog & Asuk, 2012). However, these techniques are not completely failproof and often require costly manual interventions. Moreover, these measures may still leave several fraudulent cases undetected (Sadgali et al., 2019). Naturally, it would help reduce costs to both providers and honest customers if customers could be *nudged* to file claims more honestly in the first place.

Thaler and Sunstein (2009, p. 6) define a nudge as "any aspect of the choice architecture that alters people's behavior in a predictable way without forbidding any options or significantly changing their economic incentives". Nudges have been found effective in reducing energy usage at home (Allcott, 2011; Ayres et al., 2013), promoting healthier eating (Cadario & Chandon, 2020), increasing college enrollment (Castleman & Page, 2015), promoting diversity in the workplace (Behavioral Insights Team, 2018), and so on. Nudges to stimulate honest behavior, hereafter *honesty-nudges*, try to stimulate honest behavior via moral suasion (Bott et al., 2020; Dimant et al., 2020). But, can honesty-nudges reduce *customer* dishonesty?

Several studies on stimulating honest behavior have demonstrated promising results in lab experiments (Bicchieri & Dimant, 2019; Bicchiery et al., 2021; Chou, 2015a; Chou, 2015b; Dimant et al., 2020; Koning et al., 2020). Out in the field, nudges seem to successfully influence people to file taxes more honestly (Bott et al., 2020; Hallsworth et al., 2017; John & Blume, 2018; Kettle et al., 2017). However, few studies have tested honesty-nudges with field experiments outside tax compliance. And even within nudging tax compliance, we see evidence of null (Kettle et al., 2017) or even negative (John & Blume, 2018) effects. Recent pre-registered and high-powered studies (Kristal et al., 2020) did not replicate the positive results of previous studies on honesty-nudges. Moreover, findings from a seminal paper on increasing honesty by Shu et al., 2012 have been retracted from PNAS. Consequently, there seems to be a need to revisit and further investigate the effects of honesty-nudges through high-powered field experiments.

In the current work, we implement and test three honesty-nudges, standalone and in all possible combinations of the three, using a controlled field experiment on actual customers of a large Nordic insurance provider. The customers make real choices when they file claims on the website, and those determine if and how much settlement they receive. We adapt two nudge interventions that may reduce dishonesty: signing-at-the-beginning and a descriptive social norm message. We also conceive a new informational nudge (Miesler et al., 2017) specific to an insurance context-a solidarity message-and test it for the first time.

Next, we outline the theory behind how the nudges can reduce dishonesty when customers file claims.

1.1 | Signing-at-the-beginning

On a self-report form, people usually confirm at the end that the information they provided is accurate. However, a signature-at-the-top of a form confirming a veracity statement can activate one's ethical standards right from the start of a process (Kristal et al., 2020). Salient ethics makes it harder to rationalize away the dissonant cognition of later false reporting because it contrasts vividly with one's ethical standards (Festinger, 1957). This makes it psychologically costlier to provide false information during later stages of the process.

Several studies that tested the signing-at-the-beginning nudge have found null effects in contexts of increasing honesty in tax compliance (Kettle et al., 2017), loan repayment (Bhanot, 2017), and reporting miles driven (Kristal et al., 2020). One commonality across these settings is that they involve evading or reducing payments. A nudge may not be enough to elicit honest reporting when there are strong motives to avoid financial pains of payment.

However, confirming honesty statements at the beginning has been found to make participants report more honestly in a hypothetical insurance claims setting in the lab (Leal et al., 2016). But it is yet to be tested if signing-at-the-beginning can reduce claims fraud in the real world where potential payouts (settlements) are much higher.

Self-reporting information in an insurance claims context (rather than for an insurance application) may have different mechanisms since the customer has an opportunity to gain payouts rather than reduce premiums. Lying to actively make gains may have higher psychological costs compared to lying to prevent losses. So, making ethics salient right from the start can make misreporting in later stages of the claim evoke greater dissonance. Accordingly, we hypothesize that signing-at-the-beginning can reduce claims fraud by nudging customers to self-report more honestly to avoid potential dissonance.

1.2 | Descriptive social norm message

A social norm-nudge is an informational nudge that works by giving people social information on what most people do or find appropriate to do in specific situations (Bicchieri & Dimant, 2019). A descriptive social norm message informs the target what most people do in a situation (Cialdini et al., 2006). In uncertain situations (e.g., how honestly should a customer file a claim), people may wonder what most people do (Bicchieri & Dimant, 2019). Information about what the normative behavior is can cue people and influence behavior in that situation.

Also, people are more likely to imitate the behavior of those they feel psychologically closer to (Gino & Galinsky, 2012). Therefore, descriptive social norm messages can be more effective when they are specified to a group that the target identifies with. Hotel guests re-used towels more when they were told that 75% of previous guests in the *same room* reused their towels than when told that 75% of hotel guests, *in general*, reused their towels (Goldstein et al., 2008).

Norm-nudging has been used to elicit honest self-reporting when filing tax forms (Bott et al., 2020; Hallsworth et al., 2017; John & Blume, 2018). Other settings include encouraging towel re-use in hotels (Goldstein et al., 2008), reducing energy usage at home (Allcott, 2011; Schultz et al., 2007), improving voter turnout (Gerber & Rogers, 2009), and increasing charitable giving (Agerström et al., 2016).

However, existing literature on norm-nudges also documents null effects. Norm-nudges delivered through letters increased tax compliance (Bott et al., 2020; Hallsworth et al., 2017), but norm-nudges administered on the tax filing website did not (Kettle et al., 2017).

We chose a descriptive norm message over an injunctive norm message because telling the customer that most others approve of filing claims honestly (injunctive) can make the customer think that others only say that out of social desirability. So, a descriptive norm message here can cue that filing claims honestly *is* the prevalent behavior. Given the mixed findings in nudging honesty, we test if a descriptive social norm message about similar customers can reduce insurance claims fraud.

1.3 | Solidarity message

Negative perceptions of insurance companies among customers are one of the key drivers of insurance claims fraud (Tennyson, 1997). Insurance companies may be perceived as large corporations profiting off individual customers. This can make customers incur lesser psychological costs when committing fraud. Moreover, a customer may not usually consider the consequences of fraud on other customers. An insurance customer participates in a risk pool where each member is reciprocally responsible for others' risks (Lehtonen & Liukko, 2011). An insurance provider enables individuals to come together to share the risk of an undesired event. Informing customers about these inner workings can evoke a sense of solidarity and influence customers to see the insurance provider in a more positive light.

People can be persuaded to stay honest when they expect others to reciprocate as well (Caraban et al., 2019). In a lab experiment, an additional sentence "Note that the recipient relies on you" significantly increased generous behavior in dictator games (Brañas-Garza, 2007). Moreover, several experiments suggest that people are less likely to lie for personal benefit when their lie can harm another party (Gneezy, 2005).

Evoking solidarity among customers can make consequences of actions to fellow customers more salient and help reduce an antagonistic "me versus insurance company" sentiment. Though the solidarity nudge does not directly state that misreporting can harm other customers, it tries to make the customer consider if it is fair to receive a settlement from the "common pot". Moral suasion based on fairness has been found to reduce tax fraud (Bott et al., 2020). Based on these arguments, we hypothesize that a solidarity message can nudge honest self-reporting as the customer may reflect on the consequences of their actions on other customers.

2 | MATERIALS AND METHODS

2.1 | Design

We investigate the effects of three honesty-nudges on insurance claims fraud with a controlled field experiment. Collaborating with a large Nordic insurance company, we implemented an experimental design where customers were randomly assigned to one of the eight conditions (see Table 1). We tested each of the three nudges standalone and in all possible combinations. The context is when customers file for claims online on a travel insurance policy they bought. Manipulations for the experiment were implemented on the user-interface customers faced when filing claims.

A natural field experiment (like ours) can combine features of the lab-randomized assignment, and naturally occurring data—

TABLE 1 There are eight conditions in the experiment

	Implemented nudge interventions		
Experimental condition	Signing-at-the-beginning (sign)	Descriptive social norm message (SN)	Solidarity message (sol)
Baseline			
Sign	✓		
SN		✓	
Sol			✓
Sign × SN	✓	✓	
Sign × Sol	✓		✓
SN × Sol		✓	✓
Sign × SN × Sol	✓	✓	✓

Note: Each customer was exposed to one, two, or three of the nudges. The nudges implemented in each condition are marked by a ✓ mark.

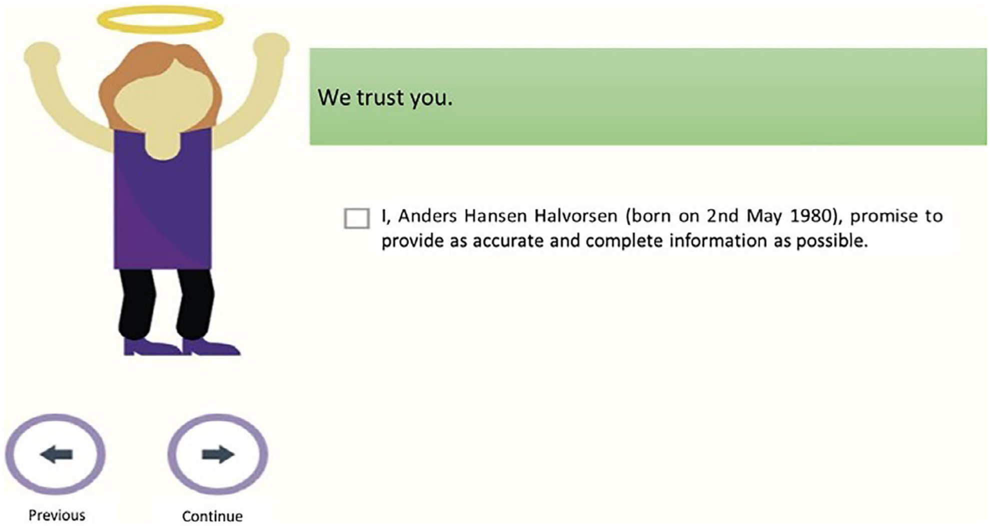


FIGURE 1 Illustration of the signing-at-the-beginning nudge. Customers were exposed to this right at the beginning of their claim [Colour figure can be viewed at wileyonlinelibrary.com]

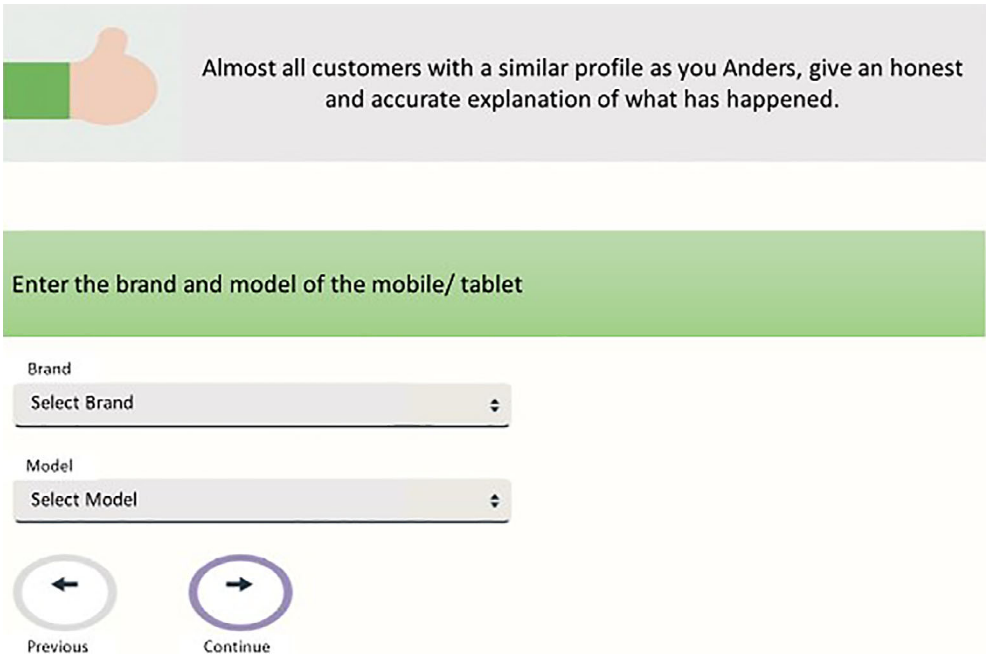


FIGURE 2 Illustration of the descriptive social norm nudge. Customers were exposed to this message when they were filling out details about their event of a loss [Colour figure can be viewed at wileyonlinelibrary.com]

realism (List, 2008). We did not collect informed consent from customers because that would affect what we are studying. Since we study the effect of honesty-nudges on claims fraud, telling customers they are participating in an experiment would make it difficult to measure the actual effect of the nudges. The customers were not debriefed because that can affect the provider's relationship with them.

As participating in the experiment does not harm customers, their data is completely anonymized, and honesty-nudge research can accrue social good-we make the case for not collecting informed consent (List, 2008). Moreover, our experiment does not ask for any additional data during the customer's claim, and the claiming user interface was the same during the experiment period as before, the experimental manipulations being the only changes.

2.2 | Procedures

When customers logged in to file a claim, they were randomly assigned to one of the eight conditions. Customers faced the same user interface except for subtle changes made to specific web pages to implement the nudges.

The travel insurance claims interface in this study has four main sections. The first section has a landing page that is displayed when customers first log in. This page is the same for customers across all conditions.

