© 2016, American Marketing Association
Journal of Marketing
PrePrint, Unedited
All rights reserved. Cannot be reprinted without the express
permission of the American Marketing Association.

# Data Privacy: Effects on Customer and Firm Performance

**Kelly D. Martin**
Associate Professor of Marketing and FirstBank Faculty Fellow
Colorado State University
Rockwell Hall 012, 1278 Campus Delivery
Fort Collins, CO 80523-1278
970.491.7269
kelly.martin@colostate.edu

**Abhishek Borah**
Assistant Professor
Michael G. Foster School of Business
University of Washington
4295 E. Stevens Way NE
485 PACCAR Hall, Box 353226
Seattle, WA 98195-3200
206.543.4569
abhi7@uw.edu

**Robert W. Palmatier**
Professor of Marketing
John C. Narver Chair in Business Administration
University of Washington
4295 E. Stevens Way NE
418 Paccar Hall, Box 353226
Seattle, WA 98195-3200
206.543.4348
palmatrw@uw.edu

***Acknowledgement***

The authors thank the Monfort Family Foundation (Colorado State University) for financial support of this research. They also thank the Marketing Science Institute (MSI) for funding and inclusion of an early draft of this paper in their Working Paper Series (4-1922).

# Data Privacy: Effects on Customer and Firm Performance


## *Abstract*

Although marketers increasingly rely on customer data, firms have little insight into the ramifications of these uses or how to prevent negative effects. Data management efforts may heighten customers' vulnerability worries or create real vulnerability. Using a conceptual framework grounded in gossip theory, this research links customer vulnerability to negative performance effects. Three studies show transparency and control in firms' data management practices can suppress the negative effects of customer data vulnerability. Experimental manipulations reveal that mere access to personal data inflates feelings of violation and reduces trust. An event study of data security breaches affecting 414 public companies also confirms negative effects, as well as spillover vulnerabilities from rival firms' breaches, on firm performance. Severity of the breach hurts the focal firm but helps the rival firm, which provides some insight into mixed findings in prior research. Finally, a field study with actual customers of 15 companies across three industries demonstrates consistent effects across four types of customer data vulnerability and confirms that violation and trust mediate the effects of data vulnerabilities on outcomes.


*Keywords*: data breach; consumer vulnerability; privacy; data transparency, data control, spillover effects, big data

Managers and academics alike contend that collecting and using customer data is an effective way to improve marketing returns (McAfee and Brynjolfsson 2012; Schumann, von Wangenheim and Groene 2014). Consultants suggest that firms can use customer information to generate "productivity and profit gains that are 5 to 6 percent higher than those of the competition" (Biesdorf, Court, and Willmott 2013, p. 40). In turn, firms spend $36 billion annually to capture and leverage customer data (Columbus 2014). However, such efforts also increase *customers' data vulnerability*, or perceptions of susceptibility to harm due to unwanted uses of their personal data, such as those that can result from data breaches or identity theft. Thus data collection efforts may have a dark side, and customers often express negative reactions to privacy practices (Marcus and Davis 2014). Yet firms have little insight into the potential ramifications of customer data management efforts or how to prevent negative outcomes. Thus, we aim to enhance understanding of *the effect of customer data vulnerabilities on customer behavior and firm performance, as well as key mediating mechanisms and mitigation strategies*.

We argue that customer perceptions of vulnerability to harm due to firm data practices better conceptualizes data management effects than privacy concerns. Using gossip theory, we predict strong negative responses to disclosures of personal information by "gossipers"—or firms in this case (Foster 2004; Richman and Leary 2009). Yet gossip theory also identifies two key factors that might suppress the damaging effects of data vulnerability: transparency and control. With a predicted continuum of potential harm, we evaluate the distinct effects of *data access vulnerability* (the firm has access to the customer's personal data), *data breach vulnerability* (the firm or a close rival suffers a data breach), and *data manifest vulnerability* (a data breach allows customer data to be misused, such as identity theft) on the firm itself. For example, if customers provide their personal information to a retailer such as Home Depot, they

experience data access vulnerability. If Home Depot suffers a data breach, the potential for harm becomes more salient, to its own and to Lowe's customers, even if they are not directly affected.

To test this conceptual model, we conduct three complementary studies. In Study 1, we run a series of experiments to delineate the effects of data access vulnerability from a customer's perspective. We examine how firms' mere access to customer information creates specific negative emotional and cognitive outcomes. By manipulating data access vulnerability, transparency, and control, we also provide a strong test of mitigation strategies. An event study in Study 2 investigates the customer vulnerability created by 414 data security breaches that affected 261 public companies. We analyze stock price data for both the breached firms and their closest rivals; we also consider firm policies that might mitigate this harm, such as the provision of more transparent information about data uses or granting greater control to consumers over the uses of their personal information. Finally, Study 3 examines all four types of data vulnerability (access, breach, spillover, and manifest) with a field study involving actual customers of 15 companies, whose transparency and control practices we captured from current privacy policies. This study confirms that suppression occurs across all types of vulnerability and substantiates proposed mediating mechanisms of violation and trust on customer outcomes.

This work contributes to extant literature in four ways. First, by assessing *customers' feelings of vulnerability, we provide a theoretical foundation for understanding how firms' data management practices affect customer behaviors and firm performance.* A customer-centric perspective is rare in descriptions of the effects of information management on performance. However, customer data vulnerability parsimoniously captures multiple salient aspects, including privacy concerns, data breaches, and identity theft, without subjecting customers to real financial harm. The negative customer effects appear mainly due to anxiety about the potential for data misuse and feelings of violation, rather than actual data misuse (Scharf 2007). Capturing the

effects of this sense of vulnerability thus is critical. The perceptions of data vulnerability negatively affect performance across the continuum; for example, among the respondents in Study 3, 10% report they would be more likely to fabricate their personal information, 23% would be more likely to speak negatively, and 22% would be more likely to switch when a firm simply accesses their personal data. In Study 2, we find that an actual data breach reduces the focal firm's stock value by -.29% and its closest rival's by -.17%.

Second, the *severity of a data breach by the focal firm determines whether spillovers to its closest rival have positive or negative effects on performance*, a finding that helps resolve some mixed prior findings (Ko and Dorantes 2006; Malhotra and Malhotra 2011). The severity of a data breach aggravates its negative effect on the firm's stock price, whereas this effect of severity switches for the rival firm. That is, as the severity of the breach at the focal firm increases, it improves the rival firm's performance. Two mechanisms operate on rival firms during a focal firm's data breach: a *negative spillover effect* due to concerns about a similar data breach at the rival firm, and an offsetting *positive competitive effect* that benefits the rival firm, because customers of the damaged focal firm might switch to this rival. Thus at low levels of severity (-2SD), the net effect of a data breach by the focal firm on a rival firm is -.7%, whereas at high levels, the net effect reaches +1.7%.

Third, with our application of gossip theory we *identify and test two managerially relevant mitigation strategies that are effective across the range of data vulnerabilities*. Making a firm's data management policies more transparent and providing customers with control over their data can suppress the negative effects of vulnerability on performance. These strategies also interact to suppress even further the negative performance effects on focal firms, spillover to rivals, and even the negative effects of identity theft. The consistent beneficial effects—shown across three studies using event study and experiment methodologies, measured at both firm and

customer levels, with different operationalizations—strongly support their mitigation of the negative effects of customer data vulnerability. For example, according to model-free median split analyses, firms with low (vs. high) transparency, experience a 1.5 times larger drop in stock price after a data breach. Firms that offer high control suffer no effect of breaches on their stock price, whereas firms that offer low control experience negative returns of -.3%. High transparency and low control is the worst strategy, prompting consumers to express willingness to pay a 5% price premium to switch firms, compared with a low–low condition (Study 3).

Fourth, *two mediating mechanisms, emotional violation and cognitive trust, effectively link all manifestations of customer data vulnerability to performance*. Access to a customer's personal or sensitive data alone increases perceptions of vulnerability, causing customers to feel violated and reduce their trust in the firm. We also show in Study 3 that emotional violation and cognitive trust mediate the effects of data vulnerability on customers' falsifying behaviors, negative word of mouth, and switching behaviors. These mediating effects prove notably robust, across industries, types of data vulnerability, and demographic characteristics.

## Understanding Customer Data Vulnerability

As firms expand their efforts to collect and use customer data, customers grow more concerned about their privacy and the potential for harm. These concerns often are labeled "privacy issues," though the construct of privacy is relatively amorphous and cannot capture the essence of customers' psychological attitudes, such that "[p]rivacy is a concept in disarray. Nobody can articulate what it means" (Solove 2006, p. 476; see also Martin and Murphy [2016] for a comprehensive review of privacy literature in marketing). We propose *customer data vulnerability*, or a customer's perception of his or her susceptibility to being harmed due to various uses of his or her personal data, instead is a critical construct for privacy literature in that it drives customers' responses to firms' efforts to collect and use their data. Gossip theory

describes how people respond to the unsanctioned collection, use, or disclosure of their personal information (Dunbar 2004; Foster 2004), such that we consider it germane for understanding how customers respond when firms collect and use their personal data too.

**Customer Data Vulnerability**

Vulnerability implies susceptibility to injury or harm (Smith and Cooper-Martin 1997). When a firm collects, stores, and uses customers' personal information, it increases the potential for harm and thus their feelings of vulnerability. Most negative customer effects resulting from data uses thus stem from customers' anxiety about the potential for damage or feelings of violation, rather than actual data misuses or financial or reputation harm (Scharf 2007). As legal perspectives argue, customers experience harm at the moment of the breach, regardless of whether their data subsequently are misused (Fisher 2013). Therefore, it is critical to capture the effects of customers' vulnerability, rather than focusing only on damages.

We delineate customer data vulnerability along a continuum of potential harm (see Panel A, Figure 1). The most benign form exists when companies have access to a customer's personal data, or *data access vulnerability*. This mere access means that firms have "detailed digital dossiers about people" and can engage in "widespread transfer of information between a variety of entities" (Solove 2003, p. 2). Customers limit how and with whom they share sensitive information to reduce this vulnerability, using disclosure management processes such as reactance or refusal (Acquisti, John, and Loewenstein 2012). Yet companies already possess and continue to actively seek increasing volumes of customer information, such that data access vulnerability is a widespread and growing concern for customers (Tucker 2014).

*Data breach vulnerability* increases customers' perceptions of susceptibility to harm even more, because it implies a firm that already has their private data, or one of its close rivals, suffers an actual security lapse. The U.S. Identity Theft Resource Center estimates that nearly

130 million personal records have been subjected to risk from data breaches (idtheftcenter.org). Ultimately, not everyone whose records have been compromised experiences victimization, but the unknown scope and lack of control over this threat makes this type of vulnerability especially troubling to customers. The perception of vulnerability increases due to data breaches at firms that possess the customer's data (focal firm) but also, indirectly, due to breaches at close competitors (rival firms), because these events increase the salience of the belief that similar breaches are possible.

This latter spillover effect (*spillover vulnerability*) arises when customers perceive greater susceptibility to harm because a firm similar to one that has their data suffers a data breach. Our proposed continuum (Figure 1) shows that spillover creates less vulnerability than a data breach at a focal firm a customer actually uses. Although vulnerability is made salient to a customer when a close competitor firm suffers a breach, we expect this vulnerability is less than when a focal firm suffers a breach. Nonetheless, to illustrate, analysts assessing the damage to Home Depot's stock price in the wake of its 2014 data breach accurately predicted negative effects for Lowe's too (Trefis Team 2014).

Finally, *data manifest vulnerability* occurs when customer data actually are misused, causing harm to the customer. Disclosures and fraudulent activities represent the most severe form of vulnerability, by moving beyond susceptibility to a state of actual harm. Even when the actual damage that a customer experiences is minor, the event significantly increases perceptions of data vulnerability. Thus, the *greatest effects tend to stem not from actual data misuse but from accompanying feelings of violation and the indeterminate nature of the threat* (Anderson 2013; Scharf 2008; Solove 2003).

In contrast with research that focuses on customer privacy perceptions, empirical studies of customer data management primarily address how customers disclose personal information

(Moon 2000; White 2004) and begin to trust firms, due to their data management processes (Bart et al. 2005; Schlosser, White, and Lloyd 2006). A separate but related literature stream investigates data security breaches and their effects on a firm (e.g., Hseih et al. 2015; Sen and Borle 2015). We summarize selected relevant literature in Table 1, revealing that research into how data management affects *both customers and the firm* is relatively limited.

−Insert Figure 1 and Table 1 about here−

**Gossip Theory**

Customers' psychological and behavioral responses to feelings of vulnerability can be informed by gossip theory, considering the common notion of unsanctioned transmissions of personal information about a vulnerable third party. *Gossip* is evaluative communication about an absent third party (Feinberg et al. 2012; Foster 2004), and gossip researchers report that approximately two-thirds of all communications in public social settings are devoted to such social topics (Dunbar 2004). Thus, most people are adept at detecting gossip, guarding against becoming a gossip target, and minimizing their vulnerability to it (Beersma and Van Kleef 2012; Mills 2010). When they learn they are the target of gossip, people typically react negatively (Baumeister, Zhang, and Vohs 2004), with a range of negative emotional and cognitive responses (Leary and Leder 2009), including heightened feelings of betrayal and violation (Richman and Leary 2009) and deteriorating levels of trust (Turner et al. 2003). Thus, applying gossip theory to a business context suggests that customer data vulnerability may lead to feelings of emotional violation and lowered cognitive evaluations of trust.

Gossip theory also identifies two factors that suppress the negative effects of unsanctioned transmissions of information: transparency and control. Transparency implies the target's awareness of and details about which information is being shared. The gossip target knows the scope of potential harm and can develop strategies to counter negative effects. Control

is the extent to which the target believes she or he can manage the flow of information (Emler 1994). A perceived lack of control over personal information, upon learning about its transmission, exacerbates negative affect surrounding a gossip event, even if the valence of the information being spread is not negative (Feinberg et al. 2012). As two forms of empowerment, control and transparency thus may help people manage the negative effects of their own vulnerability (Baker, Gentry, and Rittenburg 2005).

