Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

From Contextualizing to Context Theorizing: Assessing Context Effects in Privacy Research

Heng Xu, Nan Zhang

To cite this article:
Heng Xu, Nan Zhang (2022) From Contextualizing to Context Theorizing: Assessing Context Effects in Privacy Research. Management Science

Published online in Articles in Advance 31 Jan 2022

. https://doi.org/10.1287/mnsc.2021.4249

http://pubsonline.informs.org/journal/mnsc

MANAGEMENT SCIENCE
Articles in Advance, pp. 1–19
ISSN 0025-1909 (print), ISSN 1526-5501 (online)

From Contextualizing to Context Theorizing: Assessing Context Effects in Privacy Research

Heng Xu,[a] **Nan Zhang**[a]

[a] Kogod School of Business, American University, Washington, DC 20016
Contact: xu@american.edu, https://orcid.org/0000-0001-5642-6543 (HX); nzhang@american.edu, https://orcid.org/0002-0454-7885 (NZ)

Received: September 3, 2020
Revised: January 18, 2021; June 17, 2021
Accepted: August 30, 2021
Published Online in Articles in Advance: January 31, 2022

https://doi.org/10.1287/mnsc.2021.4249

Abstract. Over the past two decades, behavioral research in privacy has made considerable progress transitioning from acontextual studies to using contextualization as a powerful sensitizing device for illuminating the boundary conditions of privacy theories. Significant challenges and opportunities wait, however, on elevating and converging individually contextualized studies to a *context-contingent theory* that explicates the mechanisms through which contexts influence consumers' privacy concerns and their behavioral reactions. This paper identifies the important barriers occasioned by this lack of context theorizing on the generalizability of privacy research findings and argues for accelerating the transition from the contextualization of individual research studies to an integrative understanding of context effects on privacy concerns. It also takes a first step toward this goal by providing a conceptual framework and the associated methodological instantiation for assessing how context-oriented nuances influence privacy concerns. Empirical evidence demonstrates the value of the framework as a diagnostic device guiding the selection of contextual contingencies in future research, so as to advance the pace of convergence toward context-contingent theories in information privacy.

History: Accepted by Anindya Ghose, information systems.
Funding: The authors were supported in part by the National Science Foundation [Grants 1850605 and 1851637], and the Defense Advanced Research Projects Agency [Grant HR00111920023].
Supplemental Material: The data files and online appendix are available at https://doi.org/10.1287/mnsc.2021.4249.

Keywords: information privacy • context • behavioral research methods

1. Introduction

Understanding consumers' desires for privacy has emerged as a task front and center for firms and policy makers in addressing the rapidly growing collection of consumer data in today's digitally connected world. Recent privacy laws such as the California Consumer Privacy Act and the European Union's General Data Protection Regulation are frequently attributed as regulatory responses to consumers' need for privacy protection. On the research side, a rich literature has emerged across multiple disciplines (e.g., information systems, marketing, economics, psychology, and computer science), aiming to understand the factors that drive people's desires for privacy and their related behavior (e.g., Smith et al. 2011; Acquisti et al. 2015, 2016, 2020).

This rich literature also highlighted the complexity of consumers' privacy desires (Acquisti et al. 2013) and noted the importance of conceptualizing privacy in a contextualized manner (Nissenbaum 2009): for example, by implicating a particular *context* in studying consumers' attitudes, beliefs, and perceptions toward the disclosure of their private information (e.g., Solove 2002, Xu et al. 2012). Although having a specific context is important for measuring a "fluid" construct like the cognition and perceptions of privacy (Acquisti et al. 2015, 2016), the limitations of a context-specific approach should also be noted. First, most context-specific privacy studies focused on one or a few contexts (e.g., online shopping in one study, social media in another). This defers the comparisons between findings from different contexts, and the development of broad-range context-contingent theories, to the time when a sufficient number of context-specific studies have been accumulated to allow for theory-grounded meta-analyses. Second, in many research studies, researchers have the freedom to choose from numerous possible contexts (e.g., by adjusting the type of personal information (emails or online purchase records), the entity posing privacy threats (online advertisers or social media websites), etc.). Ideally, decisions as to whether a particular context is appropriate for a study should be grounded in theory. Yet, in the extant literature, the specification of a context is often only

recorded in a post hoc descriptive fashion rather than theoretically designed or empirically examined. This, once again, makes it rather slow to develop a theoretical framework that can guide the selection of contexts in future research.

Given the significant interest by firms and policy makers in understanding consumers' privacy desires, the absence of a theoretical framework guiding contextualization represents a considerable gap in the literature and is thus the focus of this paper. Specifically, the objective of the paper is to examine the multiplicity of contexts and their impact on consumers' cognition and perceptions of privacy, identify challenges researchers may face in contextualizing privacy research, and offer possible ways forward. Because our goal was to explicate how a context could affect a wide range of attitudinal and perceptual constructs about privacy, we followed Smith et al. (2011) in using *privacy concerns* as an umbrella term that encompasses the "beliefs, attitudes, [and] perceptions" (Smith et al. 2011, p. 998) about privacy. Conceptually, we drew from the situational strength theory (Mischel 1977, Meyer et al. 2020) to develop a two-dimensional framework for explicating how two pronounced effects of context, *range restriction* and *situational uncertainty*, alter the grounds on which people ascribe meanings to "privacy concerns." To operationalize this framework as a diagnostic device for contextualizing privacy studies, we drew from a recent methodological advance in psychology (Finch and Hernández Finch 2020) and sociology (Iannario et al. 2020) for the statistical analysis of self-reported data, the *combination of (discrete) uniform and (shifted) binomial distribution* (CUB) model[1] (Piccolo and Simone 2019), and adapted it in privacy research to develop a quantitative framework for assessing the magnitude of the two effects occasioned by a given context. After illustrating the usage of the framework with an empirical demonstration, we conclude the paper with a summary of the potential applications of our framework and a set of recommended practices it entails for future privacy research.

2. Conceptual Development

2.1. From Contextualization to Context Theorizing in Privacy Research

The importance of context in theory development was well recognized in information systems (Avgerou 2019) and beyond (Johns 2006). The information privacy literature, in particular, witnessed repeated calls for greater consideration of context over the last decade (Smith et al. 2011, Acquisti et al. 2015). Recently, the movement toward contextualization in privacy research was based on the notion of context serving as a *sensitizing device* that illuminates the potential boundaries to the paradigm within which privacy theories are nested. Through this lens, context is conceptualized as encompassing all situational opportunities and constraints that shape the meanings attached to privacy-related constructs and/or influence the functional relationships between these constructs. Nissenbaum (2009), for example, defined contextualization as linking privacy violations to a set of contextual factors such as the type of private information, the entities involved, and the transmission principles (e.g., buying or selling, consent, or coerced), which together make the conception of privacy more accurate and complete. Simply put, such "situational linking" allows researchers to more accurately assess the applicability of privacy theories whilst making their interpretation of empirical findings more robust.

There is little doubt that the contextualization of privacy theories is an area ripe for exploration, especially given the highly dynamic nature of the technological landscape pertaining to privacy. However, after many studies started exploring the sensitization of theory to a plethora of contexts, we began facing a vexing problem in the field: the variation of research findings from one study to another. For example, one study might find privacy concerns to be a strong predictor of privacy behavior in one context (Dienlin and Trepte 2015), another might find the two to be virtually uncorrelated in another context (Reynolds et al. 2011), and yet another could find the two to be negatively correlated in a third context (Sheehan and Hoy 1999). Treating context as a sensitizing device, we would simply accept this variation as an "error variance" and take it under advisement when setting the situational boundaries of our theory. This, unfortunately, falls well short of what firms and policy makers need in practice, which is a *context-contingent theory* that incorporates context as a *critical driver* of privacy concerns by explicating the mechanisms through which a context influences consumers' privacy concerns and their associated behavioral reactions.

When contexts are studied in a piecemeal fashion, the development of a broad-range context-contingent theory is often left to comparative studies or meta-analyses that synthesize the results of many studies (e.g., Jawahar and Williams 1997). For privacy research, this strategy faces two obstacles. First, it requires a sufficient number of context-specific studies to be coalesced over time. With the phenomena of information privacy being so amorphous, we are still far from forming a critical mass of context-specific studies to enable theoretical synthesis. Indicatively, a recent meta-analysis was only able to identify and test moderating variables at the country level (e.g., the level of privacy protection afforded by law) (Baruh et al. 2017). Second, unless the contextualization in each

study is theory grounded, it is unreasonable to expect a fast convergence toward a context-contingent theory (Kozlowski and Klein 2000). Unfortunately, there is not yet a framework that can guide the theorizing and assessment of context effects in privacy research, representing a considerable gap in the literature.

This article takes a first step toward bridging this gap. Its intended substantive contribution is to offer a conceptual and quantitative framework for assessing *how* contexts may influence people in forming and/or expressing their beliefs, attitudes, and perceptions about privacy, so as to enable a proper understanding and delineation of the context effects on privacy concerns. Such an understanding is not only theoretically salient—in illuminating the potential elements of a context-contingent theory—but practically pertinent to the *validity* (Cook and Campbell 1976) of future research for two reasons. First, it addresses a potential threat to internal validity in terms of conflating context effects with the underlying privacy concerns or the dynamics between privacy concerns and other constructs (e.g., privacy-related behavior). Such a conflation often occasions anomalous findings, especially when the underlying dynamics are countervailed by the context effects (Johns 2006). Second, it also allows researchers to maximize external validity by making more informed decisions in contextualizing their studies and demarcating the boundary conditions of their findings.

2.2. Context Effects on Privacy Concerns

Although the importance of context effects has long been recognized in behavioral research (Cronbach 1957), a notorious challenge is the lack of a consensual structure for studying context effects (Johns 2006). Some viewed context effects through the lens of validity threats (Rousseau and Fried 2001), and others examined how context conditions cognition, affect, and behavior (Johns 1991); yet others conceptualized context effects as moderating functional relationships between variables (Xie and Johns 1995), incurring sign reversals, changing causal directions, or tipping precarious relationships. In spite of the dissensus, the prevailing theoretical arguments (Johns 2006) point to two main context effects pertaining to an intrapsychic construct. One is *range restriction*, which limits the observations of the construct to only a portion of its full range. The other is *situational uncertainty*, which associates the construct with ambiguous meanings that are interpreted differently by different individuals. Both have their scholarly roots in the situational strength theory (Mischel 1977, Meyer et al. 2020), which contends that contexts vary considerably in their *situational strength* (i.e., the power of a context on abetting or constraining human agency). When the situational strength of a context on an individual is too "strong," it constrains the individual's expression (of the underlying construct) by making it more likely to conform to the norm, thereby imposing a *range restriction* on the observed expressions from multiple individuals. On the flip side, when the situational strength of a context on an individual is too "weak," it fosters a variety of meanings that the individual may discretionarily ascribe to the construct, thereby introducing considerable *situational uncertainty* to its observations. In the passages that follow, we discuss the manifestation of these two prominent context effects on people's stated privacy concerns.