On the next page of this section, customers may be prompted to confirm a veracity statement before beginning their claim. This is where the signing-at-the-beginning nudge is implemented (see Figure 1 in Section 2.3.1).

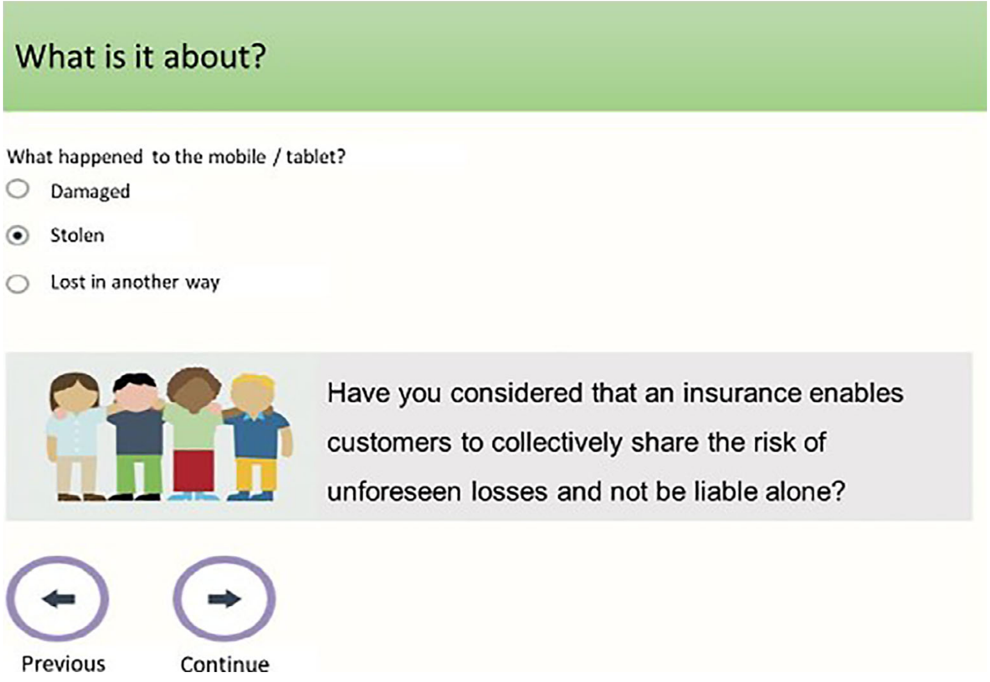


FIGURE 3 Illustration of the solidarity message nudge. Customers were exposed to this message when they were filling out details about their event of a loss [Colour figure can be viewed at wileyonlinelibrary.com]

The second section is where customers fill in information about their event of a loss. On the section's first page, the descriptive social norm message nudge is implemented by displaying it as a banner (see Figure 2 in Section 2.3.2). On the next page, the solidarity message is also implemented as a banner (see Figure 3 in Section 2.3.3).

The third section is where customers fill in details about the event of a loss. Here customers write in terms of text where and how this event happened. The fourth and final section asks customers to input contact information and submit the claim. Both these sections are the same across all conditions.

Anytime during the claiming process, the customer can choose to abort the claim. This is coded with a dummy variable. They can also choose to abort their claim online and then continue with a phone call instead. This is also coded with a dummy variable. We code these to improve the fidelity of our manipulation (see Section 3.1 for details). A summary flow of the procedures is illustrated in Figure A1 in Appendix A.

2.3 | Materials

Different (combinations of) nudges were administered to customers via subtle changes to pages in the user interface for filing claims. For example, a checkbox for a veracity statement, or a banner with a message on a particular page. All other elements of the interface were kept constant across conditions.

2.3.1 | Signing-at-the-beginning

Signing-at-the-beginning has been extended to online user interfaces where signing-at-the-beginning was used to increase tax compliance (Kettle et al., 2017) by getting people to confirm a veracity state *before* they filled in tax information. Similarly, our variant of this intervention entails customers "signing-at-the-beginning" of the claiming process.

Here, customers actively had to confirm the veracity statement-*I, [name], (born on [date of birth]), promise to provide as accurate and honest information as possible* right at the beginning of filing their claim (in addition to confirming at the end). To strengthen the manipulation, a visual element of a person with a halo was added, as well as a trust cue ("We trust you"), see Figure 1. The control condition entailed signing only at the end of the form (like standard terms and conditions agreements).

2.3.2 | Descriptive social norm message

In stimulating honesty, previous studies have used descriptive social norm message in a tax letter by adding a line-*most people file tax returns honestly* (Bott et al., 2020) or displaying directly on the online tax filing interface-*most people self-report tax information honestly* (Kettle et al., 2017).

Similarly, we adapted the intervention and reinforced it by informing the customer that those with profiles similar to them file claims honestly. Customers were shown the message-*almost all customers with a similar profile as you [name], give an honest and accurate explanation of what has happened*, in addition to a visual element, see Figure 2 below. This message appeared on a page where the customer fills in details about the incident.

2.3.3 | Solidarity message

The solidarity message, being an information nudge, follows the same mechanism of implementation as the descriptive social norm message.

We implement this information nudge by informing customers how they are in the same boat with others when they fill in the details of their claim.

Customers were shown the message—*Have you considered that an insurance enables customers to collectively share the risk of unforeseen losses and not be liable alone?*. To strengthen the manipulation, a visual element of a group of people with their arms around each other (animated) was added, see Figure 3 below. This message also appeared on a page where the customer fills in details about the event of a loss.

2.4 | Outcome measures

Since there is no single measure that detects insurance claims fraud directly, the industry uses several indicators for possible fraud. Tracy and Fox (1989) used mean repair expense estimates as an indicator to infer differences in levels of insurance fraud across groups in an auto body repair context. In collaboration with our industry partner, we developed a set of outcome measures to infer levels of fraud across conditions.

The outcome measures are (1) Claimed Amount, (2) Claim-Settlement Difference, (3) Session Cancelation, (4) Claim Rejection, and (5) Event Description Length. The measures do not confirm claims fraud or lack thereof at a case level. The goal here is to test if our treatments influence how honestly customers self-report information in their claims across conditions.

First, lower levels of Claimed Amount would suggest a positive effect of nudges in reducing indicators of claims fraud. This is because the "nudge-able" customers influenced by honesty-nudges would be less likely to exaggerate the amount of the loss. All things equal, the successfully nudged customers should lower average claimed amounts in nudged conditions.

Second, lower levels of Claim-Settlement Difference can also suggest a positive effect of nudges. Because claimed amounts that are close to settlement amounts issued by the insurance handler are more likely to be reported from honest claims. The successfully nudged customers should lower average differences in claimed amounts and settlement amounts in nudged conditions.

Third, a higher proportion of Session Cancelation can also suggest positive effects of nudges. This is because honesty-nudges can dissuade customers from filing fraudulent claims when no event of loss occurred, and thereby lead them to abort their claim. However, there could be several alternative reasons why consumers may cancel the reporting of an insurance incident. Such as, when they discover that the event is not covered, or they are interrupted. But these should not vary systematically across conditions since customers are randomly assigned. We are interested in the group that gets nudged to cancel a potentially fraudulent claim session. If customers cancel their session in coded with a dummy variable.

Fourth, a lower proportion of Claim Rejection can also suggest positive effects of nudges. If the proportion of fraudulent claims is lower in nudged groups, fewer claims would be rejected by the insurance handler. Since there is no relation between nudge assignment and handler assignment, we assume that the handler differences to also be similar across conditions. If a claim gets rejected is coded with a dummy variable.

Fifth, Event Description Length is a measure of the number of characters used by customers when they describe their event of a loss. Since our nudges try to stimulate honest behavior through moral suasion, those can influence how customers write when describing damages. Since the nudges are expected to influence customers to reflect more during filing the claim, we expect longer descriptions to be written.

Though we use several outcome measures to infer claims fraud, some customers *always* commit fraud and some customers *never* commit fraud, regardless of any intervention. Ideally, we could infer more precisely if we could remove the *always* and *never* segments and compare across groups in the "nudge-able" population. Since we randomly assign customers to conditions, the proportion of "nudge-able", always and never committing fraud customers should not differ across conditions. So, changes in outcome variables across conditions should be driven by effects on the "nudge-able" group.

A summary of these measures is outlined in Table 2.

3 | RESULTS

3.1 | Sample

The final dataset we received had cases removed where it was identified that the fidelity of treatment assignment was violated. First, cases, where the same customer logged in multiple times when filing their claim, were removed, because the effects of the honesty-nudges can dissipate over multiple logins and make them incomparable to those with single logins. Second, cases, where the customer switched to a phone call to complete their claim, were removed. This is because our medium of administering the nudge is online and it is aimed to change behavior when filling forms online rather than when claiming over the phone. The final dataset ($N = 5704$) we received for analyses had customers assigned to each condition as in Table 3.

We conducted a randomization check (with customer age, number of previously purchased insurance products, and customer loyalty) to check if customers systematically differed across groups (see Table B1 in Appendix B). In one-way ANOVAs, customers across the eight conditions seemed to have no significant differences in their age F (7, 4692) $= 1.674$, $p = .1105$, and the number of previously purchased insurance products F (7, 4994) $= .629$, $p = .7323$. Customer Loyalty seemed to be significantly different F (7, 4983) $= 2.761$, $p = .0073$ across conditions. However, post hoc Tukey's HSD Test for multiple comparisons showed no statistically significant differences between any pairs at $p < .05$ and 95% confidence intervals (see Table C1 and Figure C1 in Appendix C for details).

3.2 | Main results

Broadly, we found that the nudges do not have significant main effects in reducing indicators of insurance claims fraud. The descriptive statistics on outcome variables are reported in Table 4.

TABLE 2 Proxy measures for claims fraud

Measure	Description and rationale	Expectation
Claimed Amount	The amount claimed for settlement by the customer. Successfully nudged customers would be less likely to report false or exaggerated losses. So, averages in treatment conditions should be lower	Decrease
Claim-Settlement Difference	The difference between the claimed amount and the final settlement amount issued by the provider. Successfully nudged customers would be less likely to report false or exaggerated losses. Meaning they should have smaller differences in claimed amounts and issued settlement amounts. So, averages in treatment conditions should be lower	Decrease
Session Cancelation	If the customer cancels filing the claim. Successfully nudged customers would be more likely to cancel a fraudulent session. So, there should be a higher proportion of canceled sessions in treatment conditions	Increase
Claim Rejection	If the customer's claims are rejected. Successfully nudged customers are less likely to file fraudulent claims. Assuming fraud detection accuracy is similar across conditions, there should be a lower proportion of rejected claims in treatment conditions	Decrease
Event Description Length	The number of characters used to describe damages. All three nudges are designed to make customers think about their choice more. So, successfully nudged customers are expected to write longer descriptions of the incident in nudged conditions. However, this would need to be interpreted with other outcomes	Increase

Note: Mean differences across conditions in multiple outcomes can help shed light on which nudge seems to work and to what extent.

We found no significant differences across conditions in the Pearson's Chi-squared tests for Session Cancelation $\chi^2(7, 5704) = 13.9$, $p = .056$ and Claim Rejection $\chi^2(7, 4332) = 9.1$, $p = .244$). This suggests that the nudges did not dissuade customers from filing fraudulent claims.

The descriptive statistics of outcome measures that were continuous variables (in Table 4) suggest that distributions may not be normal (as the standard deviations are much larger than the means). To check, we made subgroups for each condition and used the Shapiro–Wilk normality test separately for all eight conditions by outcomes Claimed Amount, Claim-Settlement Difference, and Event Description Length. All 24 tests were significant at $p < .001$, suggesting violations of the normality assumption.

As the numbers of customers per group are unequal, we checked if the variances across groups are homogeneous. Levene's test for equality of variances across conditions was found to be not significant for Claimed Amount, F (7, 4942) $= 0.8036$, $p = .5842$; Claim-Settlement Difference, F (7, 4628) $= 0.8475$, $p = .5478$; and Event Description Length, F (7, 4942) $= 1.8821$, $p = .0682$. The non-significant p values of Levene's test suggest that the variances across conditions are not significantly different, thereby satisfying the assumption of homogeneity of variances.

Because of satisfying the homogeneity of variances assumption but violating the normality assumption, we choose non-parametric tests for differences across conditions. First, we used the Kruskal-Wallis rank sum tests-an alternative for one-way ANOVA for non-normal distributions. We found no significant differences across conditions for Claimed Amount, Kruskal-Wallis $\chi^2(7, 4950) = 12.03$, $p = .0996$; and Claim-Settlement Difference, Kruskal-Wallis $\chi^2(7, 4636) = 6.98$, $p = .4311$.

However, we found significant differences for Event Description Length, Kruskal-Wallis $\chi^2(7, 4950) = 39.63$, $p < .0001$, Cohen's d $= .199$. This suggests a small-sized effect of the nudges on description lengths written by customers filing claims.

We then used pairwise comparisons using the Wilcoxon rank-sum test with continuity correction for and Event Description Length. To adjust p values for multiple comparisons, we used the BH method (Benjamin & Hochberg, 1995). This controls the false discovery rate, i.e., the expected proportion of false discoveries among the rejected hypotheses. All experimental conditions showed significant differences at $p < .05$ compared with the control. Event Description Length was significantly longer in all experimental conditions than that in the control condition (see Table 4 for means, and Table D1 for pairwise significance test p values).

Interestingly, event Description Length was also significant longer with customers using more characters in approved claims ($M = 340.22$, SD $= 282.74$) as compared to rejected claims ($M = 265.92$, SD $= 225.28$). This may suggest that some of the claims may have been rejected due to insufficient description. So, we ran a two-way ANOVA[2] with Event Description Length as the

[2]A two-way ANOVA is more robust to normality violations.