## Effect of Data Access Vulnerability on Customer Behaviors (Study 1)

Our research progression reflects the proposed continuum of customer data vulnerability (Figure 1). In Study 1, we investigate customer response to data access vulnerability, the most benign form, which implies only the potential for harm when a firm has access to customer personal information. Accordingly, it constitutes a conservative test of the conceptual model. We use a series of experiments and manipulate data access vulnerability, transparency, and control to test the effects of these theoretically derived suppressors. In Study 2, we use an event study methodology to capture the effects of data breach and spillover vulnerabilities and to determine if those effects can be suppressed by transparency and control. Finally, in the field study with actual customers and firm privacy policies, Study 3 manipulates each type of vulnerability to test the suppressors and mediation across multiple outcomes.

### Data Vulnerability Effects and Suppressors

The negative reactions of gossip targets to learning about a gossip event can manifest as emotions and as cognitive-based judgments, often experienced simultaneously (Richman and Leary 2009). Negative emotions may take the form of hurt feelings, mental states of betrayal, or feelings of violation (Mills 2010; Williams 2007). In business, customers' feelings of violation appear in the form of backlash, in conjunction with their more generalized feelings of anger and betrayal (Marcus and Davis 2014). Furthermore, whether negatively or positively valenced,

gossip often leads to deteriorated trust (Turner et al. 2003), as do customers' concerns about online security (Bart et al. 2005; Schlosser, White, and Lloyd 2006). Thus, we expect customer data vulnerability to affect both the emotional mechanism of violation and the cognitive mechanism of trust. *Emotional violation* captures a customer's negative affect, resulting from a perception of a firm's failure to respect her or his peace, privacy, or other rights (Grégiore and Fisher 2008). *Cognitive trust* instead is the customer's willingness to rely on a firm in which he or she has confidence (Palmatier 2008).

Gossip theory advises that *data use transparency* (hereafter "transparency") provides customers with information about how the firm collects, shares, and protects their data. Transparency grants customers knowledge about what information they provide to the firm, how it is used, and which partner firms may access that data. In addition, *customer control* (hereafter "control") over information use and data management decisions should help customers feel empowered in high vulnerability contexts, which may suppress their feelings of violation (Kumar, Zhang, and Luo 2014; Tucker 2014). With control, a customer can determine whether to participate in certain forms of data sharing, which reduces uncertainty and perceptions of sneakiness. When data access vulnerability already is low, these perceptions likely are weak anyway, so providing customers with transparency and control should have little effect on either violation or trust. However, it could suppress damaging effects on violation and trust when data access vulnerability is high.

Specifically, we propose that transparency and control, separately and interactively, mitigate the damaging effects of all types of customer data vulnerability on firm- and customer-level performance effects, including the positive effect on violation and the negative effect on trust (Baumeister, Zhang, and Vohs 2004). Prior research shows that negative responses to gossip diminish with disclosures of the facts of the situation (Beersma and Van Kleef 2012),

suggesting the suppressing effect of transparency. Customers also might choose to engage in some company data practices but opt out of others. Providing knowledge and granting control are positive signals of the firm's intentions too, so they should suppress the negative link between high vulnerability and trust.

Finally, the interaction of transparency and control may suppress the damaging effects on both violation and trust when data access vulnerability is high. If firms provide customers with both transparency and control, the combination should generate strong feelings of empowerment, even if their vulnerability is significant (Baker, Gentry, and Rittenburg 2005). Empowerment then can reduce expectations of perceived harm due to data access vulnerability, because customers believe they have knowledge about and control over the uses of their data, which mitigates their negative emotional responses and attributions (Emler 1994).

**$H_1$**: The positive effect of data access vulnerability on emotional violation is suppressed by (a) transparency, (b) control, and (c) the interaction of transparency x control.

**$H_2$**: The negative effect of data access vulnerability on cognitive trust is suppressed by (a) transparency, (b) control, and (c) the interaction of transparency x control.

**Experimental Data and Design**

We used a series of 2 x 2, between-subjects experiments to assess customer responses to firms' mere access to their data. In three experiments (Studies 1a–c), we manipulated high and low levels of (1) data access vulnerability × transparency, (2) data access vulnerability × control, and (3) transparency × control. All constructs, definitions, and operationalizations are in Table 2. We sought participants from Amazon's Mechanical Turk to gauge customer insights across a range of demographic profiles and backgrounds. We recruited 200 respondents for each of the three experiments for 50 participants per cell. We created scenario descriptions (Appendix A) to convey high and low levels of each manipulated variable, presented in a randomized design.

After reading the descriptions of data access vulnerability, transparency, and control, respondents evaluated the scenario company on measured scales for violation and trust (see Appendix B).

In Studies 1a and 1b, we investigated the ability of transparency and control, respectively, to mitigate potential damaging effects of data access vulnerability on violation and trust. Two separate between-subject experiments served to test our hypotheses with scenarios that placed participants in a situation of high/low vulnerability and then high/low transparency or control. In Study 1c, we sought to understand whether transparency and control worked interactively to influence violation and trust in situations marked by high data access vulnerability.

−Insert Table 2 about here−

**Results**

Manipulation checks with measured variables showed the experimental conditions differed significantly ($p < .01$), as expected (Appendix A), but that transparency and control were not evaluated differently across vulnerability manipulations ($p > .10$). In Study 1a, transparency significantly suppressed the positive effect of vulnerability on violation, in support of $H_{1a}$ ($M_{\text{high vuln/high trans}} = 2.78$, $M_{\text{high vuln/low trans}} = 4.65$, $M_{\text{low vuln/high trans}} = 2.18$, $M_{\text{low vuln/low trans}} = 3.23$; $F_{(1, 196)} = 4.43$, $p < .05$). Similarly, Study 1b revealed that control significantly suppressed the positive effect of vulnerability on violation, in support of $H_{1b}$ ($M_{\text{high vuln/high cont}} = 3.29$, $M_{\text{high vuln/low cont}} = 5.11$, $M_{\text{low vuln/high cont}} = 2.45$, $M_{\text{low vuln/low cont}} = 3.33$; $F_{(1, 196)} = 4.17$, $p < .05$). Although the effects were significant and in the predicted direction, neither the vulnerability x transparency nor the vulnerability x control interaction was significant for trust, failing to support $H_{2a}$ and $H_{2b}$.

In Study 1c, we tested the interactive effects of transparency and control on violation and trust. Keeping high data access vulnerability constant, we support both $H_{1c}$ and $H_{2c}$. Specifically, the transparency x control interaction suppressed the positive effect of data access vulnerability on emotional violation ($M_{\text{high trans/high cont}} = 2.37$, $M_{\text{high trans/low cont}} = 4.50$, $M_{\text{low trans/high cont}} = 4.25$,

$M_{low\ trans/low\ cont} = 4.85$; $F_{(1, 195)} = 16.29$, $p < .01$) and the negative effect on cognitive trust ($M_{high\ trans/high\ cont} = 5.50$, $M_{high\ trans/low\ cont} = 3.72$, $M_{low\ trans/high\ cont} = 4.20$, $M_{low\ trans/low\ cont} = 3.29$; $F_{(1, 195)} = 5.68$, $p < .05$). Planned contrasts show that the high transparency x high control cell creates significantly lower violation when compared with all other combinations ($p < .01$), as well as promotes the greatest reported trust ($p < .01$). Transparency and control independently suppress the effect of high data access vulnerability on emotional mechanisms; together, they suppress both the positive effect of vulnerability on violation and its negative effects on trust. This ideal mix allows customers to understand how firms collect and use data, and also have a say in how (or whether) that happens.

## Effect of Data Breach Vulnerability on Firm Performance (Study 2)

To move along the progression of increasingly severe forms of customer data vulnerability, Study 2 uses an event study to examine the effects of data breach vulnerability and spillover, as experienced by customers and anticipated by the market, on abnormal stock returns. We examine the data breach vulnerability created by a focal firm and the impact on its closest rival. Then we evaluate the moderating effects of transparency and control, independently and interactively, to provide a more externally valid test of these potential suppressors.

### Customer Data Breach Vulnerability

Most research into data breach vulnerability narrowly emphasizes firm characteristics (e.g., industry, firm size, past breach), ignoring the customer's central role in driving performance outcomes or mitigation strategies. Cybersecurity failures and data breaches are on the rise, affecting growing numbers of firms and their customers from various industries. The damages are significant: One study puts the average firm cost per breach at $3.8 million (Ponemon Institute 2015). Extant research generally shows that data breaches lead to negative abnormal stock returns for firms (e.g., Acquisti, Friedman, and Telang 2006; Malhotra and

Malhotra 2011). Because the market captures customer sentiment, including the potential for decreased use or defection, we investigate if a customer-focused theory (i.e., gossip theory) can explain a data breach, in accordance with our customer data vulnerability conceptualization.

Customer data vulnerability becomes especially salient when customers realize that a firm has experienced a data breach. Legal perspectives argue that customers experience psychological and emotional harm at the moment of the breach, regardless of whether their data subsequently are misused (Fisher 2013). According to gossip theory, when a gossip target (customer) realizes vulnerability in the form of compromised personal information, the primary responses are negative cognitive and emotional reactions to the gossiper (firm) (Mills 2010; Richman and Leary 2009). We thus seek to extend current thinking about data breach effects by incorporating the customer (Figure 1, Panel B). The detrimental impact of a data breach on firms likely results from the anticipation of negative customer responses and the perception of insufficient data protection by the firm. These forces combine to damage firm performance, as reflected in its abnormal stock returns. That is, the firm's stock price likely decreases when the efficient market anticipates lost sales to existing customers, increased difficulty acquiring new customers, and potential legal and recovery costs.

As captured in our conceptualization of spillover vulnerability, when negative events garner unfavorable publicity, the influence often spreads to rival firms, through a "guilt-by-association" effect, such that crises can harm firms that represent close rivals to an affected firm (Borah and Tellis 2016). These *negative spillover effects* occur because customers believe the nature or root cause of the crisis is endemic to the entire category or industry (Cleeren, van Heerde, and Dekimpe 2013). After a data breach by a focal firm, customers of rival firms may feel more vulnerable, which creates a cascade of actions and negative spillover to the rival firm's performance, due to anticipation in the stock market.

Yet brand scandal literature offers an alternative perspective, in which a data breach event creates a *positive competitive effect* for the closest rival that mitigates or even offsets the negative spillover. If the breach creates severe negative publicity, customer backlash, and financial harm to the focal firm, that firm's customers might switch to a rival. Customers often shift from a firm experiencing a brand crisis, and the switch ultimately may be permanent (Roehm and Tybout 2006). The rival firm gains sales and profits from new customers, which improves its financial performance. Therefore, we propose alternative hypotheses.

**$H_{3a}$**: Data breach vulnerability negatively affects firm performance.
**$H_{3b}$**: Data breach vulnerability negatively affects a rival firm's performance (spillover effect).
**$H_{3b(alt)}$**: Data breach vulnerability positively affects a rival firm's performance (competitive effect).

**Customer Data Breach Vulnerability Suppressors**

Firms might use several strategies to lessen the detrimental effects of customer data breach vulnerability on performance. Gossip theory suggests a target's vulnerability decreases when the target has knowledge about the gossip event (transparency) and the ability to manage the spread and impact of the information (control) (Mills 2010; Smith 2014). In a customer data breach vulnerability context, we argue that firms' data use transparency (i.e., the extent to which it explains its data collection, use, storage, and protection) and its provision of customer control (i.e., grants customers the ability to determine what information they give to the firm, how it is used, and which partner firms may access those data) can mitigate the damaging effects of all types of customer data vulnerability on firm- and customer-level performance effects.

According to gossip theory, for a target to address a gossip event, it must know the gossip is occurring (Eder and Enke 1991). Transparency then should be a critical suppressor, with the potential to mitigate the harm wrought by customer data breach or spillover vulnerability on performance, because customers gain the knowledge they need to evaluate the potential harm.

Transparency implies that customers have knowledge of the nature and scope of data the firm possesses and how those data are used. Typically, firms provide transparency in the form of a privacy policy or information collection disclosure notification. In addition, company information collection strategies that are overt versus covert in nature influence how customers respond to firms' personalization efforts (Tucker 2014). In a similar sense, transparency appears critical for firms to avoid the "creepiness factor" often associated with data and analytical inferences about customers (Cumbley and Church 2013).

**$H_4$**: The negative effect of data breach vulnerability on firm performance is suppressed by transparency (i.e., suppressing both data breach and spillover effects).

Providing control is another key strategy. When they learn that gossip has occurred, targets often seek to regain control of their information (Emler 1994). Salvaging this control also represents a key restorative element after a damaging gossip event (Williams 2007). Providing customers with control enables them to manage and adjust their personal data preferences with the firm. To bestow control on customers, firms generally rely on opt-in and opt-out decisions (Kumar, Zhang, and Luo 2014) and allow them to manage their individual settings and preferences governing the use of their data. For example, after Facebook suffered a data breach in 2010, it responded with policies and systems that promised to "keep people in control of their information" (Steel and Fowler 2010, p. A1).

**$H_5$**: The negative effect of data breach vulnerability on firm performance is suppressed by control (i.e., suppressing both data breach and spillover effects).

Beyond these distinct suppressive effects, the most potent force for reducing the damaging effects of vulnerability on performance may result from their combined or interactive effect. Customer knowledge (transparency) and control represent key areas for investigation in online privacy research (Caudill and Murphy 2000), and we know of no studies that investigate them empirically as they function together. Yet the methods that gossip targets use to manage

and mitigate unsanctioned transmissions of their information suggest that transparency and control can work concurrently to benefit customers. Strong transparency and control give customers more knowledge of the firm's data management practices and the ability to manage their data portfolio through opt-out choices. Customers who achieve transparency know of the potential harm but no way to manage it; customers who have control can manage their data but have insufficient knowledge to make informed decisions.

**$H_6$**: The negative effect data breach vulnerability on firm performance is suppressed by the interaction of transparency x control (i.e., suppressing both data breach and spillover effects).