2.2.1. Range Restriction. Mischel (1977) contended that a context is "strong" to an individual if it introduces norms that replace the individual's discretion as the most salient factor in forming his or her attitudes and behavior. Consequently, if a context exerts a high situational strength to multiple individuals, then the expressions of these individuals become considerably more homogeneous than would be expected based on each individual's trait profile, indicating a *range restriction* on the associated (observed) variables (Johns 2006).

Such strong contexts frequently emerge in privacy research, making the range restriction effect widely prevalent. For example, when a context involves the disclosure of Social Security numbers (SSNs), we are likely to observe a drastic increase in the base rate of self-reported privacy concerns for almost all individuals in the United States, effectively restricting the observed range by raising its floor. Similarly, when a context involves the disclosure of Facebook profiles for behavioral targeting in a political campaign, the negative connotation fueled by the Cambridge Analytica scandal would introduce a clear norm to people who are familiar with the scandal, elevating the situational strength of the context and consequently, restricting the range of self-reported privacy concerns (again by raising its floor).

As can be seen from the two examples, a *unique characteristic* of the range restriction effect in privacy research is that the restricted range tends to be at the higher end (i.e., reflecting heightened privacy concerns). Indicatively, Marreiros et al. (2017) demonstrated that the range of people's self-reported privacy concerns trends higher after exposure to a wide variety of contextual stimuli (e.g., after reading a newspaper article), no matter if such exposure is commonly perceived as positive, neutral, or negative for privacy. Such an upward trend was well recognized in the privacy literature, as researchers and practitioners frequently cited consumers' stated privacy concerns as "inflated" (Wittes and Liu 2015) and noted that consumers often express heightened privacy concerns yet refuse to take trivial actions to protect their privacy.

For example, Acquisti and Grossklags (2005) found among people who stated "high" concerns about the collection of their personal information such as name and address that 87.5% offered exactly such information in exchange for a free loyalty card.

In terms of the implications of range restriction on the validity of research findings, the methodology literature noted that range restrictions on the independent variable tend to deplete its predictive power (and correspondingly, its imputed importance) (Hunter and Schmidt 2004), potentially leading to null findings. The range restriction effect could also limit the generalizability of a study, as its finding could apply only to a portion of the full range of an independent (and/or dependent) variable. This limited generalizability naturally manifests as inconsistencies when comparing across studies, as the contexts specified in different studies could vary considerably in the extent to which they suffer from the range restriction. In sum, the prevalence of the range restriction effect suggests a contextual imperative in privacy research to detect and appreciate the range restrictions in eliciting privacy concerns (and other related constructs).

2.2.2. Situational Uncertainty. Context, as a frame of reference, has the potential to shape the very meaning of the focal construct (Johns 2006). When the situational strength of a context is too "weak" (Mischel 1977) for an individual, he or she might attach a variety of meanings to the situational stimuli, leading to a profound effect of *situational uncertainty* in how different individuals ascribe different meanings to the construct (e.g., privacy concerns)—uncertainties that transcend people's reasoning about the construct no matter if the form of reasoning is purposive, instrumental, or calculative (Solove 2002).

Weak contexts frequently emerge in privacy research. For example, consider a context that involves the disclosure of one's political views. Compared with SSN disclosure, this context is clearly much weaker for most people in the United States. Consequently, it permits considerable latitude for one's own understanding of "political views" (or the lack thereof) to imbue the meaning of "privacy concerns." In other words, the situational uncertainty entailed by the context of "political views" makes the underlying (privacy-related) construct more malleable than stable (e.g., as indicated by an individual's lack of an attitude in a coherent form).

The privacy literature has repeatedly noted this uncertainty with regard to privacy concerns, leading to the consensual belief that people are "likely to be uncertain about their own privacy preferences" (Acquisti et al. 2015, p. 510). Even though uncertainty is common for self-reported attitudes and beliefs (Bertrand and Mullainathan 2001), what is unique here is that the uncertainty does not appear limited to just a few inattentiveness respondents but prevalent in a large part of the population for a wide range of privacy issues, from their attitudes toward privacy (Acquisti et al. 2018) to the perceived effectiveness of privacy protection (Gates 2011) to even whether they consider themselves ignorant on issues pertaining to privacy (Acquisti et al. 2018).

In terms of implications on privacy research, situational uncertainty directly entails excessive variability in people's stated privacy concerns. Moreover, it could lead to people "casting around" for subtle cues (often unrelated to privacy) in expressing their privacy concerns (Acquisti 2009; Acquisti et al. 2015, 2018), with such cues ranging from the design and appearance of a survey instrument (John et al. 2011) to even the physical environment the individual happens to be in at the time (Acquisti et al. 2018). As different people may resort to different cues in accordance with their personal identities (Powell and Baker 2014), this cue-seeking tendency could give rise to even more randomness in their stated privacy concerns.

Statistically, such randomness could attenuate the observed correlation between different variables (Hunter and Schmidt 2004). Methodologically, Van Bavel et al. (2016) noted that the attenuation may be exacerbated by the fact that the situational uncertainty induced by the same context has markedly different effects on attitudes and the associated behavior. Substantively, Hoffmann et al. (2016) posited that uncertainty could lead to "privacy cynicism," with which people deliberately "discount risks or concerns" to cope with the excessive uncertainty. All these issues, no matter statistical, methodological, or substantive, could become the "tipping point" for an already precarious relationship between different privacy-related constructs (cf. Adjerid et al. 2018), leading to inconsistent findings in the literature and giving primacy to situational uncertainty as a key context effect to study in privacy research.

2.3. Explicating Context Effects with a Two-Dimensional Framework

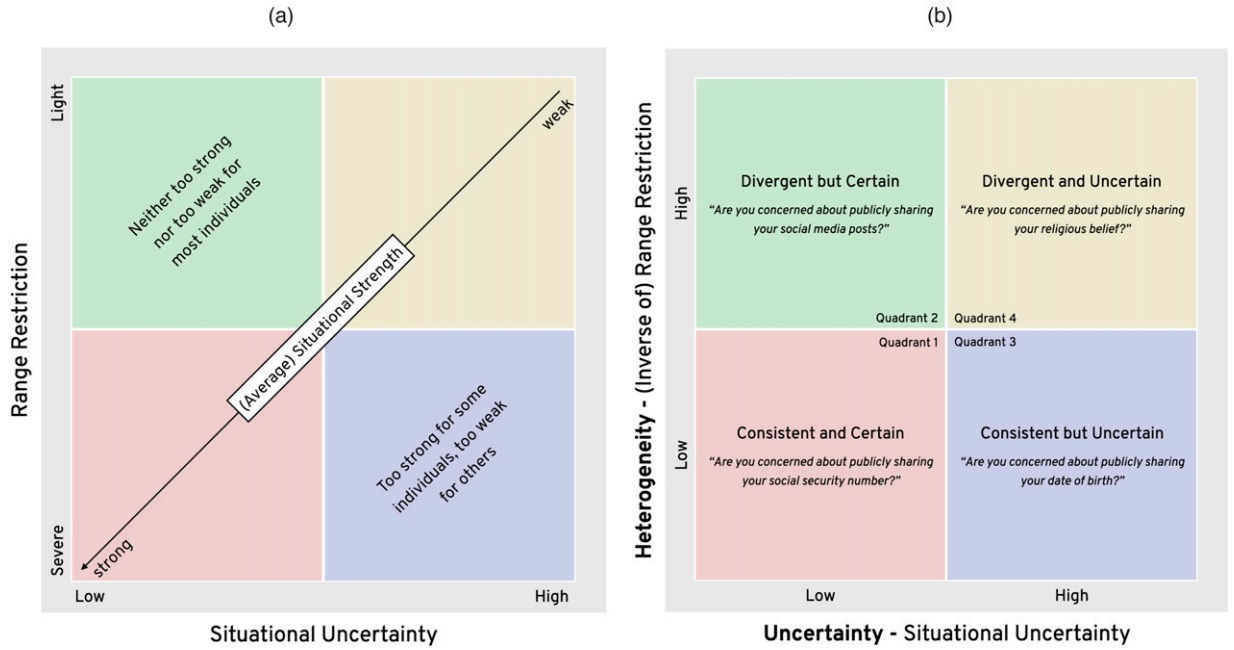
2.3.1. Why Two Dimensions? For the same individual, the two context effects are, by definition, two opposite ends of one spectrum; range restriction emerges when the situational strength of a context is too strong, whereas situational uncertainty arises when it is too weak. Yet, if we shift the unit of analysis from the individual level to the population level (i.e., by considering how a context affects the privacy concerns expressed by a large number of individuals), we must examine not only how strongly the context influences one individual but also how such influence varies across individuals in the observed population (e.g.,

how many individuals are *not* strongly affected by the context). This means that the two context effects, range restriction and situational uncertainty, could very well coexist at the population level. That is, because the situational strength of a context may vary considerably between individuals (as people are differentially sensitive to cues related to the situational strength) (Spector et al. 1995), the effect of the same context could manifest as situational uncertainty on some individuals, range restriction on some others, and neither on the rest. To fully capture the effects of a context at the population level, we need to treat (the aggregated magnitudes of) situational uncertainty and range restriction separately rather than as two ends of one spectrum. As illustrated in Figure 1(a), doing so is essential for differentiating between contexts with situational strength that is (1) consensually moderate (green), in which case neither range restriction nor situational uncertainty is significant, versus (2) too strong for some individuals yet too weak for others (purple), leading to a pronounced range restriction effect on some individuals *and* considerable situational uncertainties for others.

Reflecting the need of separate assessments for range restriction and situational uncertainty, we develop a two-dimensional framework to explicate how a context modifies the observed privacy concerns through the two effects. The goal is to enable the comparison of different contexts and to help researchers make more informed decisions when contextualizing an empirical study in future research. Traditionally, such explication is done descriptively and qualitatively, making the extent of the context effects difficult to quantify or to statistically compare in a meta-analysis. To address this issue, our framework is grounded in the *quantitative assessment* of such effects. That is, we aim to develop quantitative metrics for the effects of range restriction and situational uncertainty, so as to understand how the two context effects together influence and shape people's privacy concerns.

Note that the dimensionality of our framework (i.e., its use of two separate dimensions to capture the range restriction and situational uncertainty effects) does not mean that the two effects must be independent of each other. For example, improving the clarity of a context is known to simultaneously reduce situational uncertainty and strengthen the range restriction effect[2] (Meyer et al. 2010). Although this suggests a potential correlation between (the magnitudes of) the two effects, it by no means indicates that a context with a stronger range restriction effect *always* has a lower situational uncertainty. For example, a context associated with a conspiracy theory could exert strong range restriction effects on believers of the theory while inducing pronounced uncertainties among those who have never heard of it. Thus, the two

Figure 1. A Two-Dimensional Framework



Notes. (a) A quadrant view of the situational strength. (b) Conceptual illustration of our framework. (a) Both green and purple quadrants feature contexts with situational strength that, when averaged over all individuals in the population, falls in the middle of the spectrum. Yet, these two quadrants manifest qualitatively distinct context effects at the population level, as the situational strength of a context in the purple quadrant has a much higher heterogeneity across different individuals. (b) Our two-dimensional framework measures heterogeneity and uncertainty as the proxies for the range restriction and situational uncertainty effects, respectively.

dimensions cannot be collapsed into one when examining the context effects at the population level, indicating the necessity of a two-dimensional framework.