TABLE 3 Sample size across conditions

Condition	Baseline	Social norm (SN)	Signing (sign)	Solidarity (sol)	SN × sol	Sign × SN	Sign × sol	Sign × SN × sol
N	1299	626	593	619	633	609	631	694

Note: *N* represents the number of customers assigned to each condition.

TABLE 4 Descriptive statistics of the outcome measures

	Outcome measures				
	Claimed amount *M* (SD) *N*	Claim-settlement difference *M* (SD) *N*	Session cancelation Percentage *N*	Claim rejection Percentage *N*	Description length *M* (SD) *N*
Total	*6115*	*3730*	*18.7%*	*43.7%*	*287*
	(6053)	*(5023)*			*(266)*
	4950	*4636*	*5704*	*4332*	*4950*
Control	*6243*	*3784*	*19.9%*	*40.7%*	*254*
	(5846)	*(5846)*			*(264)*
	1110	*1041*	*1299*	*967*	*1110*
Signing (Sign)	*6208*	*3760*	*15.7%*	*44.8%*	*298*
	(6024)	*(5099)*			*(254)*
	527	*500*	*626*	*469*	*527*
Social Norm (SN)	*5869*	*3444*	*17.9%*	*44.6%*	*298*
	(5140)	*(3884)*			*(268)*
	542	*514*	*633*	*471*	*542*
Solidarity (Sol)	*6130*	*3904*	*17.3%*	*47%*	*282*
	(6112)	*(4660)*			*(250)*
	551	*512*	*619*	*481*	*551*
Sign × SN	*5904*	*3573*	*19%*	*44.6%*	*307*
	(6057)	*(5314)*			*(278)*
	522	*493*	*609*	*464*	*522*
Sign × Sol	*6231*	*3880*	*18.9%*	*46.7%*	*287*
	(6559)	*(5557)*			*(270)*
	551	*512*	*631*	*488*	*551*
SN × Sol	*5796*	*3575*	*22.7%*	*41.3%*	*294*
	(6341)	*(5421)*			*(259)*
	521	*489*	*633*	*458*	*535*
Sign × SN × Sol	*6364*	*3846*	*17.1%*	*42.9%*	*306*
	(6408)	*(4609)*			*(276)*
	612	*575*	*694*	*534*	*612*

Note: M = mean, SD = standard deviation, N = number of customers in the condition.

outcome and Experimental Condition and Claim Rejection as factors. Again, Claim Rejection seemed to be a significant factor, $F\ (1, 4950) = 88$, $p < .0001$. Even so, Condition was still significant, $F\ (7, 4950) = 3.56$, $p < .001$. The interaction between Condition and Claim Rejection was not significant, $F\ (7, 4950) = 1.46$, $p = .18$. Taken together, these suggest that even when controlling for if the claim was rejected, the nudges seem to have elicited longer descriptions than the baseline condition.

The nudges seem to have had some effect on the cognitions of customers, influencing them to write longer descriptions. However, the null results in other measures cast doubt on what this actually suggests. So, we cannot say if longer descriptions indicate more fraud-i.e., customers writing longer descriptions to “appear” more honest, or less fraud-i.e., customers feeling the need to justify their honesty claims more.

3.3 | Exploratory post hoc analyses

The effectiveness of nudges may be influenced by the target's characteristics (Ingendahl et al., 2021). We ran several exploratory analyses but do not make strong claims based on these alone. Using floodlight analyses (Spiller et al., 2013) with the *interactions* package (Long, 2019) in R, we found some suggestions of how effects of the signing-at-the-beginning nudge varied across Customer Age and Customer Loyalty (see Figures F1–F16 in Appendix F for illustrations).

We estimated simple effects of the independent variable (nudges) at all levels of Customer Age and Customer Loyalty, included together and standalone in separate models with outcomes of Claim Amount, Claim-Settlement Difference, Session Cancelation, and Claim Rejection. However, Customer Age and Customer Loyalty were found to be positively correlated, $r\ (4690) = .5358$, $p < .0001$. So, it is difficult to untangle the moderator effects of each. Details from the analyses are included in Appendix E.

Across all outcomes, there seems to be one overall trend. For more long-term customers, it may be that the signing became a signal and reminder of that opportunity that they can self-report fraudulently rather than honestly. This can indicate that long-term customers may feel more entitled to receive a settlement, and the nudge can induce reactance effects. Another reason can be that newer customers may be less familiar with the claiming process with the provider; they may pay more attention to the user interface. And being newer customers, they may be less likely to claim fraudulently because of lesser familiarity.

Though these statements are speculative, the post hoc analyses raise questions to test experimentally on the heterogeneity of honesty-nudge effects.

4 | DISCUSSION

Customer dishonesty is not a rare phenomenon among just a minority of customers. But current strategies to fight dishonesty are still limited (Fombelle et al., 2020). We add to this literature by implementing three honesty-nudges in an insurance claims context and testing them in a high-powered field experiment on real customer behavior. Specifically, we nudged customers by (1) having them sign at the beginning of filing a claim, (2) showing a descriptive social norm message, and (3) showing a solidarity message in the user interface for filing claims. We compared potential fraudulent claims using multiple outcome measures: (1) Claimed Amount, (2) Claim-Settlement Difference, (3) Session Cancelation, (4) Claim Rejection, and (5) Event Description Length.

Overall, the main results suggest that honesty-nudges implemented on the website of the insurance provider did not reduce indicators of insurance claims fraud. Strong preexisting attitudes (Vetter & Kutzner, 2016) and strong prior preferences (Venema et al., 2019) seem to be boundary conditions regarding the effectiveness of nudges. In a similar vein, our results that insurance customers could not be nudged to reduce indicators of claims fraud adds to the side of the debate that those who want to commit insurance fraud may do so even in the presence of the nudge.

Given the recent discoveries and scientific debate on the efficacy of honesty-nudges, this is not surprising. In their 2020 PNAS replication study, Kristal et al. found null main effects on signing-at-the-beginning. Similar null findings were also reported in Dimant et al. (2020)'s study on norm-nudging honesty. In fact, Kristal et al. (2020) went as far as calling it "unlikely" that the signing-at-the-beginning nudge constitutes a simple solution for reducing dishonest self-reporting.

4.1 | Implications for theory

Our study has implications for theory in pointing out a need to revisit mechanisms of nudges to reduce dishonesty. The null results are in line with how Hummel and Maedche (2019) found one-third of the effects of nudges to be non-significant in their quantitative review. Extant research on honesty-nudges also finds null results (Dimant et al., 2020; Kettle et al., 2017), or even negative results (John & Blume, 2018).

Also, in line with signing-at-the-beginning not reducing underreporting miles driven when applying for car insurance (Kristal et al., 2020), our results add evidence to support the view that this nudge may not reduce insurance claims fraud either. This also supports the view that signing-at-the-beginning may not be a viable way to elicit honest self-reporting in online settings in general (Chou, 2015a). We might be overestimating the effects of the saliency of ethics and the importance of self-concept to in reducing customer dishonesty.

Both descriptive social norms and solidarity messages did not seem to reduce indictors of claims fraud. Though previous studies found informational nudges to work in different settings (e.g., Miesler et al., 2017 in raising awareness among young adults), nudges adapted to the insurance claims setting did not seem to reduce indicators of fraud. This further suggests a need to be conservative in our expectations of how much we can nudge customers to behave honestly simply by adding information in the choice architecture.

The finding that customers in the treatment conditions used more characters to describe the events compared to customers in the control condition, can suggest that the manipulations had some effect on cognitions. One possibility can be that the nudges made respondents more self-aware, which in turn caused two different effects in different groups of customers. For consumers in a loss situation, increased self-awareness may have triggered negative emotions and a stronger motivation to be dishonest (Rick & Loewenstein, 2008). For others, the nudges may have had the hypothesized effects and decreased dishonest reporting. Next, we report some signs of heterogeneity of honesty-nudge effects in post hoc analyses.

Several null findings of signing-at-the-beginning may cast doubt on its mechanism, to begin with. Peer et al. (2020) argue that the same nudge which has a positive effect on desired behavior on one individual can have a non-significant or negative effect on another individual. Previous studies (and ours) that tested signing-at-the-beginning nudge have not tested if this nudge *actually* activates ethical considerations. Some customers may forget having signed when they self-report information on latter pages-which can lead to null effects. Another possibility can be that this nudge itself can remind people that they will now have an opportunity to commit fraud by over-reporting claimed amounts.

Although speculative, the post hoc analyses may provide clues that honesty-nudges can lead to reactance among more long-term customers and even backfire. This may be because these customers feel entitled to be treated in a particular way and that an additional signing-at-the-beginning may feel intrusive and encourage false reporting. The feeling of entitlement has been linked to hostile behavior in other settings.

4.2 | Implications for practice

Our study has important implications for businesses and policymakers. In addition to insurance claims, our results can also apply to other settings such as filing tax forms online, returning products to online retailers, credit checks, etc. No combination of honesty-nudges seemed to reduce indicators of insurance fraud. This casts more doubt on being able to nudge customers to self-report honestly online.

Since we found signs of the effects of honesty-nudges depending on characteristics (age) of the targets, A/B testing of which nudge works for which demographic can be a promising starting point. Because honesty-nudges may not be a good idea for all customer demographics. Reducing customer dishonesty may very well be a more iterative process than previously thought of. Schneider et al. (2018) advocate for adapting digital nudges to different targets. So, trial-and-error experimentation can serve well to tease out what works when and what does not. Nudging customers to behave more honestly is still a challenge and a "many-fits-many" approach to interventions may be more fruitful.

4.3 | Limitations and future research

Although the current study is novel in its large-scale and real setting, it is not without limitations. Our nudges were tested in a single product setting-travel insurance. Though claiming mechanisms share similarities across several insurance settings, honesty-nudges can be investigated in other domains to test its effectiveness (or lack thereof).

Due to constraints, we could not pre-test our nudges in the lab before administering them in the field study. The wordings on the information nudges (descriptive social norm and solidarity messages) may not have been interpreted as we intended. For instance, using almost in a descriptive social norm message can suggest there are other ways to behave. It can be interesting to test how different wordings of descriptive norm messages can influence effectiveness.

Future studies may also test honesty-nudges in other online settings such as to reduce returns fraud or shoplifting at self-service checkouts. Even within travel insurance, some types of claims provide more opportunities to file false claims than others. For example, it can be easier to over-report the value of stolen goods than falsify claims for canceled holidays. So, it would also be interesting to explore how the effects of honesty-nudges differ across types of claims.

Our outcome measures look at differences in reports across conditions to infer the likelihood of fraud. More advanced techniques, e.g., machine learning can be used to assign more accurate estimates of fraud likelihood in a case. Also, the measures Clam-Settlement Difference and Claim Rejection may depend on insurance handlers-whose judgments may vary greatly (Kahneman et al., 2021). Future field experiments may also test honesty-nudges where differences in behavior can be measured more accurately at an individual level.

Given the effects of nudges are only applicable to a select portion of customers who may commit fraud, larger sample sizes may be needed to detect effect sizes this small, especially in investigating interaction effects. Though we found some suggestions about the heterogeneity of effects of nudges, we would require even higher statistical power to find stronger evidence. Future field experiments on honesty-nudges may calculate power a priori to account for very small effect sizes as found in our study.

5 | CONCLUSION

With increasing self-service and self-reporting by customers online, nudges have been deemed as powerful tools to stimulate honest behavior. However, studies that report online field experiments on actual customers making real choices are few. We add to the debate of how much customers can be nudged to behave honestly using a large-scale randomized field experiment on the website of a large Nordic insurance provider. We found that nudges do not seem to reduce indictors of insurance claims fraud. However, we found that honesty-nudges can make customers write longer descriptions of their damages. We also speculate that honesty-nudges may influence different customers differently. We urge practitioners to be cautious in using honesty-nudges with a one-size-fits-all approach. Researchers interested in understanding customer dishonesty may find it fruitful to explore moderators that can drive the effectiveness of interventions.

CONFLICT OF INTEREST

The authors declare no conflict of interest with respect to their authorship or publication of this article.

DATA AVAILABILITY STATEMENT

Data available on request due to privacy/ethical restrictions.[3]

ORCID

Jareef Bin Martuza https://orcid.org/0000-0002-6790-5846

REFERENCES

Agerström, J., Carlsson, R., Nicklasson, L., & Guntell, L. (2016). Using descriptive social norms to increase charitable giving: The power of

[3]Involves actual claims data from an insurance company.

local norms. *Journal of Economic Psychology*, *52*, 147–153. https://doi.org/10.1016/j.joep.2015.12.007

Allcott, H. (2011). Social norms and energy conservation. *Journal of Public Economics*, *95*(9–10), 1082–1095. https://doi.org/10.1016/j.jpubeco.2011.03.003

Ayres, I., Raseman, S., & Shih, A. (2013). Evidence from two large field experiments that Peer comparison feedback can reduce residential energy usage. *Journal of Law, Economics, and Organization*, *29*(5), 992–1022. https://doi.org/10.1093/jleo/ews020

Behavioral Insights Team. (2018). The Behavioural Insights Team Annual Update Report 2017–18. Retrieved December 29, 2020, from https://www.bi.team/wp-content/uploads/2019/01/Annual-update-report-BIT-2017-2018.pdf.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)*, *57*(1), 289–300.