When more people receive gossip, the target becomes more vulnerable (Mills 2010; Smith 2014), so the negative reaction upon learning of the event should be greater (Turner et al. 2003). Paralleling this logic, we expect that greater *data breach severity*, or the scope, reach, and impact of the firm's data breach, imposes a more negative effect on the breached firm's performance. This enhanced negative effect might stem from the need for more resources to recover from a more severe breach, the greater number of disgruntled customers who potentially spread negative word of mouth, and the heightened potential for defection. Malhotra and Malhotra (2011) find no effect of breach magnitude on firm performance, but approximately half of their sample lacked information about magnitude. We also expect that the data spillover effect expands in the wake of more severe breach events, because they affect more customers and strengthen the guilt-by-association mechanism. That is, with larger breaches, more customers learn of the breach and are exposed to the negative publicity surrounding it, which makes vulnerability even more salient for customers of rival firms.

Similar to our alternative logic outlined above that a data breach by the focal firm could be beneficial to close rivals, the positive customer gains from a data breach at the rival firm might be enhanced by the severity of the focal firm's data breach. As the data breach grows more

severe, the focal firm's customers may perceive higher levels of vulnerability, increasing their likelihood of defection. The rival firm then can gain sales from customers who defect and should find it easier to acquire new customers, relative to its breached competitor.

**$H_{7a}$**: The negative effect of data breach vulnerability on firm performance (data breach effect) is aggravated by the severity of the focal firm's data breach.

**$H_{7b}$**: The negative effect of data breach vulnerability on rival firm performance (spillover effect) is aggravated by the severity of the focal firm's data breach.

**$H_{7b(alt)}$**: The negative effect of data breach vulnerability on rival firm performance (competitive effect) is alleviated by the severity of the focal firm's data breach.

**Study 2: Methodology**

In Study 2, we evaluate the effects of customer data breach vulnerability on firm performance (i.e., abnormal stock returns), as well as the influence of two managerially relevant vulnerability suppressors (transparency and control). Because of our interest in the precise effect of data breaches, we employ an event study to gauge the impact of data breaches with known timestamps on subsequent stock prices (Srinivasan and Hanssens 2009). An event study leverages the efficient market hypothesis, which states that a stock price at a particular point in time reflects all available information up to that point (Fama 1998; Sharpe 1964). Any change in the stock price due to new information reflects the present value of all expected current and future profits from that new information. We pose our hypotheses according to customers' responses to data breach events, because customer behavior is the primary driver of firm performance. Information about firms' data management practices is available to the overall market (e.g., privacy policies), and market actors try to anticipate the relevance of many diverse factors on future sales and profits. We expect customer-level effects to be manifest in immediate changes in stock price (i.e., efficient market theory). This approach is consistent with previous event studies in marketing (Borah and Tellis 2014; Homburg, Vollmayr, and Hahn 2014).

***Data.*** To analyze the relationship between data breaches and stock returns, we first identify data breaches of publicly traded firms using the Capital IQ, Factiva, Lexis-Nexis, and privacyrights.org databases. We use multiple sources to ensure that the data collection is as exhaustive as possible and to remove any ambiguous breach announcements. The unit of analysis is each specific data breach. We collect stock returns for the firm that suffered the data breach and for its closest rival. Our sample includes any global, publicly-listed firm, so we collect stock price and market index data from various stock exchanges (e.g., NYSE, Paris Stock Exchange, London Stock Exchange). Using Dun and Bradstreet's *Hoover's* Database, we identify the closest (revenues nearest to the focal firm) publicly-listed rival of each focal firm. The initial sample consists of 414 breached firm-day observations, across 261 unique firms, and 414 rival firm-day observations, across 221 unique rival firms. We drop 18 breached firm- and 10 rival firm-day observations, because we could not obtain their stock price data at the time of breach.

Event studies are subject to three important assumptions: market efficiency, unanticipated events, and confounding events. The assumption of market efficiency can be difficult to reconcile with a long event window. Assuming efficient information processing of the breach announcement, the event window ought to be as short as possible (McWilliams and Siegel 1997). Because the market should incorporate data breach information quickly, we use windows ranging from -1 to +1 days around the event to calculate abnormal returns. Finally, we control for an array of confounding events around the -1 to +1 window, including dividend declarations, contract signings, earnings information, or mergers and acquisitions. We drop any observations with confounding events around the three-day breach window, excluding 103 events for focal firms and 105 events for rival firms due to confounds. Ultimately, we get 293 breached firm-day observations across 199 unique firms and 299 rival firm-day observations for 176 unique firms.

***Measures.*** A summary of the definitions and operationalizations of the independent variables is in Table 2. We created *data use transparency* and *customer control* variables using a mix of automation and manual coding. For each focal firm and its closest rival, we obtained the privacy policy statements from the firm's website when the breach occurred, using the Wayback Machine. A newly developed Python code scraped each iteration of the focal firm and its closest rival's privacy policy documents over time, enabling us to select the documents that were current and active on the breach date. After obtaining the relevant privacy policy, we employed manual coding to measure the independent variables, with a careful reading of each privacy policy. We next created scores for both transparency and control, reflecting whether or not (0,1) firms included specific information in their privacy policy (see Table 2). Two research assistants, who did not know the study hypotheses, coded the privacy policy documents with a standardized coding schema. Their interrater agreement was greater than 85%, and all disagreements were resolved through discussion with the first author.

For the transparency independent variable, we used a count of the dummy variables across multiple elements of the privacy policy that signal openness and willingness to provide information to customers. Specifically, we coded whether the firm (1) explains its opt-out policy, (2) explains how it captures data, (3) explains how it uses data, (4) explains its data sharing internally and with third parties, and (5) provides contact information for privacy requests. If a firm's privacy policy has all five characteristics, the policy earns a score of 5 for transparency. To create the control independent variable, we counted the number of opt-out choices in the firm's privacy policy, ranging from 0 to 5. Specifically, we coded whether the consumer (1) can opt out of marketing communication, (2) can opt out of saving data usage (e.g., search history), (3) can opt out of storing personal information (e.g., credit card number), (4) can opt out of sharing data with third parties and (5) can opt out of tracking. Details of the data coding and

measure validation are in Web Appendix A. The measure of data breach severity reflected the natural logarithm of the number of customer records compromised in the data breach. Descriptive statistics and correlations for Study 2 variables are in Table 3, Panel A.

−Insert Table 3 about here−

***Model development and estimation.*** We use a market model to calculate abnormal returns. Abnormal returns to a stock, due to some event, offer controls for fluctuations in price across the whole market. The market model is superior to a capital asset pricing model for cross-sectional event studies (Campbell, Cowan, and Salotti 2010; Homburg, Vollmayr, and Hahn 2014); furthermore, Fama-French and Carhart factors are not available for firms listed in non-U.S. stock exchanges. To gather firm and market stock returns, we relied on the Center for Research in Security Prices (CRSP) and Kenneth French's website for U.S. companies and Thomson ONE and Yahoo Finance for firms listed in non-U.S. exchanges. "Returns" refer to the cumulative average abnormal returns (Web Appendix B). Because we hypothesize that the same mitigation strategies work for focal and rival firms, we pool the data for the breached firm and its nearest rival and use the variable *breach firm* (1 = breached firm, 0 = closest rival) to specify which data belong to each. In turn, the main effects model is as follows:

(1) $\text{Returns} = \beta_0 + \beta_1\text{Breach Firm} + \beta_2\text{Data Use Transparency} + \beta_3\text{Customer Control} + \beta_4\text{Data Breach Severity} + \beta_5\text{Capital Slack} + \beta_6\text{Size} + \beta_7\text{Competitive Intensity} + \beta_8\text{Number of Breaches of Breach Firm} + \beta_9\text{Number of Breaches of Rival Firm} + \beta_{10}\text{Industry Fixed Effects} + \beta_{11}\text{Year Fixed Effects} + \beta_{12}\text{Time Trend} + \varepsilon_i.$

To estimate this model, we use the xtreg command in Stata 13.0. Because the same firm can have multiple breaches in our sample, we use a panel regression and the *vce (cluster firmid)* option to account for clustering by firm. We also estimate a model that includes the interactions of transparency and control (Equation 2) and features separate coefficients for the effect of severity on focal and rival firms, such that we multiply them by breach firm and 1 – breach firm,

respectively. Thus we can test our alternative hypotheses advancing opposite predictions about the effect of severity on rival firm performance. Specifically, $\alpha_5$ captures the impact of severity on the focal firm's returns; $\alpha_6$ captures the impact of severity on the rival's returns.

(2) $\text{Returns} = \alpha_o + \alpha_1\text{Breach Firm} + \alpha_2\text{Data Use Transparency} + \alpha_3\text{Customer Control} + \alpha_4\text{Data Use Transparency*Customer Control} + \alpha_5\text{ Data Breach Severity*Breach Firm} + \alpha_6\text{Data Breach Severity*(1 – Breach Firm)} + \alpha_7\text{Capital Slack} + \alpha_8\text{Size} + \alpha_9\text{Competitive Intensity} + \alpha_{10}\text{Number of Breaches of Breach Firm} + \alpha_{11}\text{Number of Breaches of Rival Firm} + \alpha_{12}\text{Industry Fixed Effects} + \alpha_{13}\text{Year Fixed Effects} + \alpha_{14}\text{Time Trend} + \varepsilon_i.$

## Results

***Univariate analysis of data breaches on returns.*** We first analyze market response for focal and rival firms separately. Data breaches have negative and significant effects for the focal firm at 5% in the (-1, 0) and (-1, +1) window, with average returns of -.29% and -.27% respectively. The Wilcoxon signed rank test reveals that market response to focal firm breaches are negative and significant at 1% in the (-1, 0) window and at 5% in the (0, 0) and (-1, +1) windows. Market response to data breaches are negative and significant for rivals, at 5% in the (0, 0) window with average returns of -.14%, and at marginal significance of 10% in the (-1, 0) window with average returns of -.17%. Wilcoxon signed rank test confirms market response to breaches for rival firms still are negative and marginally significant at 10% for both the (0, 0) and (-1, 0) windows. In support of $H_{3a}$ and $H_{3b}$, a data breach leads to significantly negative abnormal returns for focal and rival firms, implying a negative effect of data breach vulnerability and a negative spillover effect. We reject the alternative hypothesis ($H_{3b(alt)}$) of a positive effect of a data breach on rival firm performance (competitive effect). The effect on returns is about 1.7 times stronger for the focal firm than for its closest rival in the (-1, 0) window.

To test for significant differences in damage between the focal firm and its closest rival, we ran an independent sample t-test for the (-1, 0) window surrounding the breach event. The returns for focal firms do not differ significantly from those of their closest rivals ($p = .47$).

Similarly, a two-sample Wilcoxon rank-sum (Mann-Whitney) test does not indicate any significant difference in returns ($p = .35$). Thus, we analyze abnormal market returns by pooling both focal and rival firms. The effect of data breaches for the pooled data is negative and significant (-.14%, $p < .01$ in the [0, 0] window; -.23%, $p < .01$ in the [-1, 0] window). The event window with the highest absolute value and most significant t-test for the pooled data was the event day (-1, 0) window. Therefore, consistent with previous research, we use this event window for the multivariate analyses (Raassens, Wuyts, and Geyskens 2012). Non-parametric Wilcoxon signed rank test reveals that market response to breaches is negative and significant at $p < .01$ for the (0, 0) and (-1, 0) windows and at $p < .05$ for the (-1, +1) window; the (-1, 0) event window is the most significant (Table 4).

***Multivariate analysis of data breaches x suppressors and severity on returns.*** We estimate Equations 1 and 2 with a regression model that incorporates firm fixed effects, as we have repeated observations for the same firm. We estimate the regression model by pooling both breach and rival firm observations for 583 total firm observations. We drop nine observations from the univariate analysis; we could not obtain independent variable data for these firms. The outcome reflects cumulative abnormal returns in the (-1, 0) window. We estimate Equation 1 with a random (vs. fixed) effects model, because we cannot reject the null hypothesis ($\chi^2(27) = 30.35$, $p = .30$) that the random effects model is consistent and efficient (Hausman 1978).

As we show in Table 5 (Model 1), the coefficient estimate of the dummy-coded breach firm is negative (-.08) but not significant, consistent with the results of our prior analysis. That is, the data breach does not lead to significant differences in the returns of focal versus rival firms. Regarding the role of suppressors, we find that transparency does not have a significant effect on returns (-.05, n.s.), failing to support $H_4$. However, control has a significant and positive effect on

returns (.07, $p < .05$), in support of $H_5$. Providing customers control through opt-out features thus helps suppress the negative effect of data breaches on focal and rival firms' abnormal returns.

We test the interaction hypotheses using Equation 2, again with a random effects model because we cannot reject the null hypothesis ($\chi^2(28) = 25.1$, $p = .62$), and the random effects model is consistent and efficient (Hausman 1978). In support of $H_6$ (Table 5, Model 2), the interaction of transparency and control has a positive impact on returns (.23, $p < .05$). So too, when comparing the high transparency and high control combination (via median split of the coded data), we find that this mix is more effective in securing positive abnormal returns than all other combinations ($p < .05$). For $H_7$, we test the interactions of severity with the dummy-coded breach firm and 1 - breach firm variables to evaluate the effects on focal and rival firms, respectively. The effect of data breach severity on the focal firm's returns is negative and significant (–.16, $p < .05$), as predicted in $H_{7a}$. Its effect on the returns of the closest rival instead is positive and significant (.12, $p < .05$), which offers support for $H_{7b(alt)}$ instead of $H_{7b}$. The effect of severity is asymmetric. As it increases, the effect of the breach becomes worse for the focal firm, but the benefits of the competitive effect mechanism (i.e., customers switch away from the weakened focal firm to the rival) are enhanced. The competitive benefit mechanism thus offsets the negative spillover effect mechanism.