2.3.2. Assessing Context Effects Through Between-Individual Variability. Because contexts typically rest on a unit of analysis *above* those constructs being expressly studied, most existing work on context effects recognized the importance of quantifying such effects through the *distributional properties* of the observed variable, specifically its between-individual variability (Johns 1991, 2006) (e.g., how people's stated privacy concerns vary from one individual to another). The pertinence of this variability is obvious in our framework, as range restriction clearly reduces it while situational uncertainty increases it (Johns 2006). The challenge lies in the fact that the observed variability derives from a *mixture distribution* (McLachlan and Peel 2004), as different individuals may be subject to the context effects to different degrees (or to different context effects altogether) (Spector et al. 1995). In this case, to delineate different types of context effects, we must *statistically unpack* the observed variability into different components corresponding to the magnitude of the different effects.

From a methodological perspective, the unpacking is untenable under the traditional normality assumption. Instead, one must treat the observed variable as a mixture distribution and then use a process known as *mixture decomposition* (McLachlan and Peel 2004) to identify its components. Still, the decomposition is infeasible if all components follow the exact same distribution. Fortunately, as noted recently in psychology (Finch and Hernández Finch 2020) and sociology (Iannario et al. 2020), the components corresponding to the effects of range restriction and situational uncertainty likely follow distinct distributions. This enables us to leverage the statistical tools for mixture decomposition to estimate the extent of the two context effects for privacy research. Although we defer the mathematical details to the next section, in what follows, we provide a conceptual illustration of the decomposition idea and the resulting two-dimensional framework we developed for privacy research.

Consider a self-reported variable capturing an attitudinal or perceptual construct about privacy. When the variable is elicited with a given context (from a set of individuals), we conceptualize the between-individual variability of the construct as being composed of two sources. The first is the part of variability driven by the idiosyncrasy of people's attitudes toward privacy (i.e., the deviation of their attitudes from the norm for the context (Hollander 1958)), which is clearly constrained by the range restriction effect. We refer to this part of the variability as *heterogeneity*.[3] The second source of variability stems from an individual's lack of a coherent attitude toward privacy (again, in the given context). We refer to this part of the variability as *uncertainty*:

$$\text{Variability of Responses} = \underbrace{\text{Heterogeneity}}_{\text{variability driven by the divergence of attitudes}} + \underbrace{\text{Uncertainty}}_{\text{variability driven by the uncertainty of responses}} + \underbrace{\delta}_{\text{other sources of variability}} \tag{1}$$

Although the mathematical definitions of heterogeneity and uncertainty will be discussed in the next section, the correspondence is clear between their values and the magnitude of the context effects. Uncertainty, as a measure for the situational uncertainty effect, rises when the context induces situational uncertainty for a larger number of individuals. Heterogeneity, on the other hand, decreases when the context imposes a stronger range restriction effect on each individual (not subject to the situational uncertainty effect).[4] Further, Equation (1) illustrates the qualitative relationship between heterogeneity/uncertainty and the observed variability of people's privacy concerns. When the observed variable has a small variance, both heterogeneity and uncertainty must be low (red quadrant in Figure 1(a)) because a high level of either ensures a high overall variability. On the other hand, a larger observed variance could be a result of high heterogeneity (green quadrant), high uncertainty (purple quadrant), or both (yellow quadrant). In this case, discerning the two metrics is salient for explicating the effects of a context on privacy concerns.

Considering heterogeneity and uncertainty in tandem, the two-dimensional framework depicted in Figure 1(b) elucidates how a specific context may reground people's privacy concerns. This framework maps the four-quadrant division of situational strength in Figure 1(a) to the four high/low combinations of the heterogeneity and uncertainty metrics. Each context is corresponding to a point in the two-dimensional space, and contexts in different quadrants feature qualitatively distinct characteristics. In the following discussions, we use sample contexts in privacy research as idealized types to discuss the four quadrants in a stylized manner. That is, although the characteristics for each quadrant are expected to hold, the partitioning is not a typology, and there are bound to be exceptions (especially near the boundaries between quadrants).

2.3.3. Quadrant 1. The strong situational strength here yields privacy concerns that are "consistent and certain" (i.e., low heterogeneity, low uncertainty). As predicted by the situational strength theory (Meyer et al. 2020), contexts in this quadrant tend to be psychologically proximal to the focal construct, so people are more likely to think concretely about the situational stimuli. An example is the aforementioned context of disclosing SSN. This context is obviously

proximal to the notion of privacy given SSN's de facto status as a national identification in the United States. Because almost everyone will express severe privacy concerns in this context, a strong range restriction effect emerges, leading to a low variability of privacy concerns associated with disclosing SSN. As recognized in the methodology literature (Le and Schmidt 2006), this range restriction could bias the observed association between privacy concerns and other variables. Further, if the underlying association is a more complex (e.g., U-shaped curvilinear) relationship, then different contexts in quadrant 1 could give primacy to different "segments" of the relationship, some increasing and others decreasing, because of the different range restrictions exerted by these contexts (cf. Pierce and Aguinis 2013). This could threaten both the internal and external validity of the research findings.

2.3.4. Quadrant 2. With neither context effect being particularly strong, the contexts in this quadrant elicit "divergent but certain" privacy concerns (i.e., high heterogeneity, low uncertainty). An example is a context associated with the public sharing of social media posts. Because most social media users consciously and regularly adjust the visibility of their posts (Boyd and Hargittai 2010), there is likely strong familiarity and little ambiguity associated with the context, eliminating a major source of situational uncertainty (Meyer et al. 2010). Further, the regular adjustment of visibility settings, a decision clearly related to privacy, requires social media users to pay deliberate attention to the thematic deliberations involved in privacy-related decisions. Engagement in such thematic deliberations, in turn, makes their privacy attitudes manifest and illuminated, further depleting the situational uncertainties. Meanwhile, according to a Morning Consult poll[5] in 2018, about 45% of social media users kept all their accounts private, whereas 19% shared all their social media posts publicly. Although this poll probed people's behavior rather than attitudes, their attitudes almost certainly have to be more idiosyncratic than consensual to account for such notable variation in their behavior, consistent with an absence of the range restriction effect. In sum, because neither range restriction nor situational uncertainty are likely to dominate, contexts in this quadrant help bring the underlying dynamics between privacy concerns and other variables to the fore in research studies and increase the likelihood for findings in one context to be generalizable to another in the quadrant.

2.3.5. Quadrant 3. The situational strength of a context here could be simultaneously too strong for some individuals and too weak for others, entailing privacy concerns that are "consistent but uncertain." That is, although situational uncertainty may arise for a substantial part of the population, for the other individuals (i.e., those who are "certain"), the range restriction effect applies (leading to consistent responses within the "certain" subpopulation). For example, consider a question asking individuals in the United States for their privacy concerns about disclosing a combination of ZIP code, gender, and date of birth (DOB). For an individual who is aware of the landmark finding by Sweeney (2000) that the vast majority (87%) of Americans can be uniquely identified based on ZIP code, gender, and DOB, the context in the question introduces a norm of heightened privacy concerns, leading to a marked range restriction effect. In contrast, people who are unaware of the finding may consider the question rather ambiguous, with no clear expectation pertaining to a specific course of action. This lack of clarity, in turn, leads to a weak situational strength (Meyer et al. 2010), making these individuals bear the brunt of situational uncertainties. Consider the two subpopulations in tandem; the observed variability is likely dominated by *uncertainty* (i.e., the component driven by a large part of the population lacking a coherent attitude rather than *heterogeneity*), as little divergence of attitude exists among those who are certain of their responses. Thus, this quadrant is identified by low heterogeneity and high uncertainty.

There are pronounced differences between this quadrant and the previous two. First, the overall variability of responses is not as low as in quadrant 1, ameliorating the obstacle facing the detection of the underlying true score associations. Second, unlike quadrant 2, the observed variability in quadrant 3 is mostly driven by uncertainty rather than true heterogeneity. As a result, an association empirically identified between privacy concerns and another variable could be predicated on the covariation between the other variable and the individuals' uncertainty about their privacy concerns (rather than their true attitudes). This is not a validity threat in and of itself, as uncertainty has been given prominent theoretical consideration in the privacy literature (e.g., in the development of *privacy cynicism*) (Hoffmann et al. 2016). Nonetheless, it does require researchers to be mindful of the role of uncertainty in interpreting their research findings.

2.3.6. Quadrant 4. The weak situational strength here yields privacy concerns that are "divergent and uncertain" (i.e., high heterogeneity, high uncertainty). According to the situational strength theory, contexts in this quadrant are psychologically *distal* to the focal construct, forcing people to think abstractly and depleting the situational strength (Meyer et al. 2020). An example here is a context that involves social or political issues distal to privacy, like the sharing of people's

religious belief. Uncertainty is likely high in countries like the United States because many people do not have a clear religious belief (Lipka 2015). Meanwhile, among those who are religious, their attitudes toward sharing their beliefs or keeping them private could vary widely. According to a Pew Research Center survey in 2014, about 20% of U.S. adults shared their religious belief online during the preceding week, yet another 54% stated that they have never seen anyone sharing "something about their religious faith" online (Cooperman et al. 2014). With both heterogeneity and uncertainty running high—the former because of the absence of a range restriction effect and the latter because of the presence of situational uncertainties—the observed variation of privacy concerns could conflate heterogeneity and uncertainty. In the example, when two individuals reported different levels of privacy concerns, the difference could be driven by their varying attitudes toward privacy or by one having a religious belief and the other not having a religious belief. The existence of alternative explanations confronts researchers with the challenges of first discerning heterogeneity from uncertainty and then delineating the potentially complex associations between these two factors and the other variables involved in research. Although feasible using modern latent mixture model methods (McLachlan and Peel 2004), this challenge necessarily complicates the research design and limits the generalizability of findings to contexts in the other quadrants.

2.3.7. Summary. The conceptualization of the framework can support a qualitative appraisal of how a context affects the meanings ascribed to privacy concerns and how such meanings alter the tenor of interactions between privacy concerns and other variables. Yet, the value of the framework is better appreciated when the conceptual structure is instantiated with a quantitative assessment of heterogeneity and uncertainty, which we will focus on in the next section. For this reason, we defer to the end of the paper a comprehensive discussion of how to use the framework and the recommended practices it entails for future research. Before concluding the conceptual discussions, however, we stress that the framework not only indicates what is distinctive about each context but also identifies the similarity between contexts. As we will demonstrate through an example in the results section, by identifying which contexts tend to *cluster* with each other in the two-dimensional framework, we could stimulate the future accumulation and integration of research results across contexts in a cluster, which in turn, contributes to the development and pruning of a parsimonious context-contingent theory (Leavitt et al. 2010).