Bhanot, S. P. (2017). Cheap promises: Evidence from loan repayment pledges in an online experiment. *Journal of Economic Behavior & Organization*, *140*, 246–266. https://doi.org/10.1016/j.jebo.2017.04.007

Bicchieri, C., & Dimant, E. (2019). Nudging with care: The risks and benefits of social information. *Public Choice*. https://doi.org/10.1007/s11127-019-00684-6

Bicchieri, C., Dimant, E., & Xiao, E. (2021). Deviant or wrong? The effects of norm information on the efficacy of punishment. *Journal of Economic Behavior & Organization*, *188*, 209–235. https://doi.org/10.1016/j.jebo.2021.04.002

Bott, K. M., Cappelen, A. W., Sørensen, E. Ø., & Tungodden, B. (2020). You've got mail: A randomized field experiment on tax evasion. *Management Science*, *66*(7), 2801–2819. https://doi.org/10.1287/mnsc.2019.3390

Brañas-Garza, P. (2007). Promoting helping behavior with framing in dictator games. *Journal of Economic Psychology*, *28*(4), 477–486. https://doi.org/10.1016/j.joep.2006.10.001

Cadario, R., & Chandon, P. (2020). Which healthy eating nudges work best? A meta-analysis of field experiments. *Marketing Science*, *39*(3), 465–486. https://doi.org/10.1287/mksc.2018.1128

Caraban, A., Karapanos, E., Gonçalves, D., & Campos, P. (2019). 23 ways to nudge: A review of technology-mediated nudging in human-computer interaction. *Proceedings of the 2019 CHI conference on human factors in computing systems*, 1–15. https://doi.org/10.1145/3290605.3300733

Castleman, B. L., & Page, L. C. (2015). Summer nudging: Can personalized text messages and peer mentor outreach increase college going among low-income high school graduates? *Journal of Economic Behavior & Organization*, *115*, 144–160. https://doi.org/10.1016/j.jebo.2014.12.008

Chou, E. Y. (2015a). Paperless and soulless: E-signatures diminish the Signer's presence and decrease acceptance. *Social Psychological and Personality Science*, *6*(3), 343–351. https://doi.org/10.1177/1948550614558841

Chou, E. Y. (2015b). What's in a name? The toll e-signatures take on individual honesty. *Journal of Experimental Social Psychology*, *61*, 84–95. https://doi.org/10.1016/j.jesp.2015.07.010

Cialdini, R. B., Demaine, L. J., Sagarin, B. J., Barrett, D. W., Rhoads, K., & Winter, P. L. (2006). Managing social norms for persuasive impact. *Social Influence*, *1*(1), 3–15. https://doi.org/10.1080/15534510500181459

Dimant, E., van Kleef, G. A., & Shalvi, S. (2020). Requiem for a nudge: Framing effects in nudging honesty. *Journal of Economic Behavior & Organization*, *172*, 247–266. https://doi.org/10.1016/j.jebo.2020.02.015

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Federal Bureau of Investigation. (2020). Insurance fraud. Retrieved December 29, 2020 from https://www.fbi.gov/stats-services/publications/insurance-fraud

Festinger, L. (1957). *A theory of cognitive dissonance* (Vol. 2). Stanford university press.

Fombelle, P. W., Voorhees, C. M., Jenkins, M. R., Sidaoui, K., Benoit, S., Gruber, T., Gustafsson, A., & Abosag, I. (2020). Customer deviance: A framework, prevention strategies, and opportunities for future research. *Journal of Business Research*, *116*, 387–400. https://doi.org/10.1016/j.jbusres.2019.09.012

Gerber, A. S., & Rogers, T. (2009). Descriptive social norms and motivation to vote: Everybody's voting and so should you. *The Journal of Politics*, *71*(1), 178–191. https://doi.org/10.1017/S0022381608090117

Gino, F., & Galinsky, A. D. (2012). Vicarious dishonesty: When psychological closeness creates distance from one's moral compass. *Organizational Behavior and Human Decision Processes*, *119*(1), 15–26. https://doi.org/10.1016/j.obhdp.2012.03.011

Gneezy, U. (2005). Deception: The role of consequences. *American Economic Review*, *95*(1), 384–394.

Goldstein, N. J., Cialdini, R. B., & Griskevicius, V. (2008). A room with a viewpoint: Using social norms to motivate environmental conservation in hotels. *Journal of Consumer Research*, *35*(3), 472–482. https://doi.org/10.1086/586910

Hallsworth, M., List, J. A., Metcalfe, R. D., & Vlaev, I. (2017). The behavioralist as tax collector: Using natural field experiments to enhance tax compliance. *Journal of Public Economics*, *148*, 14–31. https://doi.org/10.1016/j.jpubeco.2017.02.003

Hummel, D., & Maedche, A. (2019). How effective is nudging? A quantitative review on the effect sizes and limits of empirical nudging studies. *Journal of Behavioral and Experimental Economics*, *80*, 47–58. https://doi.org/10.1016/j.socec.2019.03.005

Ingendahl, M., Hummel, D., Maedche, A., & Vogel, T. (2021). Who can be nudged? Examining nudging effectiveness in the context of need for cognition and need for uniqueness. *Journal of Consumer Behaviour*, *20*(2), 324–336. https://doi.org/10.1002/cb.1861

Insurance Europe (2019a). The impact of insurance fraud. Retrieved May 28, 2021 from https://www.insuranceeurope.eu/mediaitem/0bf0af82-e7ef-4439-a763-d7862859d421/The%20impact%20of%20insurance%20fraud.pdf

Insurance Europe (2019b). Insurance fraud: Not a victimless crime. Retrieved February 17, 2022 from https://www.insuranceeurope.eu/mediaitem/2bf88e16-0fe2-4476-8512-7492f5007f3c/Insurance%20fraud%20-%20not%20a%20victimless%20crime.pdf

John, P., & Blume, T. (2018). How best to nudge taxpayers? The impact of message simplification and descriptive social norms on payment rates in a Central London local authority. *Journal of Behavioral Public Administration*, *1*(1), 1–11. https://doi.org/10.30636/jbpa.11.10

Kahneman, D., Sibony, O., & Sunstein, C. R. (2021). *Noise: A flaw in human judgment*. Little, Brown.

Kettle, S., Hernandez, M., Sanders, M., Hauser, O., & Ruda, S. (2017). Failure to CAPTCHA attention: Null results from an honesty priming experiment in Guatemala. *Behavioral Sciences*, *7*(4), 28. https://doi.org/10.3390/bs7020028

Kirlidog, M., & Asuk, C. (2012). A fraud detection approach with data mining in health insurance. *Procedia - Social and Behavioral Sciences*, *62*, 989–994. https://doi.org/10.1016/j.sbspro.2012.09.168

Koning, L., Junger, M., & van Hoof, J. (2020). Digital signatures: A tool to prevent and predict dishonesty? *Mind & Society*, *19*(2), 257–285. https://doi.org/10.1007/s11299-020-00237-1

Kristal, A. S., Whillans, A. V., Bazerman, M. H., Gino, F., Shu, L. L., Mazar, N., & Ariely, D. (2020). Signing at the beginning versus at the end does not decrease dishonesty. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(13), 7103–7107. https://doi.org/10.1073/pnas.1911695117

Leal, S., Vrij, A., Nahari, G., & Mann, S. (2016). Please be honest and provide evidence: Deterrents of deception in an online insurance fraud context: Deterrence of deception. *Applied Cognitive Psychology*, *30*(5), 768–774. https://doi.org/10.1002/acp.3252

Lehtonen, T.-K., & Liukko, J. (2011). The forms and limits of insurance solidarity. *Journal of Business Ethics*, *103*(S1), 33–44. https://doi.org/10.1007/s10551-012-1221-x

List, J. A. (2008). Introduction to field experiments in economics with applications to the economics of charity. *Experimental Economics*, *11* (3), 203–212.

Long, J.A. (2019). *Interactions: Comprehensive, user-friendly toolkit for probing interactions*. R package version 1.1.0. https://cran.r-project.org/package=interactions

Miesler, L., Scherrer, C., Seiler, R., & Bearth, A. (2017). Informational nudges as an effective approach in raising awareness among young adults about the risk of future disability: Informational nudges. *Journal of Consumer Behaviour*, *16*(1), 15–22. https://doi.org/10.1002/cb.1592

Morley, N. J., Ball, L. J., & Ormerod, T. C. (2006). How the detection of insurance fraud succeeds and fails. *Psychology, Crime & Law*, *12*(2), 163–180. https://doi.org/10.1080/10683160512331316325

Finanse Norge. (2019). Insurance fraud in Norway. Statistics for 2019. Retrieved November 26, 2021 from https://www.finansnorge.no/siteassets/statistikk/forsikringssvindel—svikrapport/insurance-fraud-in-norway—2020-report.pdf

Peer, E., Egelman, S., Harbach, M., Malkin, N., Mathur, A., & Frik, A. (2020). Nudge me right: Personalizing online security nudges to people's decision-making styles. *Computers in Human Behavior*, *109*, 106347. https://doi.org/10.1016/j.chb.2020.106347

Rick, S., & Loewenstein, G. (2008). *Hypermotivation: Commentary on "the dishonesty of honest people"*. American Marketing Association.

Sadgali, I., Sael, N., & Benabbou, F. (2019). Performance of machine learning techniques in the detection of financial frauds. *Procedia Computer Science*, *148*, 45–54. https://doi.org/10.1016/j.procs.2019.01.007

Schneider, C., Weinmann, M., & vom Brocke, J. (2018). Digital nudging: Guiding online user choices through interface design. *Communications of the ACM*, *61*(7), 67–73. https://doi.org/10.1145/3213765

Schultz, P. W., Nolan, J. M., Cialdini, R. B., Goldstein, N. J., & Griskevicius, V. (2007). The constructive, destructive, and reconstructive power of social norms. *Psychological Science*, *18*(5), 429–434. https://doi.org/10.1111/j.1467-9280.2007.01917.x

Shu, L. L., Mazar, N., Gino, F., Ariely, D., & Bazerman, M. H. (2012). Signing at the beginning makes ethics salient and decreases dishonest self-reports in comparison to signing at the end. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(38), 15197–15200. https://doi.org/10.1073/pnas.1209746109

Spiller, S. A., Fitzsimons, G. J., Lynch, J. G., Jr., & McClelland, G. H. (2013). Spotlights, floodlights, and the magic number zero: Simple effects tests in moderated regression. *Journal of Marketing Research*, *50*(2), 277–288. https://doi.org/10.1509/jmr.12.0420

Tennyson, S. (1997). Economic institutions and individual ethics: A study of consumer attitudes toward insurance fraud. *Journal of Economic Behavior & Organization*, *32*(2), 247–265. https://doi.org/10.1016/S0167-2681(96)00904-3

Tennyson, S. (2008). Moral, social, and economic dimensions of insurance claims fraud. *Social Research*, *75*(4), 1181–1204.

Thaler, R. H., & Sunstein, C. R. (2009). *Nudge: Improving decisions about health, wealth, and happiness*. Penguin.

Tracy, P. E., & Fox, J. A. (1989). A field experiment on insurance fraud in auto body repair. *Criminology*, *27*(3), 589–603.

Venema, T. A. G., Kroese, F. M., De Vet, E., & De Ridder, D. T. D. (2019). The one that I want: Strong personal preferences render the center-stage nudge redundant. *Food Quality and Preference*, *78*, 103744. https://doi.org/10.1016/j.foodqual.2019.103744

Vetter, M., & Kutzner, F. (2016). Nudge me if you can-how defaults and attitude strength interact to change behavior. *Comprehensive Results in Social Psychology*, *1*(1–3), 8–34. https://doi.org/10.1080/23743603.2016.1139390

AUTHOR BIOGRAPHIES

Jareef Bin Martuza is a PhD candidate in the Department of Strategy and Management at the Norwegian School of Economics. He is currently studying customer dishonesty and what to do about it. His other research interests include intergroup bias and behavioral change.

Siv Rosendahl Skard is an Associate Professor of Marketing the department of Strategy and Management at NHH. Skard is affiliated with the Center for Sustainable Business. Her research focuses on consumers' judgment and decision-making, sustainable consumption, and adoption of innovation. Her work has appeared in international journals such as Journal of Business Ethics, Journal of Business Research and Tourism Management.

Lavrans Løvlie is a Partner at PwC in Norway. His work focuses on service design, innovation, and customer experience. He is co-author of "Service Design from Insight to Implementation" and "Service Design for Business", and has published in outlets such as the Service Design Network Journal-Touchpoint.

Helge Thorbjørnsen is Professor of Marketing at the Norwegian School of Economics. His work focuses on behavioral change, technology adoption, and consumer psychology. He has published in outlets such as the Journal of the Academy of Marketing Science, Journal of Experimental Psychology, Journal of Advertising, Journal of Services Research, and Journal of Product Innovation Management.