−Insert Tables 4 and 5 about here−

***Robustness analyses.*** We performed a series of robustness analyses to ensure the validity of our assumptions and results. First, we calculated buy-and-hold abnormal returns for one year (250 trading days) and two years (500 trading days). We found no effect on long-term abnormal returns ($p > .10$) for either the breached or rival firms. Thus the effect of data breaches occurs in the short-term window, and there are no corrections in the long run. Second, we ran regression

models without the control variables for both Equations 1 and 2, and the results were consistent with our models that included the control variables.

## Integrated Model of Customer Data Vulnerabilities (Study 3)

To connect the findings from Studies 1 and 2 in a holistic framework derived from gossip theory, in Study 3, we expand our approach and examine all forms of customer data vulnerability in parallel, including data access vulnerability, data breach vulnerability, spillover vulnerability, and data manifest vulnerability (Figure 1, Panel A). We consider how company-level transparency and control might suppress an increase in felt customer vulnerability following a data breach, identity theft, or similar event. In this field study, we query actual customers of the five largest firms across three industries, then use those companies' current data management policies (i.e., same privacy policy coding used in Study 2) to blend key aspects of Studies 1 and 2. In addition, we link vulnerability to customer outcomes, including falsifying personal information, spreading negative word of mouth (WOM), and switching behaviors, to test the proposed mediating mechanisms (violation and trust) while controlling for privacy concerns and participants' prior experience with a data breach or identity theft.

*Falsifying information* occurs when customers fabricate the personal information they provide to a company (Lwin, Wirtz, and Williams 2007). With *negative WOM*, customers spread unflattering information about the firm to family, friends, and acquaintances (De Angelis et al. 2012). *Switching behavior* (switching) implies that customers defect to a competitor to avoid the focal firm's data management practices. All types of data vulnerability should lead to feelings of emotional violation, which then should lead to falsifying, negative WOM, and switching because feelings of violation and betrayal generate strong desires to punish the gossip source, such as by lying, telling others of these practices, and shifting business to other firms (Grégoire and Fisher 2008; Smith 2014). However, customers seek to reward and increase their dealings with trusted

partners, due to norms of reciprocity and reduced perceptions of the likelihood of opportunistic behaviors (Palmatier et al. 2006). Because all types of data vulnerability likely lessen such trust, they also should increase falsifying behaviors, negative WOM, and switching through this route.

**$H_8$**: The positive effects of customer data vulnerability on (a) falsifying behavior, (b) negative WOM and (c) switching behavior are mediated by emotional violation.

**$H_9$**: The positive effects of customer data vulnerability on (a) falsifying behavior, (b) negative WOM and (c) switching behavior are mediated by cognitive trust.

**Experimental Data and Design**

Using experiments, we test all four types of vulnerability with customers by asking them to evaluate companies they presently use. We recruited 202 people from Amazon's Mechanical Turk and assigned them randomly to one of three industries: retail, financial services, or technology. A list of the five largest companies in each industry then appeared, and participants selected the company whose products and services they use most frequently. If they did not use any of those firms, they were directed to one of the remaining two industry lists. Participants were to be excluded from the study only if they did not use any of the 15 total firms across the three industries. Questions related to the extent of company patronage confirmed that respondents were highly involved with their chosen firm (mean product/service use = 6.15 of 7).

After selecting a company, participants answered questions that provided baseline measures of their relationship with that company, including perceptions of vulnerability, violation, trust, falsifying behavior, negative WOM, and switching likelihood. Respondents also indicated the extent of transparency and control provided by the company. After completing these baseline questions, participants saw one of four randomly displayed email messages (50 participants per cell), reportedly from the company they selected. One email explained that the firm had been the victim of a data breach (data breach vulnerability), another indicated that the main competitor of the firm had been the victim of a data breach (spillover vulnerability), and a

third noted that the firm had been the subject of a data breach and the participant thus had been the victim of identity theft (data manifest vulnerability). A fourth and final condition simply alerted the customer to a change in firm privacy policy designed to serve as a control (data access vulnerability). The treatments and manipulation checks for Study 3 are in Appendix D.

After reading their assigned email, participants again completed the measures of vulnerability, violation, trust, falsifying, negative WOM, and switching likelihood, so we have pre- and post-assessments of the study variables. In addition to demographics, we controlled for privacy concerns and participants' experience with a data breach or identity theft. All variables exhibited desirable measurement properties (Appendix B). Descriptive statistics and correlations are in Panel B of Table 3. Manipulation checks with the measured variable showed that the four email treatments differed significantly ($p < .01$) on vulnerability, except for the data breach vulnerability and data manifest vulnerability conditions (see Appendix D). These two treatments were strong predictors of post-vulnerability but not significantly different from each other ($p > .05$). Furthermore, vulnerability measured prior to the treatment differed significantly by industry, as it was significantly higher for technology than for financial or retail companies ($p < .05$). However, removing the baseline effects from the post-vulnerability measures by looking only at the change in vulnerability, we find no systematic differences across industries ($p > .10$).

**Results**

To account for observed heterogeneity in participants' baseline assessments of vulnerability, we used the change in vulnerability (Δvulnerability) as our dependent variable for the first set of analyses. At the firm level, independent coders completed the coding schemes from Study 2 to assess the current privacy policies of the 15 company choices in Study 3 on dimensions of transparency and control. Because customer responses were nested in these 15 firms, and firms were nested in three industries, we used a three-level hierarchical linear model

with HLM 7.01 to test whether firm-level (level 2) transparency and control, measured using actual privacy policies, as well as their interaction, will mitigate Δvulnerability across customers who experience the different types of data vulnerability (level 1).

Model 1 includes the main effects across the three levels; Model 2 includes the level-2 interaction between company privacy policy transparency and control. The results in Table 6 show that separately, company-level control suppresses the change in customer vulnerability, as evidenced by a significant main effect (control: $\beta = -.11$ $p < .05$). The interaction of company-level transparency and control in suppressing Δcustomer vulnerability also is significant ($\beta = -.09$ $p < .05$). These findings indicate that customers are aware of the way company data management policies affect them while privacy policies are a valid proxy for a firm's data management policies. We also controlled for industry type (level 3), the randomly assigned data vulnerability event, personal privacy concern, prior breach experience, and demographics at the customer level, and the findings affirm the robustness of our results. Model 3 introduced transparency and control measured at the customer level using Likert scales. Although company-level analyses represent a stronger test of suppression on Δvulnerability, we confirm that customer perceptions of transparency ($\beta = -.11$ $p < .05$) and control ($\beta = -.12$ $p < .05$) also are significant, paralleling the operationalization from Study 1. Model 4 includes the customer-level interaction of transparency and control, which is significant ($\beta = -.09$ $p < .05$). Again, contrasts (via median split) show that the high transparency x high control combination led to lowest levels of violation ($p < .05$) and highest levels of trust ($p < .01$) across all groupings.

Next, for the mediation analyses, we used a partial least squares (PLS) model. To parse participant heterogeneity from our model, we relied on change variables to incorporate the pre- and post-assessments of the measured variables in our model. In the PLS model, Δvulnerability becomes the antecedent condition, predicting change in violation (Δviolation) and change in trust

(Δtrust). To examine $H_8$ and $H_9$, we investigated whether Δviolation and Δtrust mediate the customer outcome variables, namely, the changes in falsifying (Δfalsifying), negative WOM (Δnegative WOM), and switching likelihood (Δswitching). PLS conventions of resampling through bootstrapping with 500 iterations (Hulland 1999) produced findings in full support of both $H_8$ and $H_9$, which increases confidence in our conceptual model.

We also employed the PROCESS model (Preacher and Hayes 2008) to test the two mediating mechanisms of Δviolation and Δtrust on outcomes. In support of $H_{8a}$, the indirect effect of Δvulnerability on Δfalsifying through Δviolation was significant, with a confidence interval (CI) that excluded 0 (β = .17, CI = [.10, .26], $p < .01$). The indirect effect of Δvulnerability on Δnegative WOM through Δviolation also was significant (β = .07, CI = [.10, .27], $p < .01$), in support of $H_{8b}$. As we predicted in $H_{8c}$, we found a mediating path through Δviolation to Δswitching (β = .53, CI = [.42, .65], $p < .01$). The indirect effect of Δvulnerability on Δfalsifying through Δtrust was supported (β = .21, CI = [.11, .34], $p < .01$), as was the indirect effect of Δvulnerability on Δnegative WOM through Δtrust (β = .22, CI = [.12, .33], $p < .01$). Thus, we confirmed both $H_{9a}$ and $H_{9b}$. Finally, the indirect effect of Δvulnerability on Δswitching through Δtrust received support (β = 1.95, CI = [1.09, 3.01], $p < .01$), in line with $H_{9c}$. Because the direct effects of Δvulnerability on Δfalsifying and Δnegative WOM were not significant ($p > .40$), we have evidence of full mediation, whereas the direct effect on Δswitching was significant ($p < .05$), suggesting partial mediation.

**Post Hoc Examination of Post-Vulnerability Suppressors**

Research on customer privacy suggests additional strategies that might help suppress the damaging effects of vulnerability on violation and trust. Because we worked with actual customers of real companies, we used measured perceptions of fairness and value to understand whether they suppress the effect of a change in vulnerability on changes in trust or violation. We

employed an identical PLS model but included the product term interactions of the change in vulnerability with value and fairness (see Appendix B). The results (see Model 2, Table 7) indicated that neither value nor fairness moderated vulnerability's relationship with violation, yet both value ($\beta = -.44$, $p < .01$) and fairness ($\beta = -.70$, $p < .01$) moderated the vulnerability–trust relationship. That is, fairness and value reinforce a customer's trust in a firm following a data breach. These elements work in favor of trust (cognitive mechanism) rather than violation (emotional mechanism), suggesting the benefits of rational appeals to customers regarding the positive aspects of their relationship with the firm. A detailed examination of these two factors is beyond the scope of this study, but these findings offer interesting future research opportunities.

−Insert Tables 6 and 7 about here−

## Discussion, Limitations, and Research Directions

Firms increasingly emphasize the collection and use of customer data, yet backlash to these practices appears to be growing. We argue that this response is the result of customers' perceptions of vulnerability. Our examinations, at both firm and customer levels, confirm that vulnerability generates negative outcomes for firms, including negative abnormal stock returns and damaging customer behaviors (i.e., falsifying, negative WOM, and switching behaviors). Data transparency and customer control practices can suppress these detrimental effects. We provide a rigorous test of our conceptual framework by offering internally valid insights with a series of experiments using manipulated data access variables, reflecting the most benign form of vulnerability (Study 1). We provide externally valid insights into the effects of customer data breach vulnerability on firm performance, spillover vulnerability effects on rivals, and managerial tools to suppress harm (Study 2). Finally, we examine these insights collectively using a field study, to understand the effects of all types of data vulnerability from the customer's perspective (Study 3). That is, we examine how firms' data management efforts,

creating any type of customer vulnerability, can lead to negative outcomes, and we identify the emotional and cognitive mechanisms through which these negative outcomes occur.

**Theoretical Implications**

This research offers three main theoretical contributions. First, customers perceive harm and respond negatively to firms' collection and use of their data. The tests across all types of customer data vulnerability show significant negative effects, some of which are manifest even without any direct financial harm to the customer. This customer-centric view shows that people identify potential harm due to firms' data management efforts. Accordingly, vulnerability offers a more precise construct to understand customer responses to firms' use of their information than general privacy issues or financial damages. Study 3, testing multiple manifestations of data vulnerability, shows significant effects across each type. Yet in none of the data security conditions was privacy concern a significant predictor of negative customer behavior outcomes. Legal experts already have begun thinking this way about data privacy, noting that "generalized harm already exists; we need not wait for specific abuses to occur" (Solove 2003, p. 8).

Second, we use gossip theory as a unifying lens to describe how customer vulnerability creates strong negative customer responses. Gossip theory has both theoretical and intuitive appeal for evaluating how people respond to unwanted customer information access and use, when they learn of it. In confirmation of a key premise of gossip theory, we find that people have a well-developed sense of how they are perceived and evaluated by others (Richman and Leary 2009), even if those others are firms. Our data breach event study and customer experiments demonstrate that when gossip becomes salient, it produces a range of negative emotional and cognitive responses from the target toward the source (Baumeister and Leary 1995; Leary and Leder 2009). In Study 2, we find significant negative stock performance and spillover effects; in Studies 1 and 3, we demonstrate customers' heightened feelings of violation and deteriorating

trust. Consider, for example, an online post in response to a security flaw by Comcast: “As a consumer, do I just sit and wait for all my stuff to get hacked [feelings of vulnerability]?… It’s very frustrating [emotional response]!” (Macobserver 2014).

Third, we extend two peripheral elements of gossip theory that characterize how people manage gossip’s spread. In our customer-focused investigations in Study 1, the series of experiments confirms that transparency and control work synergistically to mitigate feelings of violation and enhance trust, which aligns with the Study 2 findings that show that transparency and control promises in data management practices reduce the damage to firm performance in the wake of a data breach. In the data breach event study, we coded the different elements of each firm’s privacy policy as a proxy for its data management practices. Similar coded elements in the customer field experiments in Study 3 suppress the increase in vulnerability after a data privacy event. These effects thus demonstrate that people are aware of how companies manage their data, and their practices matter for reducing felt vulnerability. The strong, significant, synergistic effects of transparency and control across three studies, using different measures and in different contexts, speaks to their powerful ability to work in combination to suppress customer vulnerability. Likewise, these combinative effects suggest ties to informed choice theory, stemming from transparency’s emphasis on knowledge and control’s emphasis on choice (Cranage 2004); extensions along these lines represent a fascinating area for further research.

**Managerial Implications**

Our findings suggest that *firms need a more tempered approach to data and analytics initiatives that involve the collection and use of customer information*. They must consider their approaches to data management carefully, to avoid negative effects. Customer data practices may help the firm identify and better understand customers and segments, but these same practices can create vulnerability throughout the customer cohort. In Study 2, we draw on firm privacy

policies to understand how firms access, manage, and communicate about customer data. Our significant findings demonstrate that firms must acknowledge privacy policy dimensions as meaningful proxies for their actual data practices. In Study 3, by blending variables from company privacy policies with individual-level responses, we show that customers are aware of data practices, which affects critical behavioral outcomes. Data practices are important for all firms, considering our spillover effect findings. Even non-compromised firms can suffer substantial financial performance detriments when a close competitor has a breach.