3. Mathematical Formulation

In this section, we instantiate our two-dimensional conceptual framework with a method that quantitatively assesses the *heterogeneity* and *uncertainty* of privacy concerns elicited in a specific context. As discussed earlier, both heterogeneity and uncertainty are *latent* variables that must be inferred from the observed variable (e.g., individuals' expressed privacy concerns). A well-established paradigm for such inferences is to develop a *latent variable model* (Shadish et al. 2002) that relates the observable variable with the latent ones. Depending on the nature (e.g., discrete or continuous) of the variables involved, a wide variety of latent variable models has been developed, including factor analysis (Child 2006), latent mixture models (McLachlan and Peel 2004), the Rasch model (Bond and Fox 2015), latent class models (Hagenaars and McCutcheon 2002), etc.

In our case, the observed variable is usually measured on an *ordinal* scale with multiple ordered options (e.g., for privacy concerns, from "very concerned" to "not at all concerned"). A latent mixture model specialized in handling an ordinal observed variable is the CUB model, which was initially developed by statisticians (Piccolo 2003) before being introduced to a variety of social science fields such as psychology (Finch and Hernández Finch 2020) and sociology (Iannario et al. 2020) for the analysis of self-reported ordinal data. In what follows, we first briefly review the CUB model before developing our latent variable model that directly links the observed variable with our target metrics (i.e., heterogeneity and uncertainty). At the end of the section, we discuss the computational methods for estimating the parameters of our latent mixture model and the metrics and indices for assessing its fit.

3.1. CUB Model

Consider an ordinal observed variable R (e.g., an individual's stated privacy concerns) with m possible responses[6] $1, 2, \ldots m$ (e.g., $m = 7$ for a seven-point Likert scale). According to the CUB model, the distribution of R (over a population sample we hereinafter refer to as *respondents*) is a function of two latent variables: $\xi \in [0,1]$, which captures the true *attitudes* of the respondents, and $\pi \in (0,1]$, which is inversely related to the *uncertainty* of the respondents toward their response. Specifically, for all possible responses $r \in [1,m]$, the CUB model specifies

$$P(R=r) = \pi \cdot \left[\binom{m-1}{r-1} \cdot \xi^{m-r} \cdot (1-\xi)^{r-1} \right] + (1-\pi) \cdot \frac{1}{m}. \tag{2}$$

As can be seen from Equation (2), the CUB model characterizes the probability distribution of R as a mixture of two distributions pertaining to the true

attitudes and the uncertainty of the respondents, respectively. The first attitude component is modeled as a (shifted) binomial distribution $B(m-1,1-\xi)$ with mean $(m-1)\cdot(1-\xi)+1$. That is, the larger ξ is, the "smaller" R is likely to be. The choice of a binomial distribution has been common in modeling ordinal data (Johnson and Albert 2006), mostly for its parsimony and also because any multinomial distribution can be factored into a sum of binomial distributions (Teicher 1954).

The second mixture component is a uniform distribution over all response categories, designed to capture the uncertainty of responses. That is, for all $r\in[1,m]$, there is $P_2(R=r)=1/m$. The choice of the uniform distribution not only achieves model parsimony but is supported by ample empirical evidence on probed answers (Manisera and Zuccolotto 2014). Note from Equation (2) that the two mixture components are weighted according to the latent variable π (i.e., $P(R=r)=\pi\cdot P_1(R=r)+(1-\pi)\cdot P_2(R=r)$). The larger π is, the more likely it is for R to be driven by true attitude rather than the randomness stemming from the uncertainty component.

3.2. Our Latent Variable Model

Although the CUB model defines the mixture composition of the ordinal observed variable R, in order to construct the latent variable model for our purpose, we still need to link R to the two latent variables of our interest, *heterogeneity* and *uncertainty*, which we henceforth denote as h and u, respectively. The link between R and u is straightforward. Because the CUB model captures the uncertainty of responses with the second mixture component P_2, we can directly map u to the weight of P_2 in the mixture distribution. That is, $u=1-\pi$. With this translation, the link between u and R naturally follows from Equation (2). To link heterogeneity h to the observed variable R, we start with a simple definition of h as the *remaining* standard deviation of R after taking into account the variance from the uncertainty component (i.e., after subtracting the second mixture component from the observed distribution). That is,

$$h=\sqrt{\sum_{r=1}^{m}\left[\frac{1}{1-u}\cdot\left(f_r-\frac{u}{m}\right)\cdot\left(r-\sum_{s=1}^{m}\left[\frac{s}{1-u}\cdot\left(f_s-\frac{u}{m}\right)\right]\right)\right]^2}, \tag{3}$$

where $f_r\in[0,1]$ is the observed frequency of R in the rth category. Although this definition of h fits the conceptual framework, it unfortunately cannot make h and u the model parameters because there are obviously many possible combinations of $f_1,\ldots,f_m$ that could result from the exact same pair of h and u. To address this identifiability issue, we consider an approximation of Equation (3) by defining the heterogeneity measure h according to the standard deviation of the first mixture component in the CUB model (i.e., $\sqrt{(m-1)\cdot\xi\cdot(1-\xi)}$ for the shifted binomial distribution $B(m-1,1-\xi)$). Specifically, after removing a constant factor of $\sqrt{m-1}$ (in order to standardize the range of h to $[0,\,1/2]$), we define the heterogeneity measure as $h=\sqrt{\xi\cdot(1-\xi)}$. Taking the definitions of u and h into Equation (2), we have the following definition of our latent variable model. Note that although we assume $\xi\geq 1/2$ (given the earlier discussions of the range restriction effect inflating the observed privacy concerns), the definition can be easily adapted to cases where $\xi<1/2$, especially because h remains constant when changing the value of ξ to $1-\xi$.

Definition 1. For given (latent) heterogeneity h and uncertainty u, our latent variable model specifies the probability distribution of the observed ordinal response variable $R\in[1,m]$ as

$$P(R=r)=(1-u)\cdot\left[\binom{m-1}{r-1}\cdot\frac{1}{2^{m-1}}\cdot\left(1+\sqrt{1-4h^2}\right)^{m-r}\cdot\left(1-\sqrt{1-4h^2}\right)^{r-1}\right]+\frac{u}{m}. \tag{4}$$

Online Appendix A, specifically Theorem 1, describes the relationship between the model parameters and the mean and variance of the observed variable, as stipulated by the latent variable model in the definition. The definition can also be extended to include covariates. The extension is discussed in Online Appendix B.

3.3. Estimating Model Parameters and Assessing Model Fit

Regardless of whether covariates are involved, the model parameters can be estimated using the expectation maximization (EM) algorithm (Dempster et al. 1977) for the log-likelihood function $\mathcal{L}=\sum_{i=1}^{n}\log(P(R_i=r_i))$, where r_i is the actual response reported by the ith respondent and $P(R_i=r_i)$ is the probability predicted by the latent variable model on the ith respondent reporting r_i. In the covariate-free case, $P(R_j=r)$ is constant for all $j\in[1,n]$, simplifying the log-likelihood function to $\mathcal{L}=\sum_{r=1}^{m}(n_r\cdot\log(P(R=r)))$, where n_r is the frequency of r in the responses and $P(R=r)$ is as defined in Equation (4). The only requirement for using EM is that the answer scale must have at least four categories (i.e., $m\geq 4$), in which case the model was proven identifiable (Iannario 2010).

The relative fit of our latent variable model can be assessed with general purpose fit indices such as the Akaike information criterion (AIC) (Vrieze 2012), the Bayesian information criterion (BIC) (Vrieze 2012), and the information complexity measure (ICOMP) (Bearse et al. 1997), all designed to balance fit with

model parsimony. In terms of the absolute model fit, we followed the F^2 statistic (Iannario 2012) developed specifically for an ordinal observed variable:

$$F^2 = 1 - \frac{1}{2} \cdot \sum_{r=1}^{m} |f_r - P(R = r)|, \tag{5}$$

where $f_r \in [0,1]$ is the observed frequency of responses in the rth category, whereas $P(R = r)$ is the predicted frequency according to the latent variable model. As can be seen from Equation (5), F^2 can be interpreted as the proportion of responses that can be accurately predicted by the model, with a value of 0.90 or higher deemed as reflecting good fit (Iannario 2012, Punzo et al. 2018).

4. Empirical Examination

In this section, we start with a discussion of the design imperatives of the empirical examination. That is, we outline the important characteristics of the two-dimensional framework that need to be empirically examined. Then, we describe the data sets used and the design of the empirical examination followed by the empirical results.

4.1. Imperatives of the Empirical Examination

The goal of the empirical examination was to study the aptness of our two-dimensional framework in capturing context effects in privacy research. To this end, three key properties arise as central issues to examine for assessing the effectiveness of the framework.

- *Prevalence of context effects*. First, we need to examine whether there is a significant degree of either effect (or both effects) in real-world contexts. After all, if both situational uncertainty and range restriction are a rarity in practice, our framework would be of little use to privacy research. Although our conceptual development elucidates the existing theories and evidence that point to the prevalence of both effects, it is important to empirically confirm such prevalence.
- *Completeness of the two-dimensional framework*. Second, we should study whether heterogeneity and uncertainty, in combination, account for most of the between-individual variability observed in practice. This verifies the completeness of our framework, ensuring that the two dimensions capture a comprehensive view of the context effects in privacy research.
- *Irreducibility into one dimension*. Finally, we need to ascertain that the two effects cannot be collapsed into a single dimension. That is, one cannot remove either heterogeneity or uncertainty in our model of response variability (i.e., Equation (1)), making it a single-dimensional model, without incurring a considerable loss of model accuracy.

Given the vast space of contextual contingencies (Johns 2006), our empirical examination is of necessity more illustrative than comprehensive. That is, as it is infeasible to exhaustively examine all possible contexts, our empirical examination was predicated on the notion of using a small number of contexts as idealized examples to appreciate the complexity of context effects in practice and to demonstrate the effectiveness of our two-dimensional framework—in particular, prevalence, completeness, and irreducibility, the three aforementioned properties.

4.2. Data Sets

To promote the practice of Open Science (National Academies of Sciences, Engineering, and Medicine 2018) and to demonstrate the value of our framework in guiding the selection of contextual contingencies in future research and practices, the data used in our empirical examination are a publicly accessible nationally representative survey of U.S. residents. The nationally representative samples featured in the data alleviate the possible confounding of context effects with the statistical distortion introduced by sampling biases.

4.2.1. Pew Research Center Privacy Panel. Commissioned by the Pew Research Center, this collection of four surveys was conducted by the GfK Group using KnowledgePanel, its nationally representative online research panel recruited through a combination of random digit dialing and address-based sampling methodologies (Madden et al. 2014). The final sample was weighted using an iterative technique that matches gender, age, education, race, Hispanic origin, household income, metropolitan area or not, and region to parameters from the March 2013 Census Bureau's Current Population Survey (Madden et al. 2014). This privacy survey was administered in four waves in the beginning of 2014, the middle of 2014, the end of 2014, and the beginning of 2015. The initial wave included a nationally representative sample of 607 adults (i.e., 18 years old or older) in the United States, of whom 417 (69%) participated in all four waves of the panel. To offset panel attrition, new participants were recruited in latter waves. Although all results reported in the paper were based on the 417 participants who participated in all four waves, we also repeated all analyses when including the latter-round recruits and found no noticeable differences in the results. The survey items used in this article are summarized in Online Appendix D.