How to cite this article: Martuza, J. B., Skard, S. R., Løvlie, L., & Thorbjørnsen, H. (2022). Do honesty-nudges really work? A large-scale field experiment in an insurance context. *Journal of Consumer Behaviour*, *21*(4), 927–951. https://doi.org/10.1002/cb.2049

APPENDIX A

Procedure flow

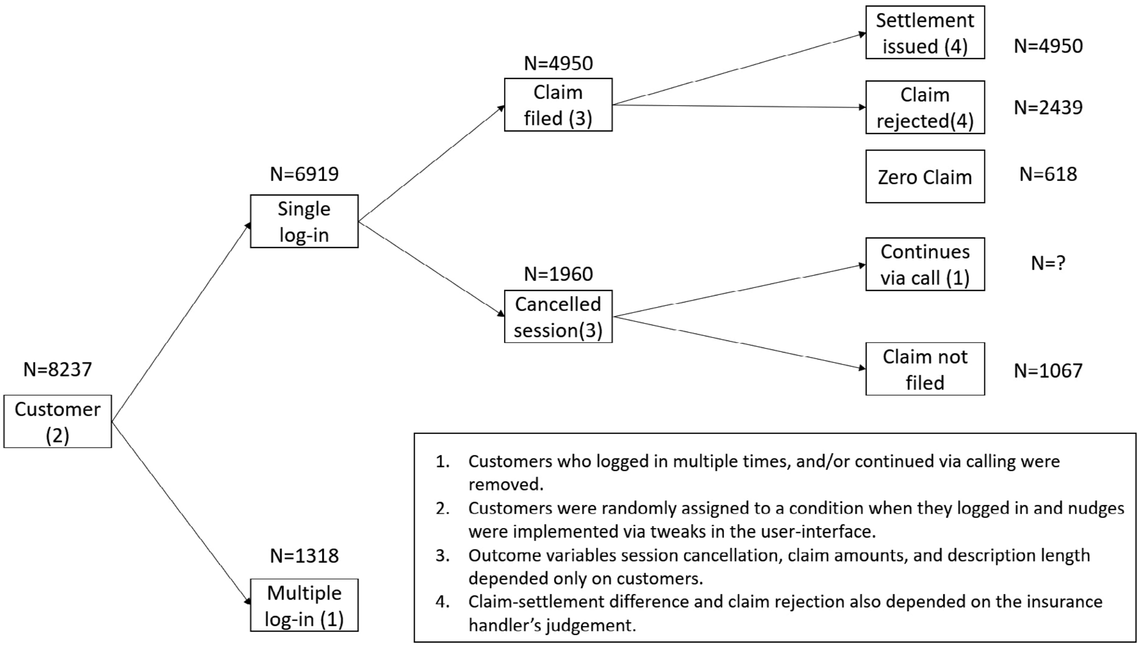


FIGURE A1 Outline of how customers moved through the user interface and measure of outcome variables

APPENDIX B

Randomization check and information on customers in the sample

TABLE B1 Randomization check and information on customers in the sample

	Customer age *M* (SD) *N*	Customer loyalty *M* (SD) *N*	# previous purchases of insurance products *M* (SD) *N*
Total	40.29	7.01	3.82
	(14.82)	(6.17)	(2.41)
	4692	4983	4994
Control	41.00	7.29	3.86
	(14.57)	(6.47)	(2.42)
	1056	1126	1127
Signature (Sign)	39.20	6.49	3.79
	(14.73)	(5.65)	(2.50)
	493	522	522
Social Norm (SN)	39.61	7.10	3.78
	(14.62)	(6.31)	(2.36)
	518	548	548
Solidarity (Sol)	39.41	6.501	3.64
	(14.69)	(5.88)	(2.25)

(Continues)

TABLE B1 (Continued)

	Customer age *M* (SD) *N*	Customer loyalty *M* (SD) *N*	# previous purchases of insurance products *M* (SD) *N*
	520	544	544
Sign × SN	39.75	6.94	3.83
	(14.56)	(6.17)	(2.41)
	491	531	531
Sign × Sol	40.64	6.54	3.88
	(15.16)	(5.63)	(2.55)
	530	562	562
SN × Sol	41.36	7.56	3.89
	(15.31)	(6.49)	(2.35)
	508	535	535
Sign × Sol × SN	40.51	7.31	3.82
	(15.05)	(6.29)	(2.48)
	576	615	615
Test for the independence of outcomes and experimental conditions	$F = 1.674$	$F = 2.761$**	$F = .629$

Note: Though customer loyalty seems to be significantly different, post hoc Tukey's HSD does not show any pairs being significantly different at $p < .05$.
**$p < .05$.

APPENDIX C

Results of Tukey's HSD comparisons of means of the number of previously purchased insurance products at 95% family-wise confidence interval

TABLE C1 Tukey's HSD multiple comparisons of means of the number of previously purchased insurance products at 95% family-wise confidence interval

	Mean differences	Lower interval	Upper interval	Adjusted *p* value
Condition				
Sign_SN-Control	−0.35099	−1.33464	0.632667	.960664
Sign_SN_Sol-Control	0.01965	−0.91724	0.956544	1
Sign_Sol-Control	−0.74907	−1.71412	0.215971	.264856
Signature-Control	−0.79744	−1.78684	0.191963	.220449
SN_Sol-Control	0.300453	−0.6807	1.281605	.983358
Social_norm-Control	−0.18268	−1.15591	0.790552	.999211
Solidarity-Control	−0.78523	−1.76087	0.190401	.222049
Sign_SN_Sol-Sign_SN	0.370638	−0.73625	1.47753	.972288
Sign_Sol-Sign_SN	−0.39808	−1.5289	0.732733	.963452
Signature-Sign_SN	−0.44645	−1.59812	0.705225	.938989
SN_Sol-Sign_SN	0.651441	−0.49316	1.796037	.670353
Social_norm-Sign_SN	0.168309	−0.9695	1.306122	.999837
Solidarity-Sign_SN	−0.43425	−1.57412	0.705624	.944264
Sign_Sol-Sign_SN_Sol	−0.76872	−1.85911	0.321664	.390608
Signature-Sign_SN_Sol	−0.81709	−1.92909	0.294914	.334921

TABLE C1 (Continued)

	Mean differences	Lower interval	Upper interval	Adjusted *p* value
SN_Sol-Sign_SN_Sol	0.280803	−0.82387	1.385472	.994554
Social_norm-Sign_SN_Sol	−0.20233	−1.29997	0.895311	.999299
Solidarity-Sign_SN_Sol	−0.80488	−1.90466	0.294887	.340208
Signature-Sign_Sol	−0.04837	−1.18418	1.087454	1
SN_Sol-Sign_Sol	1.049525	−0.07912	2.178167	.090358
Social_norm-Sign_Sol	0.566393	−0.55537	1.688156	.790758
Solidarity-Sign_Sol	−0.03616	−1.16001	1.087686	1
SN_Sol-Signature	1.09789	−0.05165	2.247428	.073576
Social_norm-Signature	0.614758	−0.52803	1.757542	.731402
Solidarity-Signature	0.012203	−1.13263	1.157035	1
Social_norm-SN_Sol	−0.48313	−1.61878	0.652519	.902855
Solidarity-SN_Sol	−1.08569	−2.2234	0.052025	.074066
Solidarity-Social_norm	−0.60256	−1.73344	0.528332	.741023

TABLE G1 Achieved power of the significant tests to compare differences across groups across outcomes

Outcome measures	Claimed amount	Claim-settlement difference	Session cancelation	Claim rejection	Event description length
Significance Tests	Kruskal-Wallis	Kruskal-Wallis	Chi-squared	Chi-squared	Kruskal-Wallis
Effect sizes	Eta squared	Eta squared	Phi	Phi	Eta squared
	0.00102	−0.0000043	0.04936	0.0458	0.0098
Total sample size	4950	4636	5704	4332	4950
Power ($1-\beta$ err prob)	0.3084	0.2778	0.7846	0.5653	0.9822

Caution: The calculated power for claimed amount, claim-settlement difference, and event description length are approximations based on the scenario if assumptions of one-way ANOVA were met.

APPENDIX D

Results of the pairwise significant tests of the conditions on the outcome of event description length

TABLE D1 Testing for treatment effects on event description length with pairwise comparisons using the Wilcoxon rank-sum test with continuity correction

	Control	(1)	(2)	(3)	(4)	(5)	(6)
Sign_SN (1)	**0.00009**						
Sign_SN_Sol (2)	**0.00009**	0.99					
Sign_Sol (3)	**0.0376**	0.39	0.39				
Sign (4)	**0.0004**	0.93	0.93	0.48			
SN_Sol (5)	**0.0005**	0.93	0.93	0.53	0.93		
SN (6)	**0.0001**	0.93	0.93	0.48	0.98	0.93	
Sol (7)	**0.0151**	0.48	0.48	0.93	0.58	0.71	0.53

Note: Values represent the *p* value (bolded when significant at <.05) when testing for differences in two pairs of conditions. The reported *p* values have been corrected for multiple tests. All treatment conditions seem to generate longer description lengths among customers. See Table 4 for means across treatments.

APPENDIX E

Documentation of floodlight analyses for the moderation of the effects of the signing-at-the-beginning nudge by customer age and customer loyalty

A floodlight analysis uses versions of the Johnson-Neyman probing tests to probe regions of significance. The region of significance refers to the range of values of the moderator (e.g., customer age or customer loyalty) for which the simple effects of an independent variable (e.g., nudges) are significant.

For Claimed Amount, floodlight analyses found that Customer Age, with a range of [18, 91] significantly moderated the effect of signing-at-the-beginning outside the interval [10.04, 47.06]. The nudge had a negative effect (but was non-significant) on claimed amounts for customers over the age of 32. However, the nudge had a significant *positive* effect on Claimed Amount by customers aged 48 and above. Similarly, Customer Loyalty, having a range of [0, 41], was a significant moderator outside the interval [−14.21, 13.01]. Though the effect of Customer Loyalty was non-significant when both interactions were included, Customer Age × signing-at-the-beginning was again significant and positive inside the interval [55.96, 86.22].

For Claim-Settlement Difference, Customer Age was a significant moderator ($p < .05$) outside the interval [20.73, 47.37]. Also, Customer Loyalty was a significant moderator ($p < .05$) outside the interval [−0.53, 10.68]. However, when both interaction terms are included, they were non-significant.

For Session Cancelation, Customer Age was a significant moderator ($p < .05$) outside the interval [−7.40, 57.88]. Though the effect of Customer Loyalty became non-significant when both Customer Age and Customer Loyalty interactions were estimated, Customer Age × signing-at-the-beginning was again significant and positive outside the interval [12.79, 51.72].

For Claim Rejection, Customer Age was a significant moderator ($p < .05$) outside the interval [−263.45, 47.54]. However, Customer Loyalty did not have any significant regions. When both interaction terms were included, Customer Age was significant ($p < .05$) inside the interval [49.28, 87.63].

APPENDIX F

Floodlight analyses on the effects of signing-at-the-beginning across customer age and relationship duration

A. Claimed amount

Customer age and duration are estimated separately.

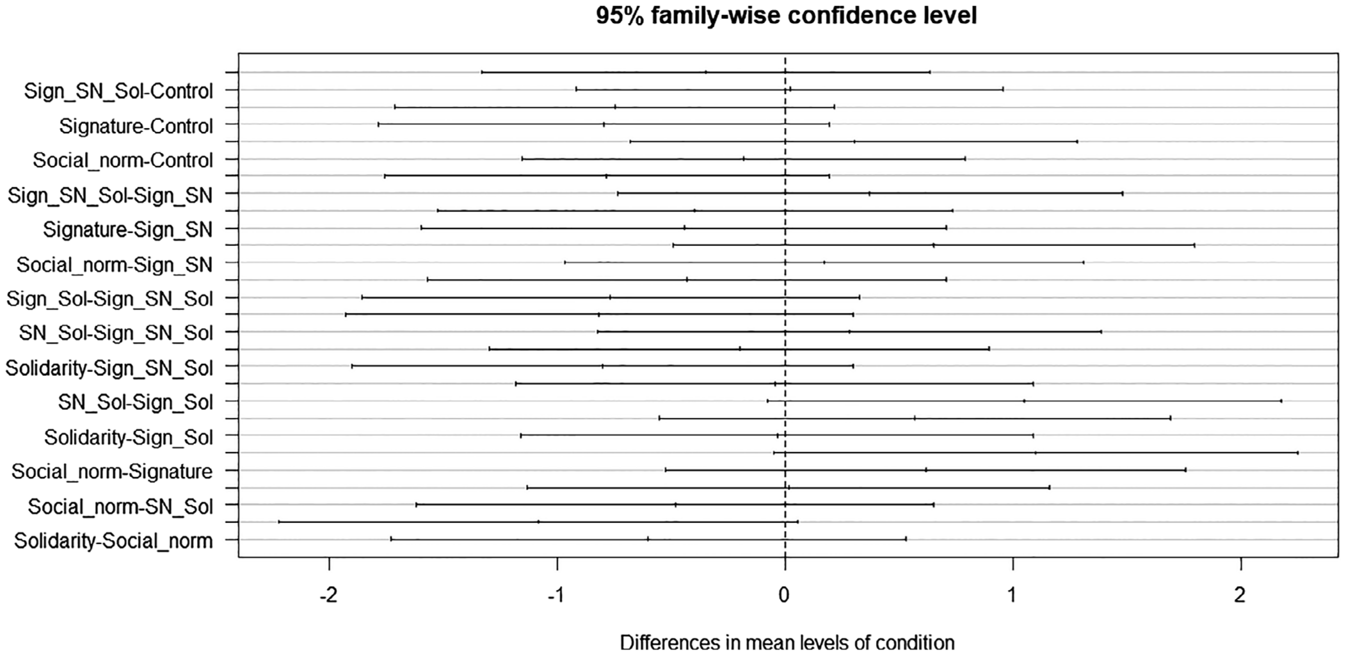


FIGURE C1 Tukey's HSD test shows no significant mean differences in the number of previously purchased insurance products among any of the pairs at $p < .05$ and 95% confidence intervals