The transparency, control, and breach severity dimensions suggest additional managerial best practices. Transparency and control combine to moderate the relationship between various types of vulnerability and performance. Across three studies and five outcome variables, we find that a potent vulnerability-suppressing combination provides customers with clear transparency and control over their personal information. High transparency and control reduces the spread of negative WOM, deters switching, and suppresses negative stock price effects. For example, Citigroup has a privacy policy that is low on both transparency and control, such that when it suffered a breach, the damages were aggravated, resulting in a loss of $836 million in value in the (-1, 0) window. According to the propensity score method for counterfactual analysis (Web Appendix C), if Citigroup had high transparency and high control, it would have suffered a loss of only about $16 million in stock value. That is, *Citigroup might have saved about $820 million had it simply offered its customers high transparency and control*.

The other combinations also suggest useful takeaways for managers. When provided with high transparency but low control, customers perceive more violation and lower trust across all studies. Thus it is a dangerous practice for firms to tell customers exactly how they will be collecting data, without also providing some say over those practices. If they lack control, customers are left to worry about the various potential uses of their data—uses that have been

made salient by their transparency. Knowledge alone has mixed effects as a vulnerability suppressor. Therefore, *if firms intend to reveal their data use practices to customers, they also need to provide them with some element of control over the information*.

The combination of low data use transparency and high control instead creates a situation of uninformed autonomy. Customers have the ability to change their preferences, so they respond favorably, even if their opt-in and -out choices are somewhat blind, without full knowledge of what and how the firm uses their information. More research is needed to reveal the full effects of this strategy, perhaps by using choice theory. In critical work on understanding choice, Iyengar (2010, p. 285) notes, "If you believed you had choice, you benefitted from it, regardless of whether you actually exercised it." Collectively, these contrasts suggest that providing customers some level of control is a powerful managerial tool for generating positive firm outcomes. The amount of customer control provided might not need to reach full and total autonomy; rather, some level of *perceived control may be sufficient to obtain the desired mitigating effects*. By allowing customers to opt in or out of various data practices, firms could promote their increased overall willingness to provide personal information.

Finally, managers need to identify their competitors' data practices, the effects on their own firm's performance, and how these effects might vary with the severity of a data breach. A more severe breach by a focal firm helps rivals, through a positive competitive effect that can overwhelm the negative spillover effect. Consider Anthem's data breach in February 2015, which affected as many as 80 million customers. The high severity of this breach led its rival Aetna to gain about $745 million (2.2% returns) on the event day, due to competitive effects. In contrast, Nvidia's breach, which affected just 400,000 user accounts in July 2012, led its rival Advanced Micro Devices to lose about $48 million (–1.4% returns) on the event day, seemingly due to the spillover effect of this less severe breach.

## Limitation and Further Research

Our investigation considers what happens to customers and firms in a relatively short period surrounding data access or a data security event. To investigate how firms engage with their customers to recover from these negative events, further research might address how firms make amends or restore benevolent aspects of their customer relationships following vulnerability-inducing events, which would represent an important theoretical and practical complement to our study.

## References

Acquisti, Alessandro, Allan Friedman, and Rahul Telang (2006), "Is There a Cost to Privacy Breaches? An Event Study," *ICIS, Association for Information Systems*, 94-116.

———, Leslie K. John, and George Loewenstein (2012), "The Impact of Relative Standards on the Propensity to Disclose," *Journal of Marketing Research*, 49 (April), 160-74.

Anderson, J. Craig (2013), "Identity Theft Growing, Costly to Victims," *USA Today*, April 14, accessed October 20, 2015, http://www.usatoday.com/story/money/personalfinance/.

Awad, Naveen Farag and M. S. Krishnan (2006), "The Personalization Privacy Paradox: An Empirical Evaluation of information Transparency and the Willingness to be Profiled Online for Personalization," *MIS Quarterly*, 30 (March), 13-28.

Baker, Stacey Menzel, James W. Gentry, and Terri L. Rittenburg (2005), "Building Understanding of the Domain of Consumer Vulnerability," *Journal of Macromarketing*, 25 (December), 128-39.

Bart, Yakov, Venkatesh Shankar, Fareena Sultan, and Glen L. Urban (2005), "Are the Drivers and Role of Online Trust the Same for All Web Sites and Consumers? A Large-Scale Exploratory Empirical Study," *Journal of Marketing*, 69 (October), 133-52.

Baumeister, Roy F. and Mark R. Leary (1995), "The Need to Belong: Desire for Interpersonal Attachments as a Fundamental Human Motivation," *Psychological Bulletin*, 117, 497-529.

———, Liqing Zhang, and Kathleen D. Vohs (2004), "Gossip as Cultural Learning," *Review of General Psychology*, 8 (2), 111-21.

Beersma, Bianca and Gerben A. Van Kleef (2012), "Why People Gossip: An Empirical Analysis of Social Motives, Antecedents, and Consequences," *Journal of Applied Social Psychology*, 42 (11), 2640-70.

Biesdorf, Stefan, David Court, and Paul Willmott (2013), "Big Data: What's Your Plan?" *McKinsey Quarterly*, March (2), 40-51.

Bollen, Kenneth A., and Kwok-fai Ting (2000), "A Tetrad Test for Causal Indicators," *Psychological Methods*, 5 (1), 3-22.

Borah, Abhishek and Gerard Tellis (2014), "Make, Buy, or Ally? Choice of and Payoff from Announcements of Alternate Strategies for Innovations," *Marketing Science*, 33 (Jan/Feb), 114-33.

——— and ——— (2016), "Halo (Spillover) Effects in Social Media: Do Product Recalls of One Brand Hurt or Help Rival Brands?" *Journal of Marketing Research*, 53 (April), 143-60.

Brown, Stephen J. and Jerold B. Warner (1985), "Using Daily Stock Returns: The Case of Event Studies," *Journal of Financial Economics*, 14 (March), 3-31.

Campbell, Cynthia J., Arnold R. Cowan, and Valentina Salotti (2010), "Multi-Country Event-Study Methods," *Journal of Banking & Finance*, 34 (December), 3078-90.

Caudill, Eve M. and Patrick E. Murphy (2000), "Consumer Online Privacy: Legal and Ethical Issues," *Journal of Public Policy & Marketing*, 19 (Spring), 7-19.

Chellappa, Ramnath K. and Raymond G. Sin (2005), "Personalization versus Privacy: An Empirical Examination of the Online Consumer's Dilemma," *Information Technology and Management*, 6 () 181-202.

Cleeren, Kathleen, Harald J. van Heerde, and Marnik G. Dekimpe (2013), "Rising from the Ashes: How Brands and Categories can Overcome Product-Harm Crises," *Journal of Marketing*, 77 (March), 58-77.

Columbus, Louis (2014), "2014: The Year Big Data Adoption Goes Mainstream in the Enterprise," *Forbes*, accessed June 9, 2014, http://www.forbes.com/sites/louiscolumbus/.

Cranage, David A. (2004), "Conservative Choice, Service Failure, and Customer Loyalty: Testing the Limits of Informed Choice," *Journal of Hospitality & Tourism Research*, 28 (August), 327-45.

Cumbley, Richard and Peter Church (2013), "Is 'Big Data' Creepy?" *Computer Law & Security Review*, 29 (October), 601-9.

De Angelis, Matteo, Andrea Bonezzi, Alessandro M. Peluso, Derek D. Rucker, and Michele Costabile (2012), "On Braggarts and Gossips: A Self-Enhancement Account of Word-of-Mouth Generation and Transmission," *Journal of Marketing Research*, 49 (August), 551-63.

Dunbar, Robin I. M. (2004), "Gossip in Evolutionary Perspective," *Review of General Psychology*, 8 (2), 100-10.

Eder, Donna and Janet Lynne Enke (1991), "The Structure of Gossip: Opportunities and Constrains on Collective Expression among Adolescents," *American Sociological Review*, 56 (August), 494-508.

Emler, Nicholas (1994), "Gossip, Reputation and Social Adaption," in Robert F. Goodman and Aaron Ben-Ze'ev (eds.) *Good Gossip* (pp. 119-140). Lawrence: University of Kansas Press.

Fama, Eugene F. (1970), "Efficient Capital Markets: A Review of Theory and Empirical Work," *Journal of Finance*, 25 (2), 383-417.

——— (1998), "Market Efficiency, Long-Term Returns, and Behavioral Finance," *Journal of Financial Economics*, 49 (3), 283-306.

Feinberg, Matthew, Robb Willer, Jennifer Stellar, and Dacher Keltner (2012), "The Virtues of Gossip: Reputational Information Sharing as Prosocial Behavior," *Journal of Personality and Social Psychology*, 102 (5), 1015-30.

Fisher, John A. (2013), "Secure My Data or Pay the Price: Consumer Remedy for the Negligent Enablement of Data Breach," *William & Mary Business Law Review*, 215 (4), 217-33.

Foster, Eric K. (2004), "Research on Gossip: Taxonomy, Methods, and Future Directions," *Review of General Psychology*, 8 (2), 78-99.

Grégoire, Yany and Robert J. Fisher (2006), "The Effects of Relationship Quality on Customer Retaliation," *Marketing Letters*, 17 (January), 31-46.

———and ——— (2008), “Customer Betrayal and Retaliation: When Your Best Customers Become Your Worst Enemies,” *Journal of the Academy of Marketing Science*, 36 (Summer), 247-61.

Hausman, J. A. (1978), “Specification Tests in Econometrics,” *Econometrica*, 46 (November), 1251-71.

Homburg, Christian, Josef Vollmayr, and Alexander Hahn (2014), “Firm Value Creation through Major Channel Expansions: Evidence from an Event Study in the United States, Germany, and China,” *Journal of Marketing*, 78 (May), 38-61.

Hsieh, Tien-Shih, Daniel Noyes, Hong Liu, and Lance Fiondella (2015), “Quantifying Impact of Data Loss Incidents on Publicly-Traded Organizations,” *Proceedings of IEEE*, 2-15.

Hulland, John (1999), “Use of Partial Least Squares (PLS) in Strategic Management Research: A Review of Four Recent Studies,” *Strategic Management Journal*, 20 (February), 195-204.

Imbens, Guido W. (2000), “The Role of the Propensity Score in Estimating Dose Response Functions,” *Biometrika*, 87 (September), 706–10.

Iyengar, Sheena (2010), *The Art of Choosing*. New York: Twelve.

John, Leslie K., Alessandro Acquisti, and George Loewenstein (2011), “Strangers on a Plane: Context-Dependent Willingness to Divulge Sensitive Information,” *Journal of Consumer Research*, 37 (February), 858-73.

Ko, Myung and Carlos Dorantes (2006), “The Impact of Information Security Breaches on Financial Performance of the Breached Firms: An Empirical Investigation,” *Journal of Information Technology Management*, 17 (2), 13-22.

Kumar, V., Xi Zhang, and Anita Luo (2014), “Modeling Customer Opt-In and Opt-Out in a Permission-Based Marketing Context,” *Journal of Marketing Research*, 51 (August), 403-19.

Leary, Mark R. and Sadie Leder (2009), “The Nature of Hurt Feelings: Emotional Experience and Cognitive Appraisals,” in *Feeling Hurt in Close Relationships*, Anita L. Vangelisti, ed. Cambridge University Press: New York, 15-33.

Lee, Peggy M. (2001), “What's in a Name.com?: The Effects of ‘.com’ Name Changes on Stock Prices and Trading Activity,” *Strategic Management Journal*, 22 (August), 793-804.

Lwin, May, Jochen Wirtz, and Jerome D. Williams (2007), “Consumer Online Privacy Concerns and Responses: A Power-Responsibility Equilibrium Perspective,” *Journal of the Academy of Marketing Science*, 35 (Winter), 573-85.

Macobserver (2014), http://www.macobserver.com/tmo/article/unpatched-comcast-security-flaw-leaves-user-data-exposed.

Malhotra, Arvind and Claudia Kubowicz Malhotra (2011), “Evaluating Customer Information Breaches as Service Failures: An Event Study Approach,” *Journal of Service Research*, 14 (1), 44–59.

Malhotra, Naresh K., Sung S. Kim, and James Agarwal (2004), "Internet Users' Information Privacy Concerns (IUIPC): The Construct, the Scale, and a Causal Model," *Information Systems Research*, 15 (December), 336-55.

Marcus, Gary and Ernest Davis (2014), "Eight (No, Nine!) Problems with Big Data," *New York Times*, accessed April 21, 2014, www.nytimes.com/2014/04/01/opinion.

Martin, Kelly D. and Patrick E. Murphy (2016), "Customer Privacy in a Data-Rich Environment," *Journal of the Academy of Marketing Science*, working paper.

McAfee, Andrew and Erik Brynjolfsson (2012), "Big Data: The Management Revolution," *Harvard Business Review*, (October), 3-9.

McWilliams, Abagail and Donald Siegel (1997), "Event Studies in Management Research: Theoretical and Empirical Issues," *Academy of Management Journal*, 40 (June), 626-57.

Mills, Colleen (2010), "Experiencing Gossip: The Foundations for a Theory of Embedded Organizational Gossip," *Group & Organization Management*, 35 (4), 371-90.

Milne, George R., Andrew J. Rohm, and Shalini Bahl (2004), "Consumers' Protection of Online Privacy and Identity," *Journal of Consumer Affairs*, 38 (Winter), 217-32.

Modi, Sachin B. and Sanjay Mishra (2011), "What Drives Financial Performance-Resource Efficiency or Resource Slack?: Evidence from U.S. Based Manufacturing Firms from 1991 to 2006," *Journal of Operations Management*, 29 (March), 254-73.

Moon, Youngme (2000), Intimate Exchanges: Using Computers to Elicit Self-Disclosure from Consumers," *Journal of Consumer Research*, 26 (March), 323-39.