Two groups of questions in the Pew surveys are particularly relevant to our empirical examination: (1) demographic variables and (2) perceptions and beliefs about privacy. In terms of demographics, the respondents indicated their gender, age, household income, and highest education completed.[7] In terms of privacy-related questions, the respondents were asked about their perceived sensitivity for various types of personal information (wave 1), their perceived

concerns on different types of companies keeping their information private (wave 2), their expectation of how long each company should store their information (wave 2), and their level of concern on the government monitoring various types of their private information (wave 3). The questionnaire design is obviously different from academic surveys. Instead of featuring a rigorous theoretical grouping of questions, the Pew surveys clustered questions by topics of interest and probed a wide variety of conceptually related yet qualitatively distinct constructs central to people's privacy concerns (e.g., perceived sensitivity, privacy expectations, etc.). Further, instead of contextualizing the questions with one or a few closely related contexts—as in most academic surveys (e.g., Xu et al. 2012)—the Pew surveys implicated each construct being probed in a plethora of contexts, from credit card companies keeping track of purchase history to government agencies monitoring search engine activities, and reported the different responses.

The design of the Pew surveys fits well with our purpose of assessing the effectiveness of the two-dimensional framework. First, the assortment of real-world contexts studied in the surveys enabled us to examine the prevalence of range restriction and situational uncertainty effects in practice. The variety of constructs probed, on the other hand, testified to the versatility of our framework and allowed questions like whether a set of contexts exerting similar effects on one construct also exerts similar effects on a related construct. Last but not least, the high profile of the Pew privacy surveys—which are known to shape public discourses, policies, and practices—demonstrates the practical value of our framework beyond academic research.

4.2.2. Gallup Internet Privacy Survey 2011/2018. This collection of two surveys was conducted by Gallup in January 2011 and April 2018, both through telephone interviews on landline and cellular phones selected by random digit dialing. To form a nationally representative sample, in both cases, the final sample was weighted by gender, age, race, Hispanic ethnicity, education, etc. and corrected for sampling biases such as the double coverage of landline and cell phone users. Note that, although the sampling processes were similar,[8] the samples in the two surveys were independently selected. Because the surveys focused on the privacy attitudes of Facebook and Google users, the questions about Facebook or Google were asked of respondents who self-reported to use the two services, respectively. The 2011 survey included 1487 adults, of whom 559 (38%) were Facebook users and 904 (61%) were Google users. The 2018 survey included 1,509 adults, of whom 785 (52%) were Facebook users and 1,106 (73%) were Google users.

Unlike the Pew surveys, the Gallup surveys provided a more holistic view of people's privacy concerns, as the surveys directly probed the respondents' level of concern on "invasion of privacy" in the context of Facebook and Google (the detailed survey items are summarized in Online Appendix D). Because the public opinions about both services shifted considerably between the two surveys (i.e., 2011–2018), the Gallup surveys provided a unique opportunity for examining how their context effects on privacy concerns changed accordingly in this period.

4.3. Design of the Empirical Examination

Recall from earlier discussions that the empirical examination was designed to focus on three key properties (i.e., *prevalence* (of context effects), *completeness* (of the two-dimensional framework), and *irreducibility* (into one dimension)). Juxtaposing different contexts with our two-dimensional framework enables the assessment of all three properties. Specifically, the estimated magnitudes (and their statistical significance) of the context effects would signal their *prevalence* across contexts. The goodness of fit between (the latent variable model underlying) our framework and the observed data would verify the *completeness*. Finally, by comparing the goodness of fit of our framework with that of the baseline single-dimensional models, the superiority of our framework would demonstrate its *irreducibility* into a single dimension.

Although juxtaposing any set of contexts would allow the assessment of the three properties, we started by grouping contexts according to the two underlying contextual factors that were given prominent roles in the existing studies of privacy attitudes and beliefs (Xu et al. 2012): (1) different types of information collected from individuals (e.g., purchase records, location, and emails) and (2) different types of organizations (e.g., specific vendors, online marketers, and social networking sites) collecting and/or using such information. Then, we continued with three additional analyses. First, we considered contexts that involve the collection of five types of information (from phone records to search engine logs) by the government (i.e., the type of organization). Second, we examined the idiosyncrasy of individual-level context effects. Given the importance of demographic variables in shaping people's privacy concerns (Sheehan 1999), we included four demographic variables (gender, age, household income, and education) as covariates in our latent variable model to examine whether there are substantial regularities in the context effects that transcend individual-level differences. Finally, we applied our framework over the Gallup data to assess the change of context effects over time. In all these studies, our goal was to examine the aforementioned *prevalence*, *completeness*, and *irreducibility* of our

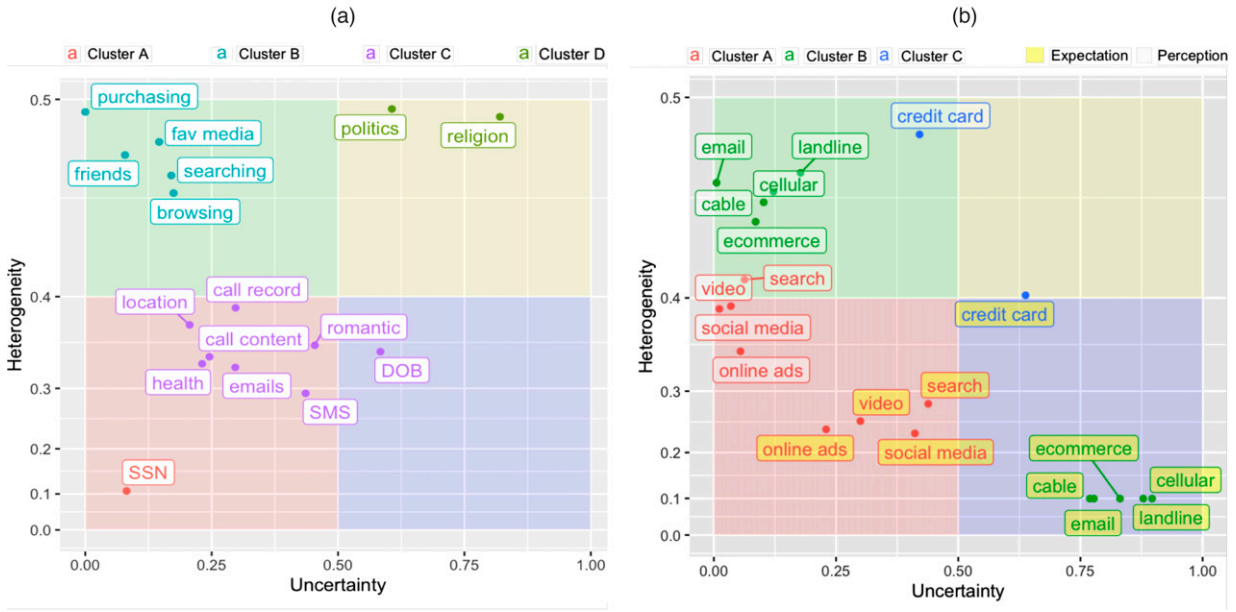
framework for assessing the context effects and *not* to delineate the specific mechanism through which each contextual factor alters the overall effect of a context. The latter requires theorizing beyond the scope of the current work, as we further elaborate in the discussion of future research at the end of the paper.

4.4. Results

4.4.1. Contextual Factor 1 (Type of Information). We first analyzed the responses to 16 survey items administered in the first wave of the Pew survey, which asked the respondents to rate their perceived sensitivity for 16 types of information on a four-point scale (from "very sensitive" to "not at all sensitive"). Figure 2(a) depicts the heterogeneity and uncertainty of the responses to each survey item, with the detailed parameter estimates for the model summarized in Table 1 in Online Appendix C. Although we refer readers to Online Appendix C for the detailed model fit (and its comparison with alternative models), we summarize here the three main findings corresponding to the three aforementioned properties (i.e., prevalence, completeness, and irreducibility).

First, our model provides ample support for the *prevalence* of context effects in the responses. In terms of the situational uncertainty effect, the parameter estimates for uncertainty were statistically significant ($p < 0.05$) for all but one type of information (i.e., purchasing records in the top left corner of Figure 2(a)). Further, the wide dispersion of estimated uncertainties (from 0.00 for purchasing records to 0.82 for religion) indicates that the magnitude of the situational uncertainty effect likely varies considerably between contexts. Although the statistical significance of the range restriction effect cannot be directly tested,[9] the estimated heterogeneity for some contexts was clearly much smaller (i.e., more restricted) than others (e.g., 0.11 for SSN compared with 0.47 for browsing records). This wide dispersion of estimated heterogeneity indicated the existence of range restriction for many contexts. Second, in terms of *completeness*, our latent variable model fits well with the data. The F^2 statistic for the absolute model fit exceeds the recommended threshold of 0.90 (Iannario 2012) for all 16 items, indicating that the model-predicted response distributions were very close to the observed distributions. Third, according to the aforementioned fit indices, our two-dimensional framework outperformed the single-dimensional baseline model in all 16 contexts even after taking into account model parsimony, speaking to the final property of *irreducibility*.

Figure 2. Comparison of Context Effects



Notes. (a) Contextual factor 1: Type of information. (b) Contextual factor 2: Type of organization. (a) Browsing = websites visited. Call content = content of phone conversations. Call record = numbers called/texted. Emails = content of email messages. Health = state of health and medicine regularly taken. Fav media = favorite media (e.g., music, movies, books). Friends = identities of friends. Location = physical location gathered from Global Positioning System (GPS) in mobile devices. Politics = political views and candidates supported. Religion = religious and spiritual views. Romantic = history of romantic relationships. Searching = searches made using online search engines. SMS = content of text messages. (b) Cable = cable television companies. Cellular = cellular telephone companies. Credit card = credit card companies. Email = email providers. Landline = landline telephone companies. Online ads = online advertisers. Search = search engine providers. Social media = social media companies. Video = online video websites. In both panels, the estimated heterogeneity/uncertainty and the fit statistics are included in Online Appendix C. Heterogeneity was plotted in the adjusted logarithmic scale, as discussed in Online Appendix C.

Having confirmed the three properties, we note that Figure 2(a) also highlights the wide dispersion of context effects across the 16 types of information, which form four obvious clusters in the two-dimensional space. Among all 16 types, SSN is the lowest on heterogeneity, forming its own cluster (cluster A) in quadrant 1. As discussed earlier, the situational strength of the SSN context depletes the variance of people's expressed privacy concerns, making 90.71% of respondents choosing the (highest) "very sensitive" category and reducing both heterogeneity and uncertainty to near zero. Forming a sharp contrast to SSN are the five types of information in cluster B, which all reside in quadrant 2. Compared with the SSN context, the situational strengths of these contexts are not as strong, as these types of information are not associated with any obvious privacy-related norm. Meanwhile, the situational uncertainty induced by these contexts remains low, as people tend to be intimately familiar with the types of information in this cluster and the implications of disclosing them. Overall, the contexts in cluster B feature a combination of high heterogeneity and low uncertainty.