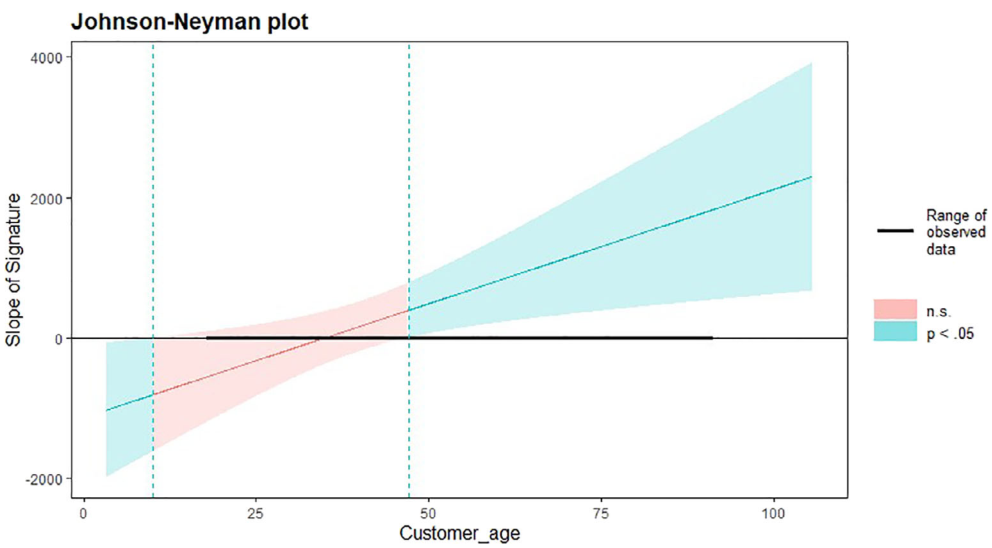
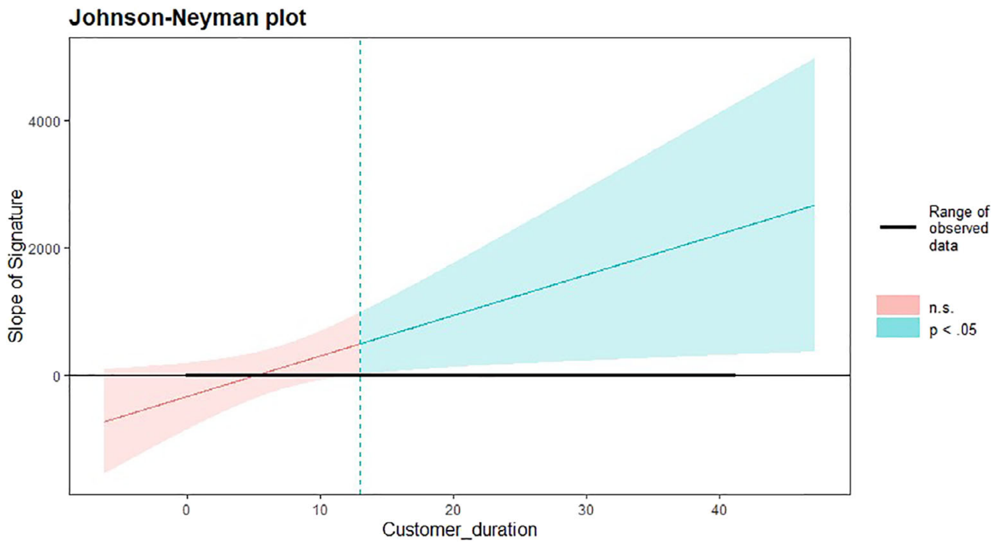


FIGURE F1 Increased claimed amounts for customers above 48 [Colour figure can be viewed at wileyonlinelibrary.com]

FIGURE F2 Increased claimed amounts for customer loyalty above 13 years [Colour figure can be viewed at wileyonlinelibrary.com]



Customer age and duration are estimated together.

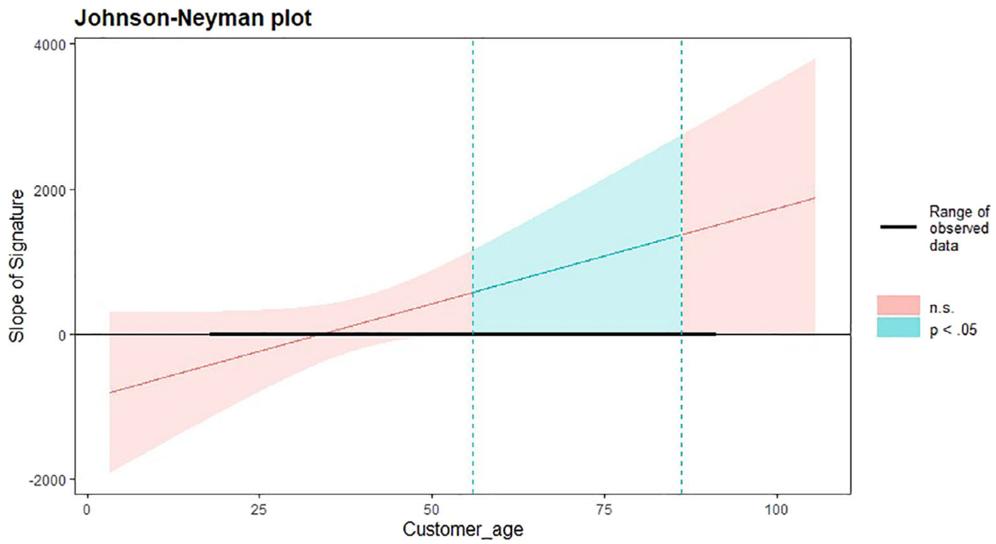


FIGURE F3 Signs of more long-term customers backfiring even when controlling for the duration [Colour figure can be viewed at wileyonlinelibrary.com]

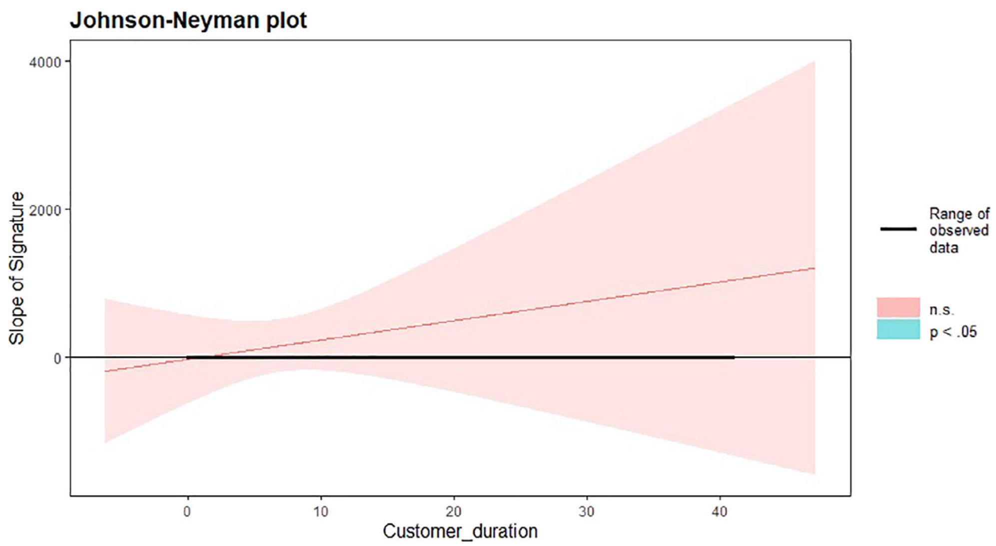


FIGURE F4 Moderation by customer loyalty disappears when controlling for age [Colour figure can be viewed at wileyonlinelibrary.com]

B. Claim-settlement difference

Customer age and duration are estimated separately.

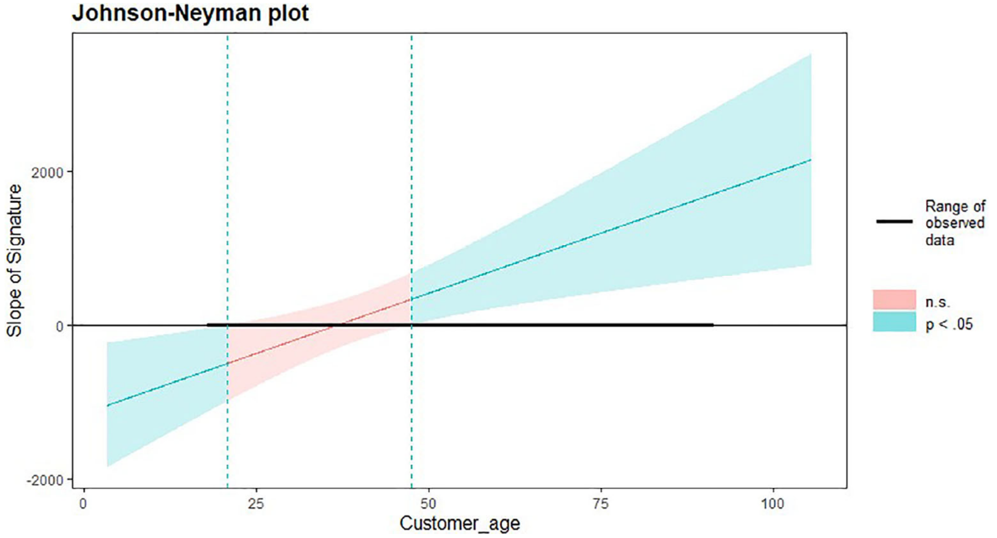


FIGURE F5 Increased claim-settlement differences for customers above 48 [Colour figure can be viewed at wileyonlinelibrary.com]

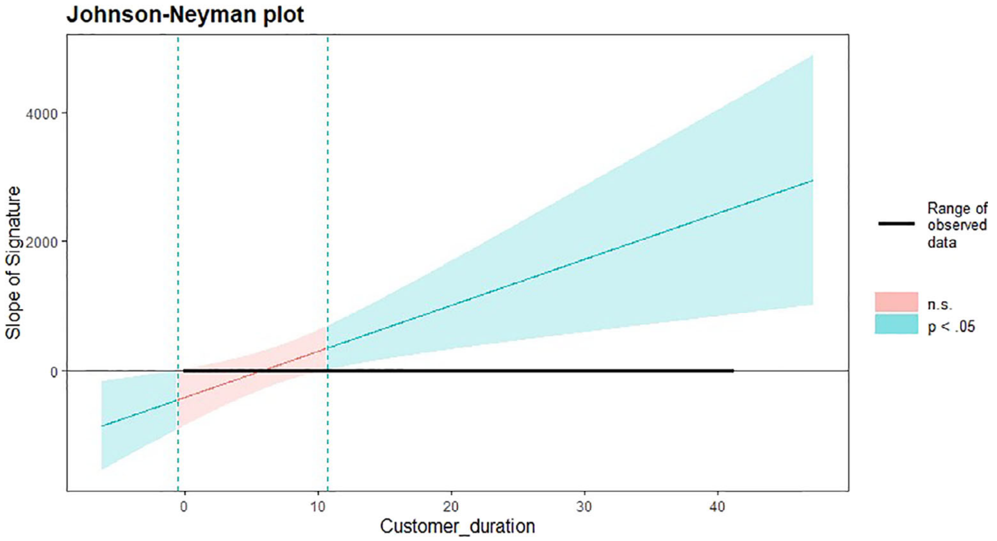


FIGURE F6 Increased claimed amounts for customer loyalty above 11 years [Colour figure can be viewed at wileyonlinelibrary.com]

Customer age and duration are estimated together.