Mothersbaugh, David L., William K. Foxx II, Sharon E. Beatty, and Sijun Wang (2012), "Disclosure Antecedents in an Online Service Context: The Role of Sensitivity of Information," *Journal of Service Research*, 15 (1), 76-98.

O'Gorman, Rick, Kennon M. Sheldon, and David S. Wilson (2008), "For the Good of the Group? Exploring Group-Level Evolutionary Adaptations Using Multilevel Selection Theory," *Group Dynamics: Theory, Research, and Practice*, 12 (1), 17-26.

Palmatier, Robert W. (2008), "Interfirm Relational Drivers of Customer Value," *Journal of Marketing*, 72 (July), 76-89.

———, Rajiv P. Dant, Dhruv Grewal, and Kenneth R. Evans (2006), "Factors Influencing the Effectiveness of Relationship Marketing: A Meta-Analysis," *Journal of Marketing*, 70 (October), 136-53.

———, Lisa K. Scheer, and Jan-Benedict E. M. Steenkamp, (2007), "Customer Loyalty to Whom? Managing the Benefits and Risks of Salesperson-Owned Loyalty," *Journal of Marketing Research*, 44 (May), 185-199.

Ponemon Institute (2015), *2015 Cost of Data Breach Study: Global Analysis*, IBM and Ponemon Institute, LLC.

Preacher, Kristopher J. and Andrew F. Hayes (2008), "Asymptotic and Resampling Strategies for Assessing and Comparing Indirect Effects in Multiple Mediator Models," *Behavior Research Methods*, 40 (3), 879-91.

Raassens, Néomie, Stefan Wuyts, and Inge Geyskens (2012), "The Market Valuation of Outsourcing New Product Development," *Journal of Marketing Research*, 49 (October), 682-95.

Richman, Laura Smart and Mark R. Leary (2009), "Reactions to Discrimination, Stigmatization, Ostracism, and Other Forms of Interpersonal Rejection: A Multimotive Model," *Psychological Review*, 116 (2), 365-83.

Roehm, Michelle L. and Alice M. Tybout (2006), "When Will a Brand Scandal Spill Over, and How Should Competitors Respond?" *Journal of Marketing Research*, 43 (August), 366-73.

Romanosky, Sasha, Rahul Telang, and Alessandro Acquisti (2011), "Do Data Breach Disclosure Laws Reduce Identity Theft?" *Journal of Policy Analysis & Management*, 30 (Spring), 256-86.

Rosenbaum, Paul R., and Donald B. Rubin (1983), "The Central Role of the Propensity Score in Observational Studies for Causal Effects," *Biometrika*, 70 (1), 41-55.

Scharf, S. (2007), "Report Casts Doubt on the Impact of Data Breaches on Identity Theft," *Internal Auditor*, 64 (August), 23.

Schatz, Daniel and Rabih Bashroush (2016), "The Impact of Repeated Data Breach Events on Organisations' Market Value," *Information and Computer Security*, 24 (1), 73-92.

Schlosser, Ann E., Tiffany Barnett White, and Susan M. Lloyd (2006), "Converting Web Site Visitors into Buyers: How Web Site Investment Increases Consumer Trusting Beliefs and Online Purchase Intentions," *Journal of Marketing*, 70 (April), 133-48.

Schumann, Jan H., Florian von Wangenheim, and Nicole Groene (2014), "Targeted Online Advertising: Using Reciprocity Appeals to Increase Acceptance among Users of Free Web Services," *Journal of Marketing*, 78 (January), 59-75.

Sen, Ravi and Sharad Borle (2015), "Estimating the Contextual Risk of Data Breach: An Empirical Approach," *Journal of Management Information Systems*, 32 (2), 314-41.

Sharpe, William F. (1964), "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk," *Journal of Finance*, 19 (3), 425-42.

Smith, Eliot R. (2014), "Evil Acts and Malicious Gossip: A Multiagent Model of the Effects of Gossip in Socially Distributed Person Perception," *Personality and Social Psychology Review*, 18 (4), 311-25.

Smith, N. Craig and Elizabeth Cooper-Martin (1997), "Ethics and Target Marketing: The Role of Product Harm and Consumer Vulnerability," *Journal of Marketing*, 61 (July), 1-20.

Solove, Daniel J. (2003), "Identity Theft, Privacy, and the Architecture of Vulnerability," *Hastings Law Journal*, 54 (1227), 1-47.

——— (2006), "A Taxonomy of Privacy," *University of Pennsylvania Law Review,* 154 (January), 477-564.

Srinivasan, Raji, Gary L. Lilien, and Shrihari Sridhar (2011), "Should Firms Spend More on Research and Development and Advertising During Recessions?" *Journal of Marketing,* 75 (May), 49-65.

Srinivasan, Shuba and Dominique M. Hanssens (2009), "Marketing and Firm Value: Metrics, Methods, Findings, and Future Directions," *Journal of Marketing Research*, 46 (June), 293-312.

Steel, Emily and Geoffrey A. Fowler (2010), "Facebook in Privacy Breach," *Wall Street Journal*, 256 (92), A1-A2.

Stuart, Elizabeth A., (2010), "Matching Methods for Causal Inference: A Review and a Look Forward," *Statistical Science,* 25, (1), 1-21.

Trefis Team (2014), "Has Optimism Been Guiding Home Depot and Lowe's to Glory?" *Forbes.com*, July 10, accessed October 20, 2015, www.forbes.com/sites/.

Tucker, Catherine (2014), "Social Networks, Personalized Advertising and Privacy Controls," *Journal of Marketing Research*, 51 (October), 546-62.

Turner, Monique Mitchell, Michelle A. Mazur, Nicole Wendel, and Robert Winslow (2003), "Relational Ruin or Social Glue? The Joint Effect of Relationship Type and Gossip Valence on Liking, Trust, and Expertise," *Communication Monographs*, 70 (June), 129-41.

White, Tiffany Barnett (2004), "Consumer Disclosure and Disclosure Avoidance: A Motivational Framework," *Journal of Consumer Psychology*, 14 (1 & 2), 41-51.

Williams, Kipling D. (2007), "Ostracism," *Annual Review of Psychology*, 58, 425-52.

**TABLE 1**

**Customer Data Vulnerability: Selected Relevant Literature**

| Study | Areas of Focus | Context | Key Findings |
|---|---|---|---|
| *Data Access Vulnerability* | | | |
| Bart et al. (2005) | Online trust, privacy, security, website presentation, brand strength | Model estimation with data from 6831 customers | Navigation and presentation, advice, and brand strength are more influential predictors of online trust than are privacy and security. Online trust mediates the relationship between website characteristics and behavioral intentions more strongly for some product categories than others. |
| Schlosser, White, and Lloyd (2006) | Trusting beliefs (ability, benevolence, integrity), website investment, privacy/security | Website manipulation experiments | Website investment/design is the strongest factor leading to purchase intentions and trust. Privacy and security statements increase benevolence and integrity dimensions of trust but do not increase searchers' willingness to buy online. |
| John, Acquisti, and Loewenstein (2011) | Environmental cues, privacy concerns, willingness to divulge highly sensitive information | Online experiments | Contextual information encourages more or less customer information disclosure, including both intrusiveness and the professional look of the response format. Priming with a privacy statement decreases disclosure. |
| Acquisti, John, and Loewenstein (2012) | Conformity, reciprocity, injunctive and descriptive norms, herding effect | Online experiments | Customers are willing to disclose increasingly sensitive information when they believe others have done so. Respondents disclose sensitive information more freely when placed at the beginning of a questionnaire (cf. random or end placement). |
| Schumann, Wangenheim, and Groene (2014) | Social norms and reciprocity, privacy concerns, advertising effectiveness, user-generated content | Field studies, online experiments | Customers increasingly accept targeted advertising in exchange for a website's free services. Customers report targeted advertising as an alternative form of online currency to voluntarily repay a website for customization and other marketing benefits. |
| Tucker (2014) | Customer privacy controls, targeted ads, personalized ads, reactance theory | Facebook campaign-level click-through data | For a nonprofit using personalized (vs. nonpersonalized) and targeted (vs. nontargeted) ads on Facebook, people responded more favorably to personalized ads when they had the ability to control their personal privacy settings. |
| *Data Breach Vulnerability* | | | |
| Schatz and Bashroush (2016) | Stock market value, data breach (single and repeated) | Event study with 25 publicly-traded U.S. companies | Preliminary evidence suggests that data breaches are bad for performance. The effect worsens when a firm has experienced more than one breach. |
| Hsieh et al. (2015) | Stock market value, data loss events, firm size, data loss costs | Event study of 103 U.S. public firm data breaches | Data loss events negatively affect firm performance. Companies should invest more in data security efforts. |
| Sen and Borle (2015) | Data breach risk, firm location, industry, type of past breach | Data breaches between 2005-2012 | State-level data breach disclosure laws can influence breach risk in certain industries. Because greater security spending heightens breach risk, IT dollars may be suboptimally allocated. |
| Malhotra and Malhotra (2011) | Stock market effect of firm data, breach of customer data, severity | Event study of 93 publicly traded firm data breaches | Firm market value is negatively affected by a breach in both the short and long terms, but it is more detrimental in the long term. Larger firms suffer greater market value loss than smaller firms, and larger firms suffer more from large breaches. No effect of severity on smaller firms. |
| Acquisti, Friedman, and Telang (2006) | Firm performance, data breach scale, scope, type (e.g., employees, customers, third party), information type, industry | Event study of 79 publicly traded firm data breaches | A data breach has a significant negative effect on stock market value the day that breach is announced. The cumulative effect increases the day of the announcement but then decreases and loses statistical significance over time. |
| Ko and Dorantes (2006) | Firm performance, data breach | Matched-sample comparison methodology with 19 data breaches | Focal firm's performance decreased relative to unaffected peer firms (examined as a control group). This study finds both short-term and long-term negative effects of a data breach on performance and identifies a fourth-quarter recovery effect. |
| *Data Manifest Vulnerability* | | | |
| Romanosky, Telang, and Acquisti (2011) | Consumer identity theft, data breach disclosure laws | Victim identity theft reports via U.S. FTC spanning 2002-2009 | Research asks whether data breach disclosure laws actually reduce identity theft. They find that on average, statutes reduce identity theft caused by breaches by six percent in evidence of the effectiveness of a particular privacy protecting law. |
| Milne, Rohm, and Bahl (2004) | Consumer identity theft, online shopping behavior, privacy attitudes, offline data protection practices | Consumer surveys | Findings across three surveys suggest consumers are not sufficiently protecting themselves from identity theft. Authors advocate for greater firm and government protection, given consumers' reported lack of understanding about adequate ways to protect themselves from harm online. |

**TABLE 2**
**Constructs, Definitions, and Operationalizations**

| Constructs | Definitions | Studies 1 and 3 | Study 2 |
|---|---|---|---|
| Data access vulnerability | Customer expectation of susceptibility to the harm that can come from the disclosure of their personal data. | Experimentally manipulated extent (high/low) to which company has access to personal, sensitive, or private customer information. | N/A |
| Spillover vulnerability | Extent to which customer feels vulnerable due to the data breach of a firm that is a close rival of a firm they use. | Experimentally manipulated event in which customers learn a close competitor of a company they use is the victim of a data breach. | Closest competitor firm data breach event (yes/no). |
| Data breach vulnerability | Extent to which customer feels vulnerable due to a firm's security lapse, making data vulnerability salient. | Experimentally manipulated event in which customers learn a company they use has been the victim of a data breach. | Corporate data breach event (yes/no). |
| Data manifest vulnerability | Extent to which customer feels vulnerable due to actual misuse of personal information, making data vulnerability salient. Can occur through fraudulent activity including, but not limited to, identity theft. | Experimentally manipulated event in which customers learn a company they use has been the victim of a data breach and that their information has been used fraudulently, in the form of identity theft. | N/A |
| Data breach severity | The scope, reach, and impact of a firm's data security breach. | N/A | Log of number of customer records compromised in data breach. |
| Data use transparency | Customer knowledge of a firm's access to her or his data and understanding of how it is going to be used (Awad and Krishnan 2006) | Experimentally manipulated extent (high/low) to which a company's data management policies are clear, straightforward, and easy to understand. | Count of whether various elements explained: opt-out policy, data capture, data use, data sharing with third parties, contact information available for privacy requests. |
| Customer control | Customer perception of the extent to which she or he can manage a firm's use of her or his personal data (Tucker 2014). | Experimentally manipulated extent (high/low) to which customer has control over firm's use of data. | Count of number of opt-out choices as detailed in the firm's privacy policy. |
| Emotional violation | Customer's perception of a firm's failure to respect peace, privacy, or other rights (Grégoire and Fisher 2008). | Extent to which customer feels violated or betrayed by firm's use of data. | N/A |
| Cognitive trust | Willingness to rely on an exchange partner in whom one has confidence (Palmatier 2008). | Extent to which a customer reports trusting a firm and its behaviors. | N/A |
| Financial performance | Focal firm (rival firm) financial performance. | N/A | Firm's abnormal stock market returns calculated by the market model. |
| Falsifying behavior | Customer fabrication of their personal information when transacting with a company (Lwin, Wirtz, and Williams 2007). | Customer reported likelihood of providing inaccurate personal information to a company with which they interact. | N/A |
| Negative word of mouth | Customer negative communications to others about a company (De Angelis et al. 2012). | Customer reported likelihood of spreading negative word of mouth about the firm to friends, family. | N/A |
| Switching behavior | Customer likelihood of discontinuing the relationship in favor of a similar alternative (Palmatier, Scheer, and Steenkamp 2007) with reduced data access vulnerability. | Customer reported likelihood of switching to a comparable firm, including trying its products/services and paying a premium to switch. | N/A |

**TABLE 3**
**Descriptive Statistics and Correlations**

***Panel A: Study 2***

| Variables | Mean | SD | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. Firm Performance | .00 | .02 | 1.00 | | | | | | | | | | | |
| 2. Data Breach Vulnerability | .49 | .50 | -.03 | 1.00 | | | | | | | | | | |
| 3. Data Use Transparency | 3.71 | 1.46 | -.01 | .05 | 1.00 | | | | | | | | | |
| 4. Customer Control | 1.03 | .91 | .05 | -.01 | .51 | 1.00 | | | | | | | | |
| 5. Data Breach Severity | 10.30 | 4.01 | -.01 | -.02 | -.05 | -.03 | 1.00 | | | | | | | |
| 6. Capital Resource Slack | .24 | .76 | .00 | -.05 | .11 | -.01 | .10 | 1.00 | | | | | | |
| 7. Firm Size | 10.49 | 1.63 | -.01 | -.03 | .08 | .00 | -.05 | .05 | 1.00 | | | | | |
| 8. B2C vs. B2B | .53 | .50 | -.01 | .03 | .06 | .10 | -.09 | .06 | .39 | 1.00 | | | | |
| 9. Services vs. Goods | .69 | .46 | .02 | .01 | -.15 | -.05 | .06 | .00 | -.22 | -.15 | 1.00 | | | |
| 10. Competitive Intensity | .01 | 1.02 | .01 | -.01 | .00 | .09 | -.22 | -.08 | .18 | .02 | -.26 | 1.00 | | |
| 11. Firm Prior Number of Breaches | .40 | 1.07 | .02 | .38 | .06 | .05 | .05 | .03 | .15 | .00 | .12 | .01 | 1.00 | |
| 12. Rival Prior Number of Breaches | .48 | 1.31 | -.01 | -.36 | -.06 | -.06 | .07 | .08 | .24 | .10 | .13 | -.01 | -.14 | 1.00 |

*Note*: Correlations of .08 or greater are significant at $p < .05$.