Compared with cluster B, the contexts in cluster C have smaller heterogeneity but larger uncertainty. Consider the context of disclosing DOB as an example. On one hand, it introduces an obvious norm to people who are aware of the association between SSN and DOB (Acquisti and Gross 2009) or to those who use their DOB as a security question for online authentication (Pinchot and Paullet 2012). On the other hand, to those who are unaware of these associations, DOB would appear rather distal to privacy, as they are unlikely to have devoted any previous thought on DOB being sensitive. Because people are known to be uncomfortable with acknowledging ignorance or even the appearance of being ignorant, a "distal" context such as DOB likely manifests as a strong situational uncertainty effect. When a context exerts a significant degree of both effects (albeit on two different parts of the population), an assessment of the context effects at the population level tends to exhibit (1) a considerable uncertainty, reflecting the situational uncertainty effect on one part of the population, and (2) a low heterogeneity—reflecting the range restriction effect on the other part of the population—given that heterogeneity is inversely related to the magnitude of the range restriction effect. This is what we observe in the figure for cluster C. Finally, politics and religion form their own cluster D in quadrant 4, where heterogeneity and uncertainty both run high. This is again consistent with our earlier discussions, as both items fit into the type of "controversial issues" that tend to incur considerable situational uncertainty.

4.4.2. Contextual Factor 2 (Type of Organization).

Next, we analyzed the responses to survey items in the second wave, which asked two different questions for (each of) 10 types of companies ranging from online service providers like social media to traditional financial companies like credit card issuers. The first, which we refer to as the *perception* items, asked about the respondents' perceived confidence in personal information being kept private by the companies (four-point scale, from "not at all confident" to "very confident"). The other, which we refer to as the *expectation* items, asked the respondents about privacy expectations, specifically in terms of how long a company *should* retain their personal information (five-point scale, "should not save any information," "a few weeks," "a few months," "a few years," "as long as they need to").

Figure 2(b) depicts the heterogeneity and uncertainty imputed from the responses, with the detailed parameter estimates summarized in Table 2 in Online Appendix C. Like in the first study, the results supported *prevalence*, *completeness*, and *irreducibility*. In terms of prevalence, uncertainty was statistically significant ($p < 0.05$) for 18 of 20 contexts, whereas heterogeneity varied widely across contexts (between 0.10 and 0.49). In terms of completeness, our F^2 exceeded 0.80 for all 20 items and 0.90 for 15 of them, demonstrating a good fit. Further, the fit indices suggest that our model outperformed the baseline model on all 20 items, indicating irreducibility.

Other than affirming the three properties, the contrast between perception and expectation items allowed us to examine two additional questions: (1) whether the types of companies that manifest similar context effects for one item also have similar effects for the other and (2) whether the context effects manifested by the same type of company differ between the perception and expectation items. The answers to the two questions are obvious from the figure. First, note from Figure 2(b) that the context effects exerted by the 10 types of companies form three clearly separated clusters. Interestingly, the companies that cluster together for the perception item are also clustered together for the expectation one. This is an encouraging observation for the generalizability of privacy research across contexts, as it demonstrates that the crosscontext generalizability (e.g., between two companies in the same cluster) for one privacy-related construct likely extends to other privacy-related constructs. Second, equally obvious from the figure is that the context effects exerted by the same company could vary drastically from one construct to another. Specifically, for every 1 of the 10 types of companies, the heterogeneity of responses is considerably lower—and the uncertainty considerably higher—for the expectation item than the perception item. This is remarkably

consistent with the arguments by Hong and Thong (2013) that survey questions focusing on people's expectations may not be able to elicit their idiosyncratic attitudes or beliefs toward privacy because of the absence of a real trade-off in the question design.

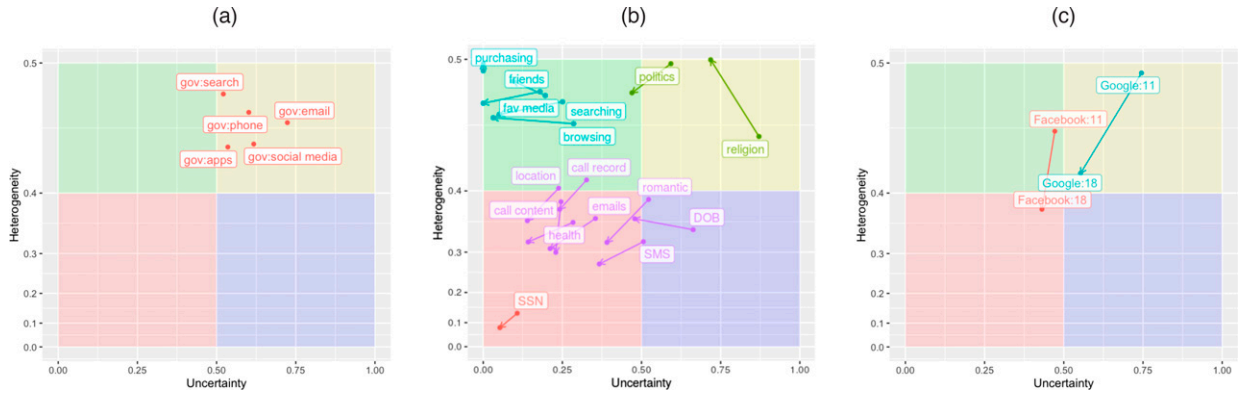
4.4.3. Other Empirical Evaluations. Figure 3 depicts the results of the three remaining analyses, with their detailed parameter estimates and fit statistics summarized in Tables 3–5 in Online Appendix C, respectively. Before discussing each study in detail, we note that the results, once again, supported the three key properties of our framework (i.e., prevalence, completeness, and irreducibility). As can be seen from Tables 3 and 5 in Online Appendix C, for all items, the estimates for uncertainty were statistically significant (*prevalence*), the fit statistic F^2 was above 0.90 (*completeness*), and our model outperformed the single-dimensional baseline model on all fit indices (*irreducibility*).

4.4.3.1. Government Surveillance. We analyzed the responses to five survey items administered in the third wave of the Pew survey, which asked the respondents to rate, on a four-point scale (from "very concerned" to "not at all concerned"), their level of concern on government collecting five types of information. As can be seen from Figure 3(a), all five contexts fell in quadrant 4, indicating pronounced situational uncertainty. This echoes our earlier conceptual development suggesting that situational uncertainty arises from contexts involving "controversial issues" (e.g., the role of the government in balancing national security with individual privacy). The social discourse regarding such issues and the divided opinions (Dinev et al. 2008) suggest the lack of an established norm (and a dearth of range restriction), driving the contexts into quadrant 4.

4.4.3.2. Demographics as Covariates. To understand how context effects covary with demographic variables, we introduced gender, age, income, and education as covariates in the model. Table 4 in Online Appendix C summarizes the coefficient estimates linking heterogeneity and uncertainty to the demographic covariate for each item. Overall, the demographics of respondents do *not* appear to substantially improve the fit of the model. Indicatively, two of three fit indices, BIC and ICOMP, always suggest the covariate-free model as a better fit.[10] This fit deficit is indeed good news for future studies of context effects, as it suggests that such effects are unlikely to vary radically from one population sample to another. Figure 3(b) also testifies to this generalizability. Although income level does covary with the heterogeneity and uncertainty of people's responses in all contexts, it does *not* change the insights we discussed earlier for the contexts, as the composition of the four clusters remains stable regardless of the income level.

Notwithstanding the overall observation, Table 4 in Online Appendix C highlights a few demographic variables that may shift the context effects. For example, more education generally reduces uncertainty ($p < 0.05$ for the perceived sensitivity of DOB, SSN, and romantic relationships). Age tends to increase it ($p < 0.001$ for the perceived sensitivity of DOB and the expected time frame for search engines to store private information). In terms of heterogeneity, women

Figure 3. Other Empirical Evaluations: Government Surveillance, Demographics, and Temporal Changes



Notes. (a) Government surveillance. (b) Demographics as covariates. (c) Temporal changes. (a) All items asked about respondents' level of concern on government monitoring of certain activities. Social media = activity on social media websites. Search = on search engines. Phone = on cell phone. Apps = on mobile apps. Email = email messages. (b) See the note in Figure 2 for the meaning of item names. Each arrow is pointing from the parameter estimates for respondents with household income less than or equal to the median (i.e., category 13, $60,000 to $74,999) to those with income greater than or equal to the median income. The label (i.e., the variable name) is positioned close to the origin of each arrow. (c) All items asked about respondents' level of concern of invasion of privacy when using the corresponding service. 11 = survey conducted in 2011. 18 = survey conducted in 2018.

tend to have less divergent attitudes than men ($p < 0.001$ for the perceived sensitivity of text messages (SMS), call content, and call record), with the only exception on political views, for which the heterogeneity is higher for women ($p < 0.01$). Similarly, more education usually reduces heterogeneity, with one notable exception of religion, for which highly educated people tend to have more divergent attitudes ($p < 0.01$). Considering this in tandem with the well-known correlation between income and education, we have a potential explanation for the "outlier" direction of the arrow for religion in Figure 3(b). Overall, the shift of context effects with these demographic variables suggests additional subtleties to be considered when examining the association between demographics and people's privacy-related behavior in a specific context (cf. Goldfarb and Tucker 2012).

4.4.3.3. Temporal Change of Context Effects. Finally, we analyzed the responses to the Gallup surveys in 2011 and 2018, both of which probed the respondents on their level of concern over the invasion of privacy when using Facebook and Google. As can be seen from Figure 3(c), for both Facebook and Google, the context effect shifted between 2011 and 2018, in particular with a reduction of both heterogeneity and uncertainty, indicating an increase of the situational strength. This is consistent with our earlier conceptual development, as the growing usage of these web services[11] and the constant social discourse about the privacy implications of such use (e.g., Rader 2014) both contributed to the institution of societal norms that increase their situational strengths.

5. Discussions

5.1. Research Implications

Contextualization is increasingly important in privacy research thanks (at least in part) to the rapidly diversifying landscape of technological issues from which privacy concerns arise. The need to contextualize will only be reinforced in the future by the ever-growing diversity of technologies and perspectives. This makes it all the more important for researchers to be mindful of the theoretical and methodological implications of the contextual contingencies they chose. For each study, a theory-driven contextualization is integral to the proper interpretation and generalizability of the findings. For the longer-term goal of developing a context-contingent theory, theory-driven contextualization is even more critical, as it accelerates the convergence of context-specific research into an integrative understanding of the effects of context contingencies.

With this backdrop, the research implications of our findings are threefold. First, our findings demonstrated how context can substantially shift the distributional properties of people's perceptions, attitudes, and beliefs about privacy. Some contexts, like those involving SSN or the collection of personal data by online advertisers, could institute an obvious norm that severely constrains people's idiosyncratic differences on a privacy-related construct. Other contexts, like those involving the disclosure of political views, could permit so much latitude in an individual's interpretation of privacy that a privacy-related construct could be conflated with factors such as whether the individual has a clear political view. The wide variability of context effects in our findings substantiated the need for privacy researchers to assess such effects, both conceptually and quantitatively, when contextualizing future research studies.