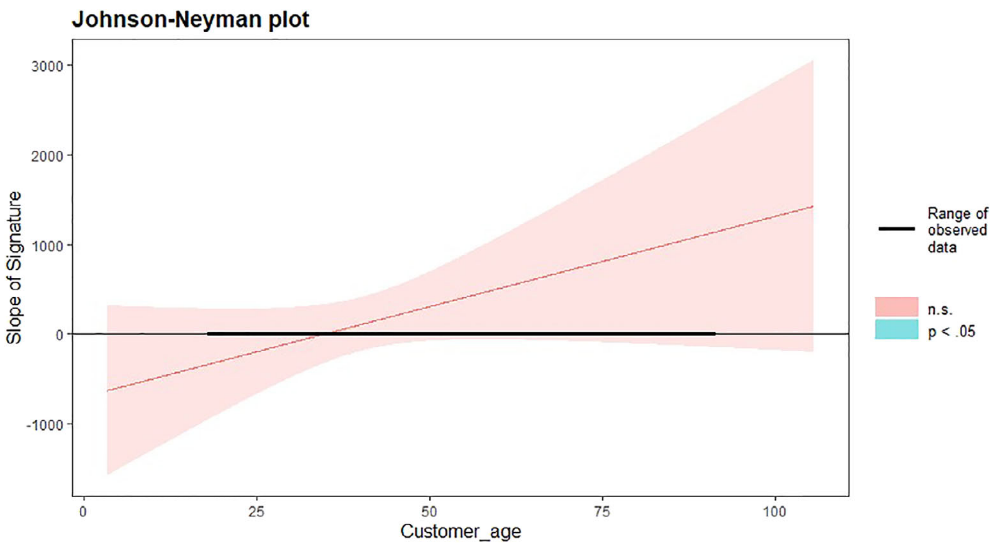


FIGURE F7 Moderation by customer loyalty disappears when estimating both [Colour figure can be viewed at wileyonlinelibrary.com]

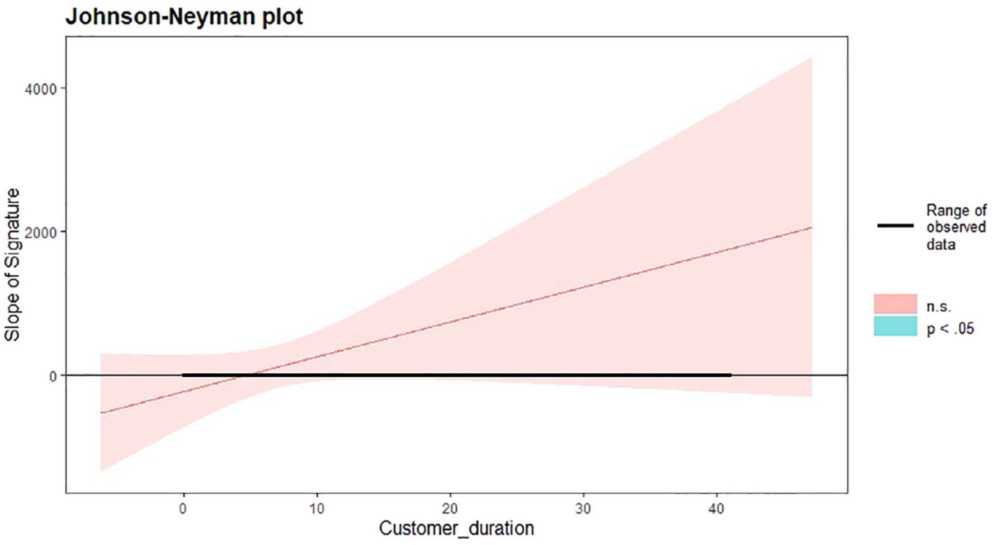


FIGURE F8 Moderation by customer loyalty disappears when estimating both [Colour figure can be viewed at wileyonlinelibrary.com]

C. Session cancelation

Customer age and duration are estimated separately.

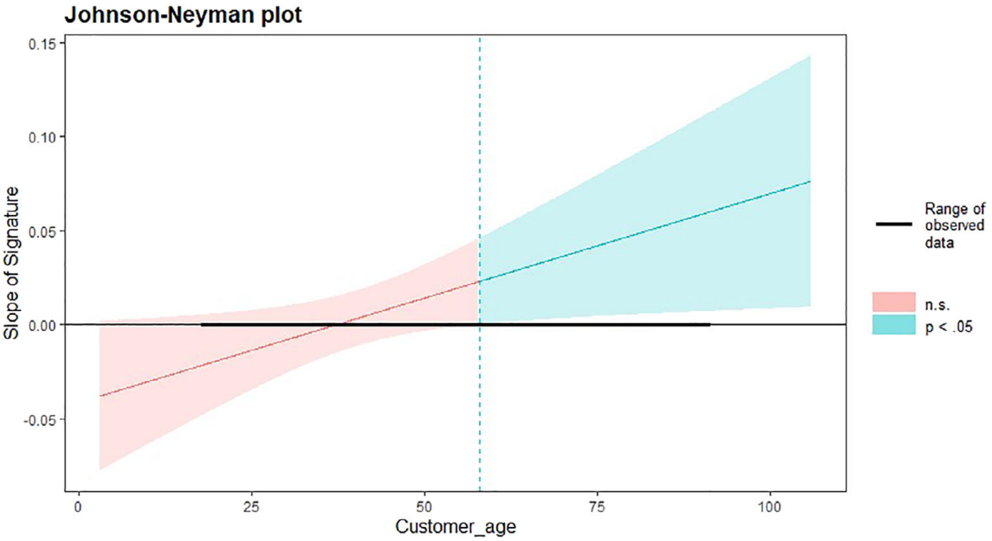


FIGURE F9 Increased session cancelations for customers above 57 [Colour figure can be viewed at wileyonlinelibrary.com]

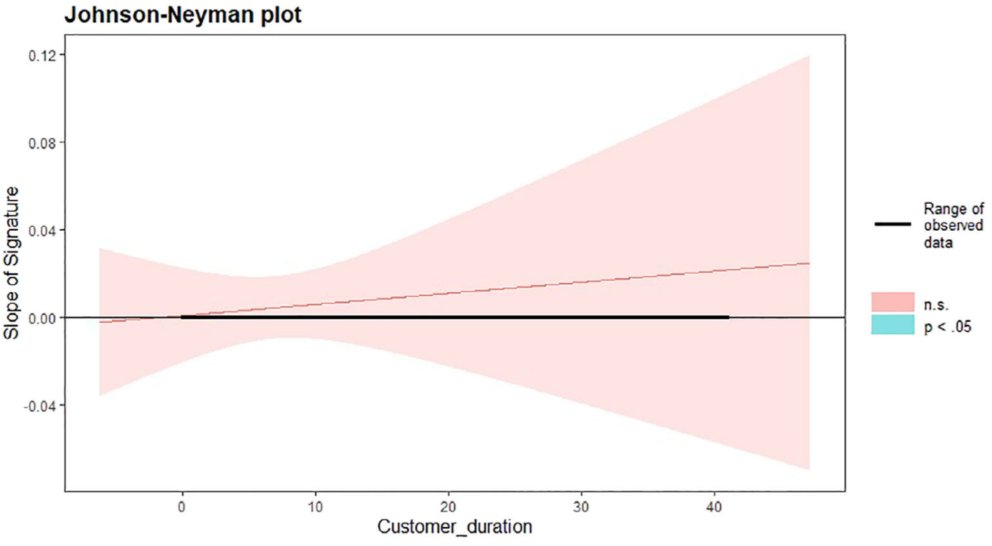


FIGURE F10 No region of significance for session cancelations across duration [Colour figure can be viewed at wileyonlinelibrary.com]

Customer age and duration are estimated together.

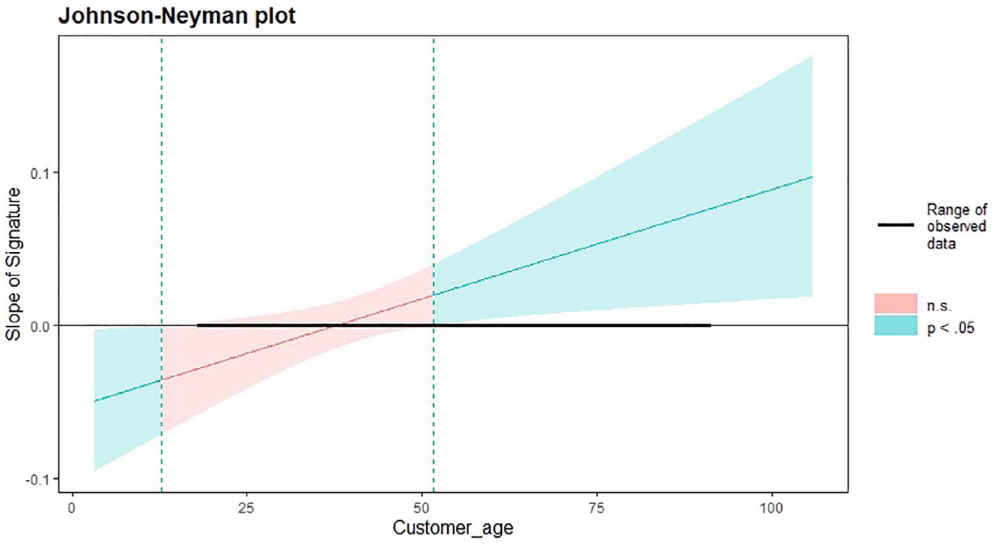


FIGURE F11 Increased session cancelations for customers above 52 when controlling for the duration [Colour figure can be viewed at wileyonlinelibrary.com]

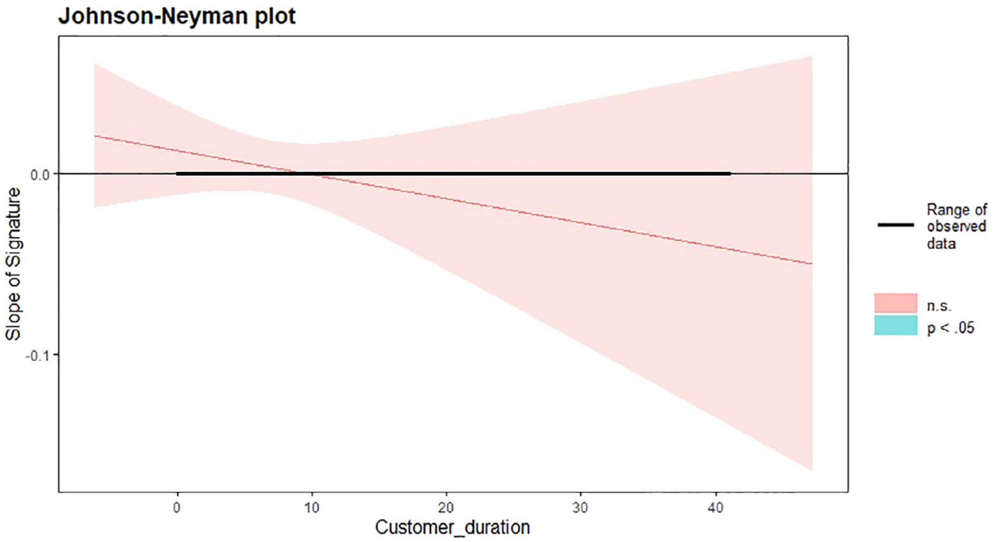


FIGURE F12 No region of significance across duration when controlling for age [Colour figure can be viewed at wileyonlinelibrary.com]

D. Claim rejection

Customer age and duration are estimated separately.

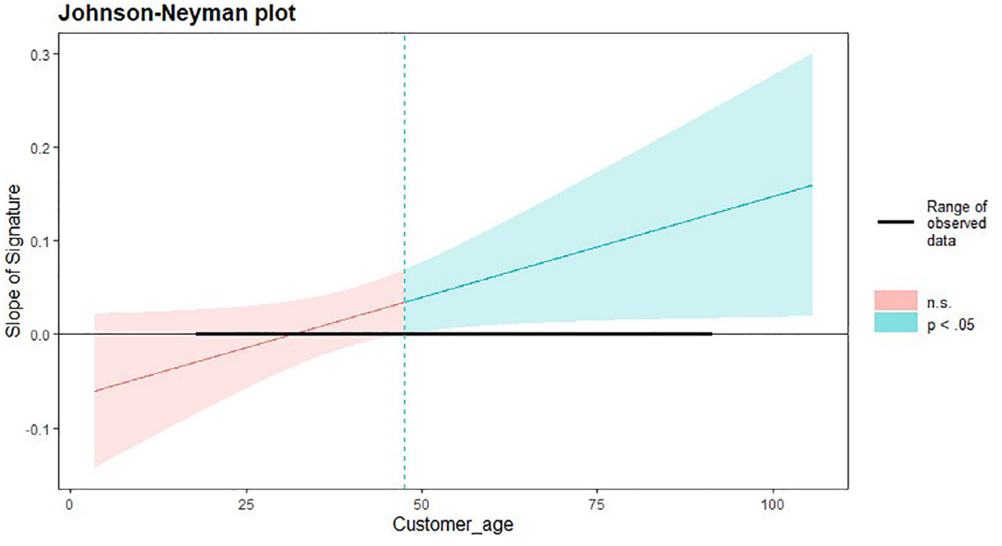


FIGURE F13 Increased claim rejections for customers above 48 [Colour figure can be viewed at wileyonlinelibrary.com]

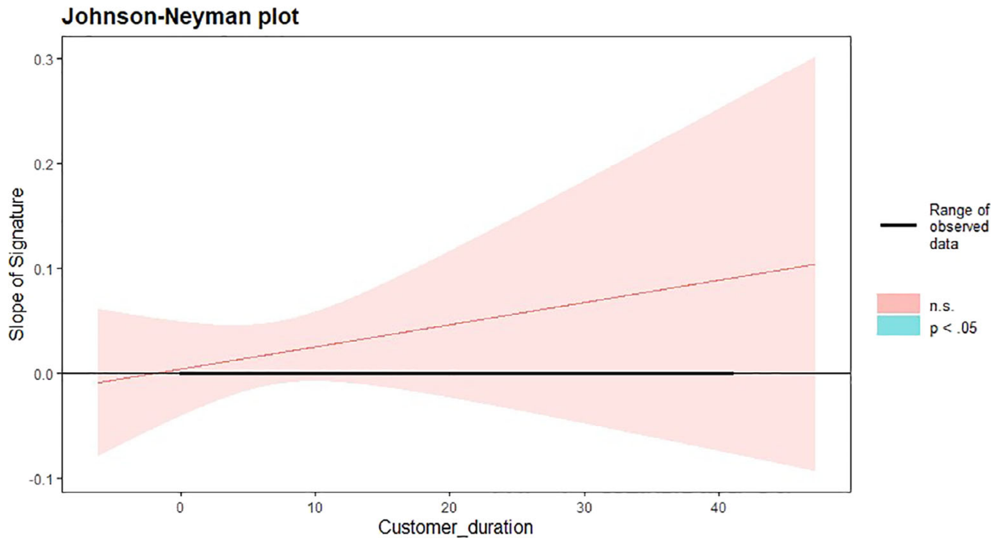


FIGURE F14 No region of significance for claim rejection across duration [Colour figure can be viewed at wileyonlinelibrary.com]

Customer age and duration are estimated together.