***Panel B: Study 3***

| Variables | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. Change in Vulnerability | 1.00 | | | | | | | | | | | | | | | | |
| 2. Data Use Transparency (Privacy Policy | -.01 | 1.00 | | | | | | | | | | | | | | | |
| 3. Customer Control (Privacy Policy) | -.16 | .03 | 1.00 | | | | | | | | | | | | | | |
| 4. Data Use Transparency | -.15 | .53 | .11 | 1.00 | | | | | | | | | | | | | |
| 5. Customer Control | -.18 | .12 | .72 | .57 | 1.00 | | | | | | | | | | | | |
| 6. ΔEmotional Violation | .55 | -.09 | -.23 | -.01 | -.03 | 1.00 | | | | | | | | | | | |
| 7. ΔCognitive Trust | -.62 | .01 | .12 | .07 | .03 | -.51 | 1.00 | | | | | | | | | | |
| 8. ΔFalsifying Information | .38 | -.07 | -.05 | -.06 | -.05 | .52 | -.54 | 1.00 | | | | | | | | | |
| 9. ΔNegative Word of Mouth | .48 | -.01 | -.01 | -.07 | -.12 | .57 | -.61 | .60 | 1.00 | | | | | | | | |
| 10. ΔSwitching Behavior (%) | .43 | -.03 | -.03 | -.08 | -.06 | .35 | -.47 | .33 | .46 | 1.00 | | | | | | | |
| 11. Value | -.15 | .19 | .12 | .50 | .58 | -.09 | .05 | -.05 | -.01 | -.06 | 1.00 | | | | | | |
| 12. Fairness | -.27 | .14 | .05 | .51 | .56 | -.14 | .10 | -.08 | -.08 | -.18 | .59 | 1.00 | | | | | |
| 13. Privacy Concern | .12 | .09 | .01 | .02 | .02 | .12 | -.18 | .02 | .06 | .08 | -.02 | -.06 | 1.00 | | | | |
| 14. Data Vulnerability Event Experience | -.15 | -.07 | -.07 | -.09 | -.06 | -.14 | .10 | -.15 | -.02 | .00 | .04 | -.09 | .16 | 1.00 | | | |
| 15. Age (categorical) | -.05 | .09 | .06 | -.20 | -.28 | -.02 | .04 | -.05 | .03 | .04 | -.31 | -.13 | .01 | .04 | 1.00 | | |
| 16. Gender (1 = male; 0 = female) | -.06 | .01 | .03 | .00 | .00 | -.07 | .18 | -.06 | -.03 | -.01 | .05 | .06 | -.12 | .05 | -.07 | 1.00 | |
| 17. Population Size (categorical) | .02 | .03 | .04 | -.04 | -.01 | .02 | .01 | -.10 | -.02 | .06 | -.02 | -.06 | -.03 | -.07 | -.14 | -.16 | 1.00 |
| **Mean** | 1.29 | 3.80 | 2.13 | 4.62 | 3.84 | 1.01 | -1.27 | .51 | .85 | 4.28 | 4.37 | 5.02 | 4.53 | 3.38 | 3.48 | .58 | 3.01 |
| **SD** | 1.62 | 1.08 | 1.13 | 1.56 | 1.51 | 1.59 | 1.51 | 1.43 | 1.45 | 17.21 | 1.40 | 1.21 | 1.43 | 1.56 | .95 | .51 | 1.33 |

*Note*: The computation of the privacy policy correlations attached company-level values to each individual case. Correlations of .15 or greater are significant at $p < .05$

**TABLE 4**

**Results: Effect of Data Breach Events on Abnormal Stock Returns for Focal and Rival Firms (Study 2)**

| Windows | Returns | T-stat | Sig. Level | Wilcoxon | Sig. Level |
|---|---|---|---|---|---|
| ***Panel A: Abnormal Returns to a Breach Event for Both Focal and the Nearest Rival of Focal Firms*** | | | | | |
| (0,0) | -.14% | -2.61 | 0.01 | -2.77 | 0.01 |
| (-1,0) | -.23% | -2.90 | 0.00 | -3.30 | 0.00 |
| (0,1) | -.05% | -0.70 | 0.49 | -1.63 | 0.10 |
| (-1,1) | -.14% | -1.45 | 0.15 | -2.23 | 0.03 |
| ***Panel B: Abnormal Returns to a Breach Event for Focal Firms*** | | | | | |
| (0,0) | -.15% | -1.72 | 0.09 | -2.02 | 0.04 |
| (-1,0) | -.29% | -2.38 | 0.02 | -2.87 | 0.00 |
| (0,1) | -.13% | -1.31 | 0.19 | -1.96 | 0.05 |
| (-1,1) | -.27% | -2.01 | 0.05 | -2.48 | 0.01 |
| ***Panel B: Abnormal Returns to a Breach Event for the Nearest Rival of Focal Firms*** | | | | | |
| (0,0) | -.14% | -2.00 | 0.05 | -1.90 | 0.06 |
| (-1,0) | -.17% | -1.69 | 0.09 | -1.76 | 0.08 |
| (0,1) | .03% | 0.29 | 0.77 | -0.42 | 0.68 |
| (-1,1) | -.01% | -0.04 | 0.97 | -0.72 | 0.47 |

**TABLE 5**

**Results: Data Breach Vulnerability, Suppressors, and Severity on Abnormal Stock Returns (Study 2)**

| Variables | Hypotheses | Model 1 | | Model 2 | |
|---|---|---|---|---|---|
| | | Coefficients | Std. Coef. | Coefficients | Std. Coef. |
| ***Main Effects*** | | | | | |
| Breach (1 = Focal; 0 = Rival) | | -.0019 (.0014) | -.0830 | .0096** (.0038) | .0100 ** |
| Transparency | $\mathbf{H_4}$ | -.0009 (.0009) | -.0544 | -.0018* (.0010) | -0.113 * |
| Control | $\mathbf{H_5}$ | .0017* (.0009) | .0660 * | -.0029 (.0024) | -.1126 |
| ***Interaction Effects*** | | | | | |
| Transparency * Control | $\mathbf{H_6}$ | | | .0012* (.0006) | 0.225 * |
| Severity of Breach * Breach | $\mathbf{H_{7a}}$ | | | -.0006* (.0003) | -0.163 * |
| Severity of Breach * (1 - Breach) | $\mathbf{H_{7b/7b(alt)}}$ | | | .0005* (.0002) | 0.124 * |
| ***Controls*** | | | | | |
| Data Breach Severity | | .0000 (.0002) | .0016 | | |
| Capital Resource Slack | | .0001 (.0007) | .0042 | -.0001 (.0007) | -.0043 |
| Firm Size | | -.0001 (.0004) | -.0068 | -.0001 (.0004) | -.0031 |
| B2C vs. B2B | | .0004 (.0014) | .0152 | .0000 (.0014) | -.0004 |
| Services vs. Goods | | .0012 (.0016) | .0522 | .0012 (.0016) | .0515 |
| Competitive Intensity | | .0010 (.0050) | .0060 | -.0005 (.0052) | -.0031 |
| Prior Number of Breaches for Focal Firm | | .0008 (.0008) | .0368 | .0009 (.0008) | .0446 |
| Prior Number of Breaches for Rival Firm | | -.0001(.0006) | -.0056 | -.0001 (.0006) | -.0040 |
| Time Trend | | .0000 (.0000) | .0003 | .0000 (.0000) | .0003 |
| Industry Fixed Effect Included[a] | | Yes | | Yes | |
| Year Fixed Effect Included[b] | | Yes | | Yes | |
| | | | | -.0296* (.0143) | |
| Fit Statistics | | Overall R-Square: 0.03 | | Overall R-Square: 0.045 | |
| Wald Chi Square | | 70.93 | | 91.3 | |
| N | | 583 | | 583 | |

*Note*: Standard errors are reported in parentheses; *$p$ < .05, **$p$ < .01, ***$p$ < .001 (one-tailed hypothesis tests).

[a] Industry fixed effects include finance, retail, online, high-tech, food/health, and services (manufacturing as reference category).

[b] Year fixed effects for 2004 (reference category) through 2015.

**TABLE 6**
**HLM Results: Change in Vulnerability and Suppressors (Study 3)**

| Variables | Model 1 | Model 2 | Model 3 | Model 4 |
|---|---|---|---|---|
| | β (s.e.) | β (s.e.) | β (s.e.) | β (s.e.) |
| ***Industry-Level (Level 3) Effects*** | | | | |
| Retail | .01 (.19) | .03 (.17) | .02 (.16) | .04 (.16) |
| Technology | .05 (.20) | .06 (.22) | .03 (.19) | .06 (.21) |
| ***Company-Level (Level 2) Effects*** | | | | |
| Transparency | -.01 (.04) | -.05 (.05) | -.01 (.03) | -.04 (.05) |
| Control | -.11 (.06) * | -.16 (.07) * | -.11 (.05) * | -.16 (.07) * |
| Transparency * Control | | -.09 (.07) * | | -.09 (.07) * |
| ***Customer-Level (Level 1) Effects*** | | | | |
| *Data Vulnerability Event* | | | | |
| Data Transparency | | | -.11 (.07) * | -.14 (.07) * |
| Customer Control | | | -.12 (.07) * | -.10 (.07) * |
| Transparency * Control | | | | -.09 (.06) * |
| Data Breach Event | .70 (.17) *** | .69 (.17) *** | .71 (.17) *** | .69 (.17) *** |
| Identity Theft Event | .52 (.17) *** | .51 (.17) *** | .53 (.16) *** | .51 (.16) *** |
| Spillover Event | -.54 (.17) *** | -.54 (.17) *** | -.56 (.17) *** | -.57 (.17) *** |
| *Customer-Level Controls* | | | | |
| Privacy Concern | .11 (.06) | .12 (.06) | .11 (.06) | .10 (.06) |
| Prior Event Experience | -.13 (.06) * | -.14 (.06) * | -.11 (.06) * | -.12 (.06) * |
| Age | -.09 (.06) | -.09 (.06) | -.03 (.06) | -.03 (.06) |
| Gender | -.05 (.12) | -.04 (.12) | -.06 (.12) | -.04 (.12) |
| Population Size | -.17 (.13) | -.17 (.18) | -.14 (.13) | -.15 (.13) |
| Level 3 $R^2$ | .06 | .06 | .06 | .06 |
| Level 2 $R^2$ | .28 | .31 | .28 | .30 |
| Level 1 $R^2$ | .57 | .58 | .60 | .61 |

*$p$ < .05, **$p$ < .01, ***$p$ < .001.