Second, our findings represented the first steps toward a systematic understanding of context effects in privacy research. Specifically, drawing from the situational strength theory, we identified range restriction and situational uncertainty as the two main effects a context may exert on people's perceptions, attitudes, and beliefs about privacy. Further, by introducing to the privacy literature an emerging method for survey data analysis in psychology (Finch and Hernández Finch 2020) and sociology (Iannario et al. 2020), we were able to quantitatively assess the degree of the two effects through two metrics, heterogeneity and uncertainty, which together form our two-dimensional quantitative framework for assessing the effects of a context on privacy-related constructs. Through an empirical examination of the results of two nationally representative surveys, we showed that the two-dimensional framework fit well with the data, as the two dimensions together accounted for the vast majority of the observed variability between individuals. Further, both dimensions are essential, as the two-dimensional framework was far superior to a wide spectrum of baseline single-dimensional models in terms of model fit.

Finally, with an easy to use visualization, our two-dimensional framework also serves as a diagnostic device for comparing and contrasting the effects of different contexts, so as to help understand how the meanings people ascribe to an underlying privacy-related construct shift from one context to another. Such a diagnostic device not only provides scholarly guidance for the selection of contextual contingencies in future privacy research but also, allows a proper understanding of whether research findings in one context are likely generalizable to another context. It is our hope that our framework would motivate more privacy researchers to make the choice of context a matter of research design integrative to the framing of a study, so as to advance the pace of convergence

toward context-contingent theories in information privacy.

5.2. Recommended Practices for Contextualizing Future Privacy Research

Our two-dimensional framework provides a conceptual and quantitative account for how a context regrounds privacy concerns in terms of (1) the range restriction effect, as measured by *heterogeneity*, and (2) the situational uncertainty effect, as measured by *uncertainty*. Together, they underpin the following practices we recommend for contextualizing future privacy research:

First, we recommend researchers consider the use of heterogeneity and uncertainty measures to identify potential caveats stemming from the context effects. For example, if both measures firmly place the context effects in the lower left corner of the first quadrant, a researcher might want to further scrutinize whether enough variability could emerge from the range-restricted privacy concerns to support the detection of any underlying true score associations.

Second, we recommend researchers demarcate the boundary conditions of their findings based on the context effects quantified by our framework. As discussed earlier, whether to situate privacy concerns in the context of social media or e-commerce could considerably alter the meanings a respondent ascribes to privacy. Similarly, contexts may exert varying degrees of influences on the range of attitudes and behaviors captured in a study. As a result, although a finding from a social media context might readily generalize to the context of online ads, it might not apply as well to e-commerce websites (see Figure 2(b)). Our framework could help researchers demarcate the contextual boundaries based on the quantified effects of the study context and the target one.

Finally, we echo the call for the incorporation of context effects into theoretical models when feasible (Johns 2006). This means not to gloss over the differences between contexts but to explicitly examine the potential roles the context might play in interpreting the focal phenomena and/or moderating the focal relationship. For observational studies in particular, we caution against the practice of simply "controlling away" contextual factors. The remarkable strengths of the context effects are evident from the results in the paper. Such strengths, in turn, challenge the notion that the association between substantive variables holds roughly constant across contexts, a tacit assumption made in including contexts as control variables.

5.3. Limitations and Future Research

One limitation of our work relates to the data sets we used. Although rich in contextualization, they lack in the complexity of dispositional factors when compared with the context-free multidimensional operationalizations of privacy concerns (e.g., Smith et al. 1996, Malhotra et al. 2004). Such emphasis on contextual contingencies fits our purpose, as it arguably simplifies the differential unfolding of the effects that context variations occasion. However, it cannot fully explicate the nuances of people's privacy concerns. Future work can apply our framework on a data set that incorporates context into the multidimensional operationalization of privacy concerns, so as to glean new insights on how contexts may affect different dimensions of people's privacy concerns.

Although we designed the two-dimensional framework to assess the effects of a context on one privacy-related construct, future research could use the framework to study how the effects of a context vary *across* different privacy-related constructs. For example, our empirical examination demonstrated that the effects exerted by the same (or substantively isomorphic) context varied substantially when the focal construct was perceived confidence versus privacy expectation. This introduces an intriguing question to research studies that examine the relationships between different privacy-related constructs. If a context affects different measured variables in different ways, how should we capture the influence of context effects on a relationship that involves multiple such variables? To answer this question, future research could examine the various ways in which context effects may modify empirical findings about a relationship (e.g., reducing the statistical power, inducing false positives) (Le and Schmidt 2006) when applying to an independent variable, the dependent variable, or one of the moderating or mediating variables.[12]

Another important topic for future research is how the various characteristics of a context (e.g., its clarity, as discussed earlier in the paper) influence and shape the effects of the context. To this end, researchers may draw from the various contextual factors examined in the situational strength theory (Meyer et al. 2010), vary each contextual factor individually, and examine the change of heterogeneity and uncertainty associated with the resulting context. Understanding the link between contextual factors and the overall context effects is not only helpful for the proper synthetization of studies (with different contexts) in a meta-analysis but also, instrumental for the design and selection of contexts in future research studies.

6. Conclusion

The notion of context has long been a subject of substantive debates both in the field of information systems (e.g., whether context is ephemeral and transient or stable and enduring) (Avgerou 2019) and in the

literature of information privacy (Nissenbaum 2004). Unfortunately, although researchers have grappled with the question of what context is, still nascent is a systematic understanding of how contexts alter the grounds on which people form their privacy concerns. In this paper, we take the first step toward addressing this problem by developing a conceptual and quantitative framework for assessing how different contextual factors influence people's stated privacy concerns. The resulting two-dimensional framework enables the quantitative appraisal of context effects and provides researchers with a diagnostic device guiding the selection of contextual contingencies in future privacy research. It is our hope that the findings of this paper inspire more studies on context effects in privacy research, which together advance the pace of our transition from using contextualization as a sensitizing device to developing and testing context-contingent theories for information privacy.

Acknowledgments

The authors are grateful for helpful feedback and comments from the department editor, the associate editor, and three anonymous reviewers. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the funding agencies.

Endnotes

[1] Note that the full name of the model has also been referred to as "covariates in a uniform and shifted binomial mixture" (Finch and Hernández Finch 2020). The substance of the model remains unchanged regardless of the name variation.

[2] The reason why removal ambiguity also strengthens the range restriction effect is because the clarified context (now) clearly defines the specific behaviors expected from individuals, further restricting the expressions of those who are familiar with the norm.

[3] Note that our use of the term "heterogeneity" derives from its meaning in econometrics (Arellano 2003) and is distinct from statistical heterogeneity as measured by Shannon's entropy, Gini index, etc.

[4] It is important to note that heterogeneity does not indicate the absolute magnitude of the range restriction effect. Measuring the absolute magnitude would require comparing the heterogeneity of a variable elicited with a context against that of a "vanilla" variable without any context effect. The latter, unfortunately, is infeasible to capture as any attempt in doing so would give rise to important conceptual challenges, such as what privacy concerns mean without any context (Nissenbaum 2009). For this reason, we adopt a different set of names for the metrics (i.e., heterogeneity and uncertainty) rather than reusing those for the context effects.

[5] See https://www.statista.com/statistics/934874/.

[6] We do not consider nonresponses (e.g., refusal to answer a question, "Don't Know" responses if such an option is available) for both theoretical and practical reasons. Theoretically, survey researchers have long debated whether allowing nonresponses improves or degrades response quality, with recent arguments favoring the latter given findings that nonresponses are more often from respondents who lack motivation rather than those who are truly uncertain (e.g., Krosnick et al. 2002). Practically, the number of nonresponses is very small in our main data set, rendering the point moot. For variants of the CUB model that do consider nonresponses, see Manisera and Zuccolotto (2014).

[7] The household income was categorized into 19 levels according to the standard categorization used by the Federal Reserve in the annual Survey of Household Economics and Decision making (SHED). Education was categorized into four levels, again according to the definition in SHED.

[8] One exception was the weighting process used to obtain a nationally representative sample. The demographic weighting target was the March 2010 version of the U.S. Current Population Survey (published by the U.S. Census Bureau and the U.S. Bureau of Labor Statistics) for the 2011 data and the March 2017 version for the 2018 data.

[9] As discussed earlier, this is because of the infeasibility of measuring the full range in a context-free manner.

[10] The covariate-free model is the one with two parameters, heterogeneity and uncertainty. Although AIC does suggest the model with demographic covariates to be better fitting for some survey items, AIC is known to emphasize less on model parsimony, and thus, it is more likely to prefer more complex models like the one with covariates (Vrieze 2012).

[11] For example, as discussed earlier, the percentages of respondents who use Facebook and Google grew from 38% and 61% in 2011 to 52% and 73% in 2018, respectively.

[12] Note that many privacy-related constructs have been theorized to serve different roles in a relationship. For example, *trust* has been theorized as an antecedent to privacy concerns (Belanger et al. 2002), an outcome of privacy concerns (Malhotra et al. 2004), a moderator for the influence of privacy concerns on behavior (Bansal and Zahedi 2008), and a mediator between privacy concerns and behavior (Xu et al. 2005).

References

Acquisti A (2009) Nudging privacy: The behavioral economics of personal information. *IEEE Security Privacy* 7(6):82–85.

Acquisti A, Gross R (2009) Predicting Social Security numbers from public data. *Proc. Natl. Acad. Sci. USA* 106(27):10975–10980.

Acquisti A, Grossklags J (2005) Privacy and rationality in individual decision making. *IEEE Security Privacy* 3(1):26–33.

Acquisti A, Brandimarte L, Loewenstein G (2015) Privacy and human behavior in the age of information. *Science* 347(6221): 509–514.

Acquisti A, Brandimarte L, Loewenstein G (2018) Privacy and human behavior in the information age. Selinger E, Polonetsky J, Tene O, eds. *The Cambridge Handbook of Consumer Privacy* (Cambridge University Press, Cambridge, United Kingdom), 184–197.

Acquisti A, Brandimarte L, Loewenstein G (2020) Secrets and likes: The drive for privacy and the difficulty of achieving it in the digital age. *J. Consumer. Psych.* 30(4):736–758.

Acquisti A, John LK, Loewenstein G (2013) What is privacy worth? *J. Legal Stud.* 42(2):249–274.

Acquisti A, Taylor C, Wagman L (2016) The economics of privacy. *J. Econom. Literature* 54(2):442–492.

Adjerid I, Pe'er E, Acquisti A (2018) Beyond the privacy paradox: Objective vs. relative risk in privacy decision making. *Management Inform. Systems Quart.* 42(2):465–488.

Arellano M (2003) *Panel Data Econometrics* (Oxford University Press, Oxford, United Kingdom).

Avgerou C (2019) Contextual explanation: Alternative approaches and persistent challenges. *Management Inform. Systems Quart.* 43(3):977–1006.

Bansal G, Zahedi F (2008) The moderating influence of privacy concern on the efficacy of privacy assurance mechanisms for

building trust: A multiple-context investigation. *Proc. Internat. Conf. Inform. Systems* (Association for Information Systems).

Baruh L, Secinti E, Cemalcilar Z (2017) Online privacy concerns and privacy management: A meta-analytical review. *J. Comm.* 67(1): 26–53.

Bearse PM, Bozdogan H, Schlottmann AM (1997) Empirical econometric modelling of food consumption using a new informational complexity approach. *J. Appl. Econometrics* 12(5):563–592.