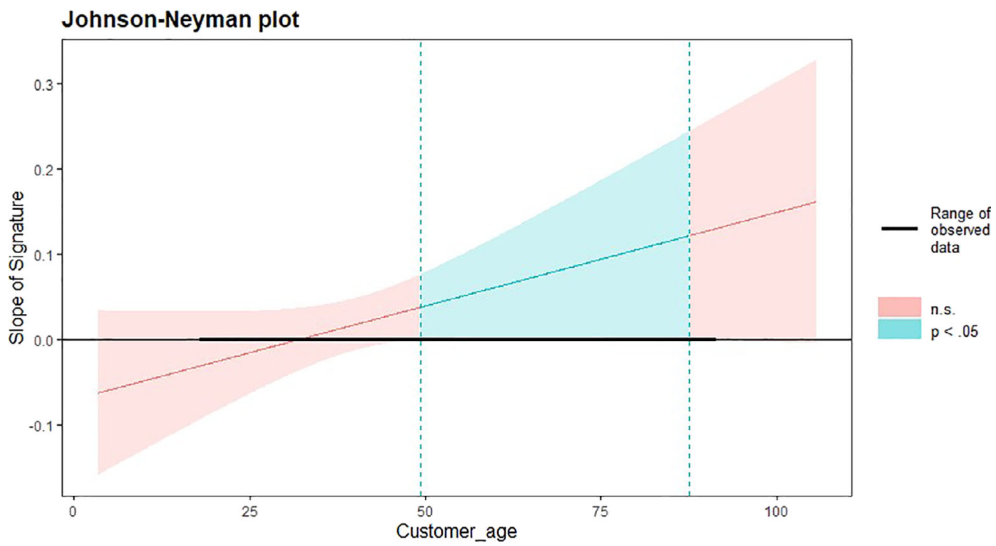


FIGURE F15 Increased rejection rates for customers above 49 when controlling for the duration [Colour figure can be viewed at wileyonlinelibrary.com]

APPENDIX G

Effect sizes and power with α err prob at .05

Note that post hoc power values are a transformation of *p* values. High *p* values correspond to low power, and vice versa. This can give indications to future field studies regarding probable effect sizes and plan sample sizes accordingly (see Table G1 and Figures G1 and G2).

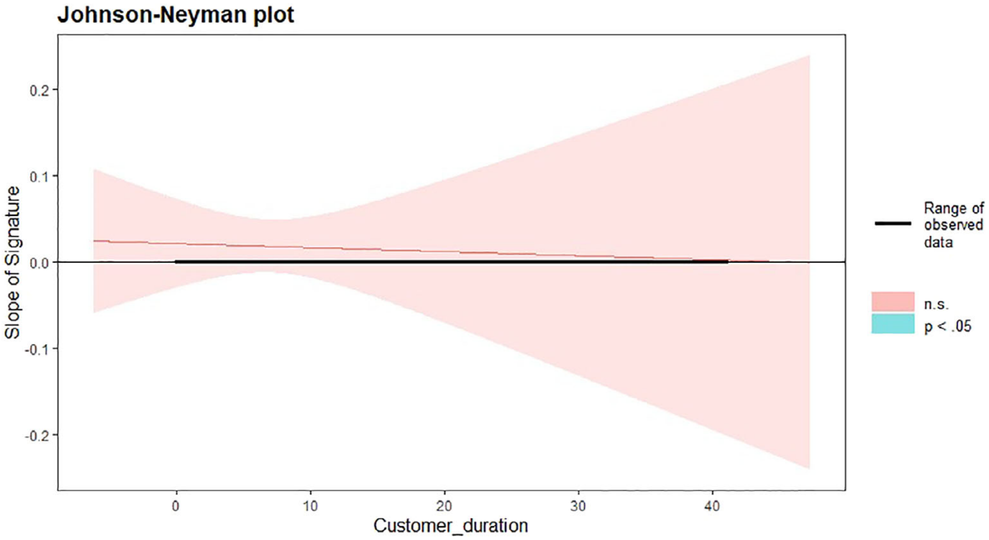


FIGURE F16 No region of significance across duration when controlling for age [Colour figure can be viewed at wileyonlinelibrary.com]

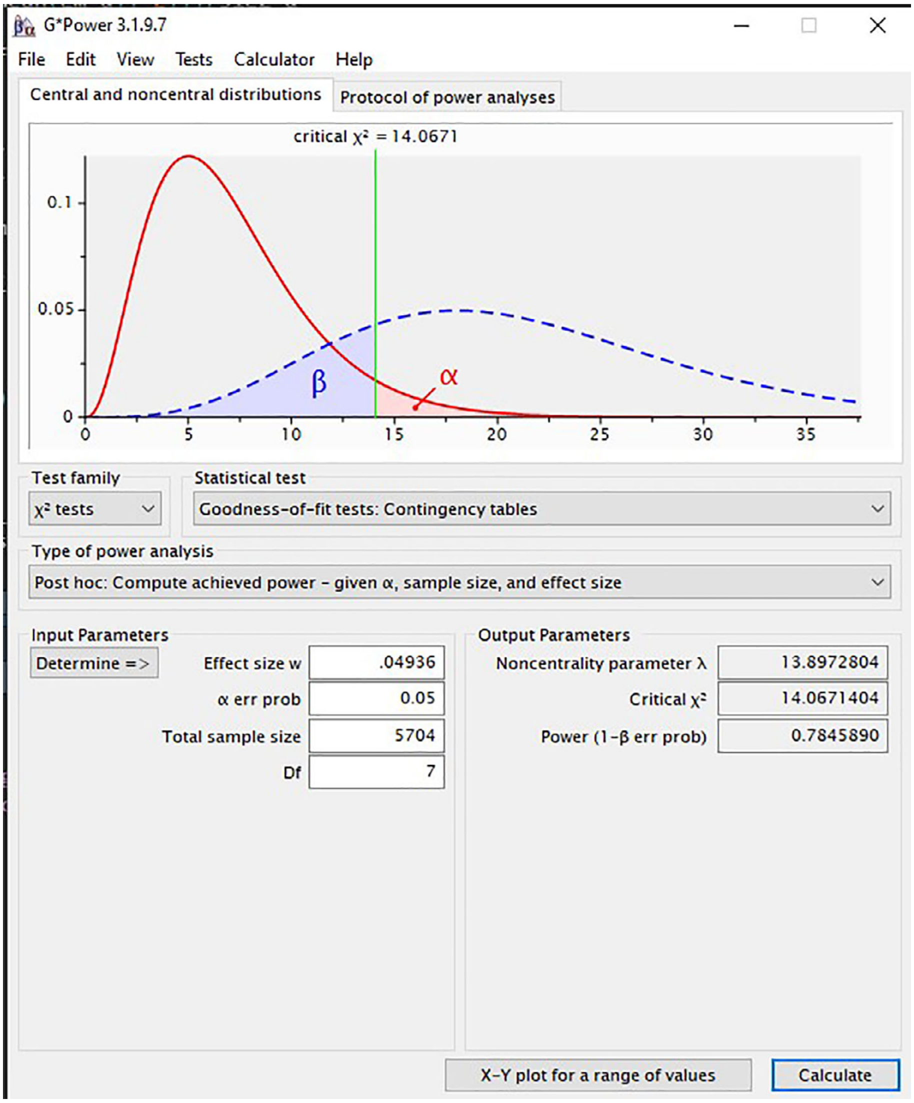


FIGURE G1 Post hoc achieved power for chi-squared tests with session cancelation as the outcome. Power = 0.7846. Computed using G*power 3 (Faul et al., 2007) [Colour figure can be viewed at wileyonlinelibrary.com]

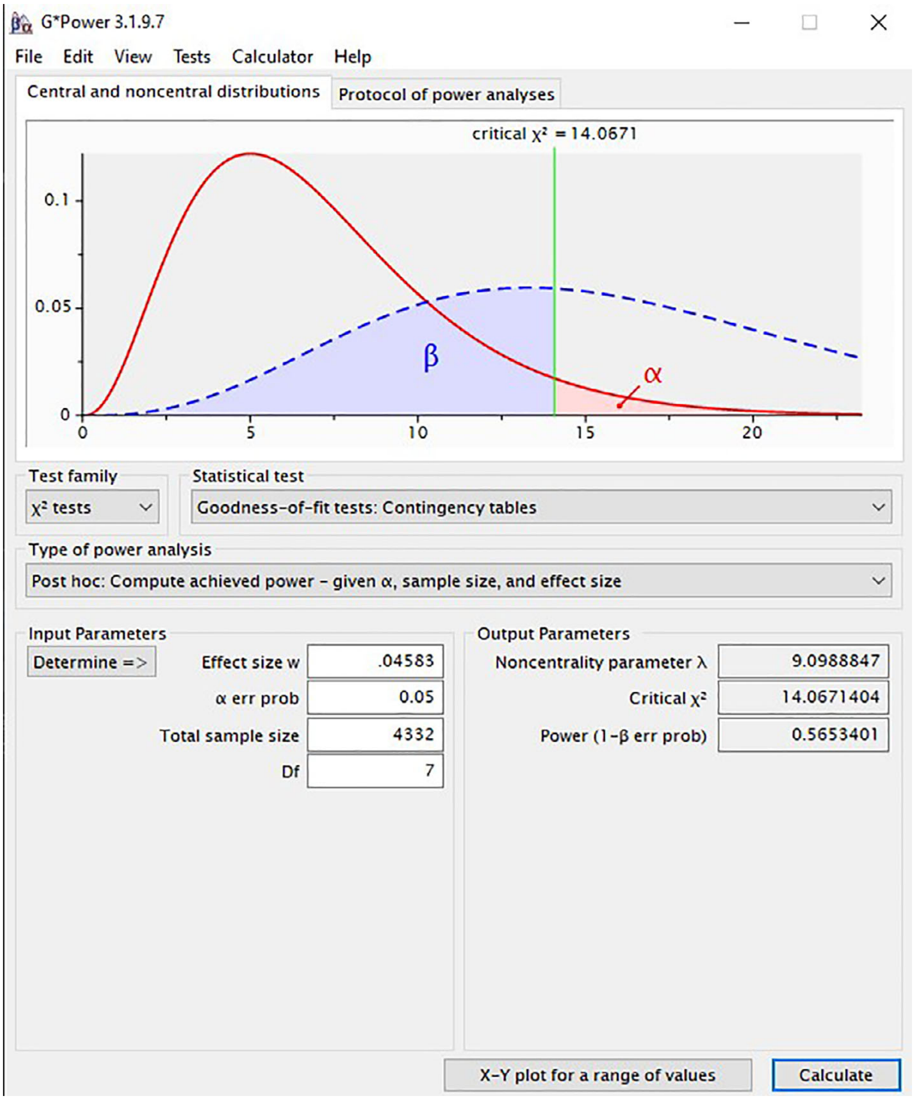


FIGURE G2 Post hoc achieved power for chi-squared tests with claim rejection as the outcome. Power = 0.5653. Computed using G*power 3 (Faul et al., 2007) [Colour figure can be viewed at wileyonlinelibrary.com]

APPENDIX H

Calculation codes of effect sizes and power calculations

This can be directly pasted on an R script and run to retrieve the values [Colour figure can be viewed at wileyonlinelibrary.com].

```
###### Calculations of effect sizes ########
############################################

# calculating effect size for Session Cancellation
phi_cancel <- sqrt(13.9/5704)


# calculating effect size for Claim Rejection
phi_reject <- sqrt(9.1/4332)

# calculating effect size for Claim_Amount
eta2_claim <- ((12.03 - 8 + 1)/ (4950 - 8))


# calculating effect size for Claim-Settlement Difference
eta2_dif <- ((6.98 - 8 + 1)/ (4636 - 8))


# calculating effect size for Event Description Length
eta2_desc <- ((39.63 - 8 + 1)/ (4950 - 8))


###### Calculations of post hoc power for one-way ANOVA scenarios ########
############################################

# setting alpha at .05 and number of groups at 8
a <- .05; i = 8

# calculating power for Claim_Amount
n_c <- c(1110, 527, 542, 551, 522, 551, 521, 612) # cell sizes
gm_c <- c(6243, 6208, 5869, 6130, 5904, 6231, 5796, 6364) # means
sig2_c <- 6053^2
lambda_c <- (sum(n_c*((gm_c-mean(gm_c))^2)))/sig2_c

claim_pwr <- 1 - pf(qf(1-a,i-1, sum(n_c-1)), i-1, sum(n_c-1),lambda_c)

# calculating power for Claim-Settlement Difference
n_d <- c(1041, 500, 514, 512, 493, 512, 489, 575) # cell sizes
gm_d <- c(3784, 3760, 3444, 3940, 3573, 3880, 3575, 3846) # means
sig2_d <- 5023^2
lambda_d <- (sum(n_d*((gm_d-mean(gm_d))^2)))/sig2_d

dif_pwr <- 1 - pf(qf(1-a,i-1, sum(n_d-1)), i-1, sum(n_d-1),lambda_d)

# calculating power for Event_Length_Description
n_t <- c(1110, 527, 542, 551, 522, 551, 535, 612) # cell sizes
gm_t <- c(254, 298, 298, 282, 307, 287, 294, 306) # means
sig2_t <- 266^2
lambda_t <- (sum(n_t*((gm_t-mean(gm_t))^2)))/sig2_t

text_pwr <- 1 - pf(qf(1-a,i-1, sum(n_t-1)), i-1, sum(n_t-1),lambda_t)
```