Belanger F, Hiller JS, Smith WJ (2002) Trustworthiness in electronic commerce: The role of privacy, security, and site attributes. *J. Strategic Inform. Systems* 11(3-4):245–270.

Bertrand M, Mullainathan S (2001) Do people mean what they say? implications for subjective survey data. *Amer. Econom. Rev.* 91(2):67–72.

Bond T, Fox CM (2015) *Applying the Rasch Model: Fundamental Measurement in the Human Sciences* (Routledge, New York).

Boyd D, Hargittai E (2010) Facebook privacy settings: Who cares? *First Monday* 15(2010):8.

Child D (2006) *The Essentials of Factor Analysis*, 3rd ed. (Bloomsbury Academic Press, New York).

Cook TD, Campbell DT (1976) Design and conduct of quasi-experiments and true experiments in field settings. Dunnette MD, ed. *Handbook of Industrial and Organizational Psychology* (Rand McNally, Chicago), 223–326.

Cooperman A, Smith G, Alper B, Ritchey K (2014) Religion and electronic media: One-in-five Americans share their faith online. Accessed March 1, 2020, https://www.pewforum.org/2014/11/06/religion.

Cronbach LJ (1957) The two disciplines of scientific psychology. *Amer. Psych.* 12(11):671–684.

Dempster AP, Laird NM, Rubin DB (1977) Maximum likelihood from incomplete data via the EM algorithm. *J. Roy. Statist. Soc. B* 39(1):1–22.

Dienlin T, Trepte S (2015) Is the privacy paradox a relic of the past? an in-depth analysis of privacy attitudes and privacy behaviors. *Eur. J. Soc. Psych.* 45(3):285–297.

Dinev T, Hart P, Mullen MR (2008) Internet privacy concerns and beliefs about government surveillance–an empirical investigation. *J. Strategic Inform. Systems* 17(3):214–233.

Finch WH, Hernández Finch ME (2020) Modeling of self-report behavior data using the generalized covariates in a uniform and shifted binomial mixture model: An empirical example and Monte Carlo simulation. *Psych. Methods* 25(1):113–127.

Gates GW (2011) How uncertainty about privacy and confidentiality is hampering efforts to more effectively use administrative records in producing us national statistics. *J. Privacy Confidentiality*, https://doi.org/10.29012/jpc.v3i2.599.

Goldfarb A, Tucker C (2012) Shifts in privacy concerns. *Amer. Econom. Rev.* 102(3):349–353.

Hagenaars JA, McCutcheon AL (2002) *Applied Latent Class Analysis* (Cambridge University Press, Cambridge, United Kingdom).

Hoffmann CP, Lutz C, Ranzini G (2016) Privacy cynicism: A new approach to the privacy paradox. *Cyberpsychology* 10(4):7.

Hollander EP (1958) Conformity, status, and idiosyncrasy credit. *Psych. Rev.* 65(2):117–127.

Hong W, Thong JY (2013) Internet privacy concerns: An integrated conceptualization and four empirical studies. *Management Inform. Systems Quart.* 37(1):275–298.

Hunter JE, Schmidt FL (2004) *Methods of Meta-Analysis: Correcting Error and Bias in Research Findings* (Sage, Los Angeles).

Iannario M (2010) On the identifiability of a mixture model for ordinal data. *Metron* 68(1):87–94.

Iannario M (2012) Modelling shelter choices in a class of mixture models for ordinal responses. *Statist. Methods Appl.* 21(1):1–22.

Iannario M, Manisera M, Piccolo D, Zuccolotto P (2020) Ordinal data models for no-opinion responses in attitude survey. *Sociol. Methods Res.* 49(1):250–276.

Jawahar I, Williams CR (1997) Where all the children are above average: The performance appraisal purpose effect. *Personnel Psych.* 50(4):905–925.

John LK, Acquisti A, Loewenstein G (2011) Strangers on a plane: Context-dependent willingness to divulge sensitive information. *J. Consumer Res.* 37(5):858–873.

Johns G (1991) Substantive and methodological constraints on behavior and attitudes in organizational research. *Organ. Behav. Human Decision Processes* 49(1):80–104.

Johns G (2006) The essential impact of context on organizational behavior. *Acad. Management Rev.* 31(2):386–408.

Johnson VE, Albert JH (2006) *Ordinal Data Modeling, Statistics for Social and Behavioral Sciences* (Springer Science & Business Media, New York).

Kozlowski SW, Klein KJ (2000) A multilevel approach to theory and research in organizations: Contextual, temporal, and emergent processes. Kozlowski SW, Klein KJ, eds. *Multilevel Theory, Research and Methods in Organizations: Foundations, Extensions, and New Directions* (Jossey-Bass, San Francisco), 3–90.

Krosnick JA, Holbrook AL, Berent MK, Carson RT, Michael Hanemann W, Kopp RJ, Cameron Mitchell R, et al. (2002) The impact of "no opinion" response options on data quality: Non-attitude reduction or an invitation to satisfice? *Public Opinion Quart.* 66(3):371–403.

Le H, Schmidt FL (2006) Correcting for indirect range restriction in meta-analysis: Testing a new meta-analytic procedure. *Psych. Methods* 11(4):416–438.

Leavitt K, Mitchell TR, Peterson J (2010) Theory pruning: Strategies to reduce our dense theoretical landscape. *Organ. Res. Methods* 13(4):644–667.

Lipka M (2015) Religious 'nones' are not only growing, they're becoming more secular. URL https://www.pewresearch.org/fact-tank/2015/11/11/religious.

Madden M, Rainie L, Zickuhr K, Duggan M, Smith A (2014) Public perceptions of privacy and security in the post-Snowden era. Technical report, Pew Research Center, Washington, DC.

Malhotra NK, Kim SS, Agarwal J (2004) Internet users' information privacy concerns (IUIPC): The construct, the scale, and a causal model. *Inform. Systems Res.* 15(4):336–355.

Manisera M, Zuccolotto P (2014) Modeling" don't know" responses in rating scales. *Pattern Recognition Lett.* 45:226–234.

Marreiros H, Tonin M, Vlassopoulos M, Schraefel M (2017) "Now that you mention it": A survey experiment on information, inattention and online privacy. *J. Econom. Behav. Organ.* 140:1–17.

McLachlan GJ, Peel D (2004) *Finite Mixture Models* (John Wiley & Sons, New York).

Meyer RD, Dalal RS, Hermida R (2010) A review and synthesis of situational strength in the organizational sciences. *J. Management* 36(1):121–140.

Meyer RD, Kelly ED, Bowling NA (2020) Situational strength theory: A formalized conceptualization of a popular idea. Rauthmann JF, Sherman RA, Funder DC, eds. *The Oxford Handbook of Psychological Situations* (Oxford University Press, Oxford, United Kingdom), 79–95.

Mischel W (1977) The interaction of person and situation. Magnusson D, Endler NS, eds. *Personality at the Crossroads: Current Issues in Interactional Psychology* (Lawrence Erlbaum Associates, Mahwah, NJ), 333–352.

National Academies of Sciences, Engineering, and Medicine (2018) *Open Science by Design: Realizing a Vision for 21st Century Research* (National Academies Press, Washington, DC).

Nissenbaum H (2004) Privacy as contextual integrity. *Washington Law Rev.* 79:119.

Nissenbaum H (2009) *Privacy in Context: Technology, Policy, and the Integrity of Social Life* (Stanford University Press, Stanford, CA).

Piccolo D (2003) On the moments of a mixture of uniform and shifted binomial random variables. *Quaderni Statistica* 5(1): 85–104.

Piccolo D, Simone R (2019) The class of CUB models: Statistical foundations, inferential issues and empirical evidence. *Statist. Methods Appl.* 28(3):389–435.

Pierce JR, Aguinis H (2013) The too-much-of-a-good-thing effect in management. *J. Management* 39(2):313–338.

Pinchot JL, Paullet KL (2012) What's in your profile? Mapping Facebook profile data to personal security questions. *Issues Inform. Systems* 13(1):284–293.

Powell EE, Baker T (2014) It's what you make of it: Founder identity and enacting strategic responses to adversity. *Acad. Management J.* 57(5):1406–1433.

Punzo G, Castellano R, Buonocore M (2018) Job satisfaction in the "big four" of Europe: Reasoning between feeling and uncertainty through CUB models. *Soc. Indicators Res.* 139(1): 205–236.

Rader E (2014) Awareness of behavioral tracking and information privacy concern in Facebook and Google. *Proc. 10th Sympos. Usable Privacy Security*, 51–67.

Reynolds B, Venkatanathan J, Gonçalves J, Kostakos V (2011) Sharing ephemeral information in online social networks: Privacy perceptions and behaviours. *IFIP Conf. Human-Comput. Interaction*.

Rousseau DM, Fried Y (2001) Location, location, location: Contextualizing organizational research. *J. Organ. Behav.* 22(1):1–13.

Shadish WR, Cook TD, Campbell DT (2002) *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*, 2nd ed. (Houghton Mifflin, Boston).

Sheehan KB (1999) An investigation of gender differences in on-line privacy concerns and resultant behaviors. *J. Interactive Marketing* 13(4):24–38.

Sheehan KB, Hoy MG (1999) Flaming, complaining, abstaining: How online users respond to privacy concerns. *J. Advertising* 28(3):37–51.

Smith HJ, Dinev T, Xu H (2011) Information privacy research: An interdisciplinary review. *Management Inform. Systems Quart.* 35(4):989–1016.

Smith HJ, Milberg SJ, Burke SJ (1996) Information privacy: Measuring individuals' concerns about organizational practices. *Management Inform. Systems Quart.* 20(2):167–196.

Solove DJ (2002) Conceptualizing privacy. *Calif. Law Rev.* 90:1087.

Spector PE, Jex SM, Chen PY (1995) Relations of incumbent affect-related personality traits with incumbent and objective measures of characteristics of jobs. *J. Organ. Behav.* 16(1):59–65.

Sweeney L (2000) Simple demographics often identify people uniquely. *Health* 671:1–34.

Teicher H (1954) On the factorization of distributions. *Ann. Math. Statist.* 25(4):769–774.

Van Bavel JJ, Mende-Siedlecki P, Brady WJ, Reinero DA (2016) Contextual sensitivity in scientific reproducibility. *Proc. Natl. Acad. Sci. USA* 113(23):6454–6459.

Vrieze SI (2012) Model selection and psychological theory: A discussion of the differences between the AIC and the BIC. *Psych. Methods* 17(2):228–243.

Wittes B, Liu JC (2015) *The privacy paradox: The privacy benefits of privacy threats.* Reprot, Center for Technology Innovation at Brookings, Washington, DC.

Xie JL, Johns G (1995) Job scope and stress: Can job scope be too high? *Acad. Management J.* 38(5):1288–1309.

Xu H, Teo HH, Tan B (2005) Predicting the adoption of location-based services: The role of trust and perceived privacy risk. *Proc. Internat. Conf. Inform. Systems*.

Xu H, Teo HH, Tan BC, Agarwal R (2012) Effects of individual self-protection, industry self-regulation, and government regulation on privacy concerns: A study of location-based services. *Inform. Systems Res.* 23(4):1342–1363.