**TABLE 7**

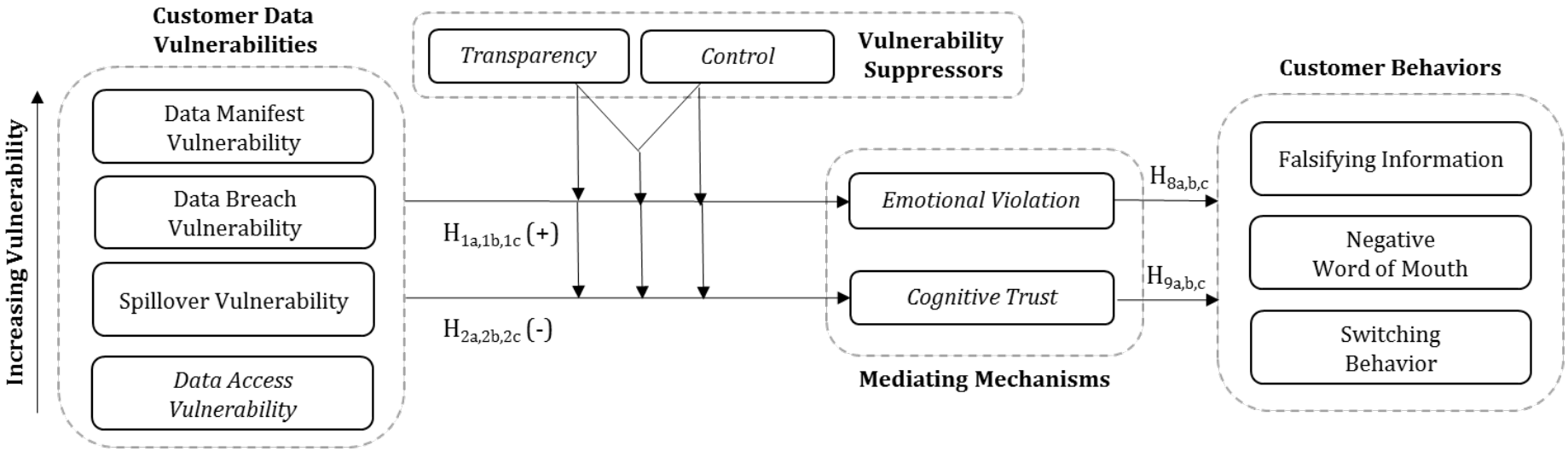
**PLS Results: Change in Vulnerability on Customer Outcomes (Study 3)**

| Structural Paths | Hypotheses | Model 1 β (SE) | Model 2 β (SE) |
|---|---|---|---|
| ***PROCESS Test of Indirect Effects*** | | | |
| Δ Vulnerability → Δ Violation → Δ Falsifying | $\mathbf{H_{8a}}$ | .17 (.04) *** | (CI: .10, .26) |
| Δ Vulnerability → Δ Violation → Δ Negative WOM | $\mathbf{H_{8b}}$ | .07 (.04) *** | (CI: .10, .27) |
| Δ Vulnerability → Δ Violation → Δ Switching | $\mathbf{H_{8c}}$ | .53 (.05) *** | (CI: .42, .65) |
| Δ Vulnerability → Δ Trust → Δ Falsifying | $\mathbf{H_{9a}}$ | .21 (.06) *** | (CI: .11, .34) |
| Δ Vulnerability → Δ Trust → Δ Negative WOM | $\mathbf{H_{9b}}$ | .22 (.05) *** | (CI: .12, .33) |
| Δ Vulnerability → Δ Trust → Δ Switching | $\mathbf{H_{9c}}$ | 1.95 (.50) *** | (CI: 1.09, 3.01) |
| ***Effects on Mediating Mechanisms*** | | | |
| Δ Vulnerability → Δ Emotional Violation | | .55 (.10) *** | .56 (.11) *** |
| Δ Vulnerability → Δ Cognitive Trust | | -.63 (.08) *** | -.58 (.11) *** |
| ***Effects of Mediating Mechanisms on Performance*** | | | |
| Δ Emotional Violation → Δ Falsifying Behavior | | .34 (.12) ** | .33 (.12) ** |
| Δ Cognitive Trust → Δ Falsifying Behavior | | -.36 (.12) ** | -.36 (.12) ** |
| Δ Emotional Violation → Δ Negative WOM | | .38 (.10) *** | .37 (.10) *** |
| Δ Cognitive Trust → Δ Negative WOM | | -.40 (.11) *** | -.41 (.12) *** |
| Δ Emotional Violation → Δ Switching Behavior | | .16 (.13) * | .16 (.13) * |
| Δ Cognitive Trust → Δ Switching Behavior | | -.38 (.12) *** | -.39 (.12) *** |
| ***Controls on Mediating Mechanisms*** | | | |
| Δ Vulnerability x Value → Δ Violation | | | .06 (.20) |
| Δ Vulnerability x Fairness → Δ Violation | | | -.08 (.19) |
| Δ Vulnerability x Value → Δ Trust | | | -.44 (.08) ** |
| Δ Vulnerability x Fairness → Δ Trust | | | -.70 (.04) *** |
| $\boldsymbol{R^2}$ | | | |
| Δ Emotional Violation | | .30 | .30 |
| Δ Cognitive Trust | | .38 | .41 |
| Δ Falsifying | | .37 | .39 |
| Δ Negative WOM | | .46 | .46 |
| Δ Switching Behavior | | .24 | .24 |

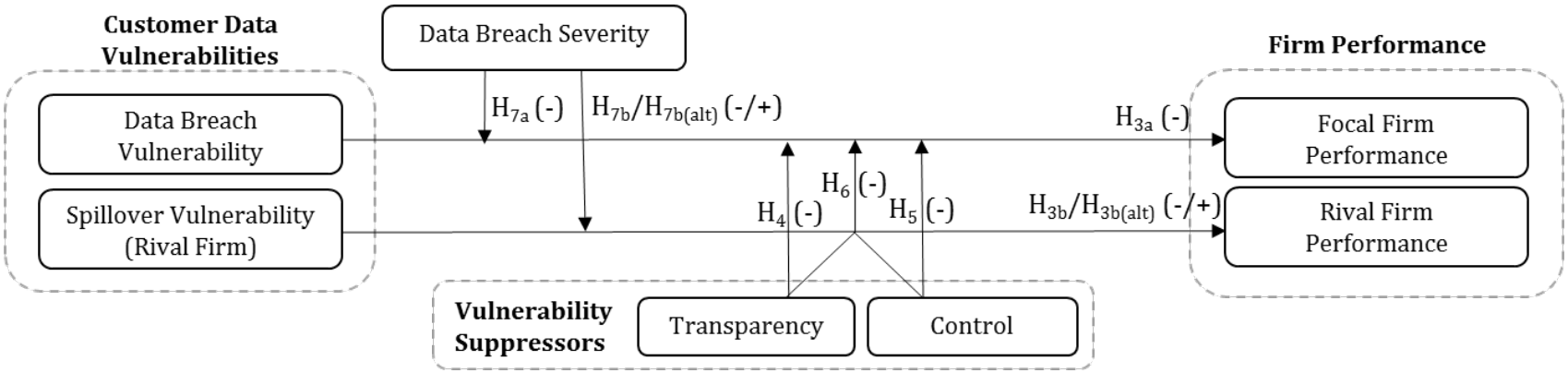
*$p < .05$, **$p < .01$, ***$p < .001$.

**FIGURE 1**
**Conceptual Framework: Studies 1, 2, and 3**

***Panel A: Effect of Customer Data Vulnerabilities and Mediators on Customer Behaviors (Studies 1 and 3)***




***Panel B: Effect of Customer Data Breach and Spillover Vulnerabilities on Firm Performance (Study 2)***




**Note**: Study 1 test the constructs in italics. Study 3 tests the full model in Panel A. Although we predict suppressing effects of transparency, control, and their interaction on all types of customer data vulnerability, the samples, measures, and study contexts differ across the three studies.

**APPENDIX A**

**Study 1 Manipulation Checks and Experiment Scenarios**

| Study | Mean (high) | Mean (low) | Sum of Squares | F-test | *p* |
|---|---|---|---|---|---|
| ***Study 1a: Vulnerability x Transparency on Emotional Violation*** | | | | | |
| Vulnerability | 5.50 | 3.31 | 240.18 | (1, 198) = 99.67 | .000 |
| Transparency | 5.85 | 2.50 | 135.16 | (1, 198) = 75.15 | .000 |
| ***Study 1b: Vulnerability x Control on Emotional Violation*** | | | | | |
| Vulnerability | 5.74 | 3.74 | 198.19 | (1, 198) = 90.41 | .000 |
| Control | 4.77 | 2.77 | 201.00 | (1, 198) = 93.50 | .000 |
| ***Study 1c: Transparency x Control on Violaton and Trust*** | | | | | |
| Transparency | 5.47 | 2.64 | 399.29 | (1, 197) = 151.00 | .000 |
| Control | 4.73 | 2.72 | 200.56 | (1, 197) = 76.20 | .000 |

**Experiment Scenarios**

***Introduction***

The following scenario asks you to imagine a company you often deal with to buy products and services. You shop with this retailer an average of once a week, both in the store and online. You make a large number of purchases with them.

***Vulnerability***

**High:** This firm has access to all your personal information, including your financial information and background, your credit card numbers, and your detailed purchase history.

**Low:** This firm has access to only limited personal information, including basic demographics and recent purchase information. They do not store credit card numbers or other financial information, and do not keep your detailed purchase history.

***Transparency***

**High:** This company is very transparent in how they manage your personal information. For example, their data management activities are clear to you, and their policies are easy to understand.

**Low:** This company is very vague in how they manage your personal information. For example, their data management activities are unclear to you and their policies are difficult to understand.

***Control***

**High**: This company gives you great control in how they manage your personal information. For example, you may change at any time your personal settings that dictate how your information is used.

**Low:** This company does not give you any control in how they manage your personal information. For example, you do not have the ability to choose the ways in which your personal information is used.

**APPENDIX B**
**Studies 1 and 3 Measures and Measurement Properties**

| Construct Items | Pre-Event/Post-Event Loading | CR | AVE |
|---|---|---|---|
| ***Study Variables*** | | | |
| Vulnerability *(new scale)* | | | |
| *The personal information that the company has about me makes me feel:* | | .97/.99 | .84/.91 |
| Insecure. | .94/.97 | | |
| Exposed. | .92/.97 | | |
| Threatened. | .93/.97 | | |
| Vulnerable. | .91/.95 | | |
| Susceptible. | .94/.96 | | |
| Data Use Transparency *(new scale)* | | .98 | .91 |
| *The company's customer data management activities are:* | | | |
| Unclear to me/Clear to me | .95 | | |
| Confusing/Straightforward | .96 | | |
| Difficult to understand/Easy to understand | .96 | | |
| Vague/Transparent | .95 | | |
| Customer Control *(adapted from Mothersbaugh et al. 2012)* | | .96 | .87 |
| I believe I have control over what happens to my personal information. | .91 | | |
| It is up to me how much the company uses my information. | .94 | | |
| I have a say in how my information is used by the company. | .95 | | |
| I have a say in whether my personal information is shared with others. | .94 | | |
| Emotional Violation *(adapted from Gregoire and Fisher 2008)* | | | |
| *Regarding the company's customer data activities, I feel:* | | .98/.98 | .91/.91 |
| ...violated. | .95/.95 | | |
| ...betrayed. | .96/.95 | | |
| ...not respected. | .96/.95 | | |
| ...taken advantage of. | .95/.95 | | |
| Cognitive Trust *(adapted from Palmatier 2008)* | | | |
| *Regarding this company's customer data activities, I think:* | | .96/.98 | .85/.94 |
| I trust the company. | .96/.97 | | |
| The company is very trustworthy. | .96/.97 | | |
| I have confidence in the company's behaviors. | .94/.97 | | |
| The company is reliable. | .94/.96 | | |
| Falsifying Information *(adapted from Lwin, Wirtz, and Williams 2007)* | | | |
| *When thinking about how I provide personal information to the company:* | | .96/.96 | .88/.89 |
| I am likely to give the company false information. | .95/.95 | | |
| I purposely try to trick the company when providing my personal data. | .96/.96 | | |
| I think it is fine to give misleading answers on personal questions. | .91/.92 | | |
| Negative Word of Mouth *(adapted from Gregoire and Fisher 2006)* | | | |
| *I would likely...* | | .97/.98 | .92/.95 |
| ...spread negative word of mouth about the company. | .97/.98 | | |
| ...bad-mouth the company to my friends, relatives, or acquaintances. | .97/.98 | | |
| ...tell others not to choose them if asked about their products/services. | .93/.96 | | |
| Switching Behavior *(adapted from Palmatier, Scheer, and Steenkamp 2007)* | | | |
| *If another company offered the same product/services but did not collect any data about your activities, how likely would you be to...* | | .95/.96 | .85/.90 |
| Shift all of my business to this new company? | .91/.94 | | |
| Try this new company's offering? | .94/.96 | | |
| Pay a premium to use this new company? | .92/.94 | | |
| ***Control Variables*** | | | |
| Value *(adapted from Chellappa and Sin 2005)* | | .94 | .80 |
| I receive value from the ways this company uses my customer data. | .90 | | |
| I save money (or can use free services) by providing my information. | .86 | | |
| I value how my information is used to customize my experience. | .93 | | |
| This company's saves me time by using my personal information. | .90 | | |
| Fairness *(new scale)* | | .97 | .89 |
| *Regarding this company's use of your customer information:* | | | |
| I believe their use of my customer information is fair. | .94 | | |
| I believe the company accesses my information in a fair way. | .96 | | |
| I believe the company's use of my information is ethical. | .94 | | |
| The company manages my information in an equitable way. | .94 | | |
| Privacy Concern *(adapted from Malhotra, Kim, and Agarwal 2004)* | | .94 | .80 |
| I am sensitive to the way companies handle my personal information. | .85 | | |
| It is important to keep my privacy intact from online companies. | .90 | | |
| Personal privacy is very important, compared to other subjects. | .92 | | |
| I am concerned about threats to my personal privacy. | .90 | | |
| Data Breach/Identity Theft Experience | | | |
| Customer Demographics (Gender, Age, City Population) | | | |

*Note:* Responses range from 1 = "strongly disagree" to 7 = "strongly agree," unless otherwise indicated.

**APPENDIX C**
**Study 2 Control Variables, Rationale for Model Inclusion, and Operationalization**

| Control Variables | Rationales | Operationalizations |
|---|---|---|
| Capital Resource Slack | Uncommitted resources can enable or prevent a firm from effectively managing a breach. | Ratio of a firm's annual sales to gross property, plant, and equipment (PPE) relative to its industry at a four-digit standard industrial classification (SIC) level (Modi and Mishra 2011) (μSalesi/PPEi)i−/(Salesi/PPEi))/(μSalesi/PPEi)i, for all firms in the same four-digit SIC. |
| Firm Size | Bigger firms might garner more negative | Log (Number of employees) |
| Industry | Effect of breach may vary by industry. | Dummy coding for industries: financial, retail, technology, online, or healthcare. |
| Competitive Intensity | Competitive rivalry may affect the market cost of the breach. | Herfindahl index: Sum of squared market share of the firm i, with the industry (I) defined at the three-digit SIC level, for the year prior to the breach. |
| Time | Reaction to breaches may strengthen/weaken over time. | Days since first breach in the sample timeframe. |
| Year Dummies | Year dummies control for macroeconomic effects. | Date of earliest public report of breach, converted into binary variables for years 2006–2015. |
| B2B vs. B2C | Effect of breach may vary by whether the firm focuses on business or end-customers. | The firm's primary four-digit SIC code classifies it as B2B (e.g., chemicals, primary metal, business services, engineering, accounting, research, management and related services) or B2C (e.g., food and kindred products, apparel, hotels, travel agents), using the scheme by Srinivasan et al. (2011) and Borah and Tellis (2014). |
| Goods vs. Services | Effect of breach may vary by whether the firm focuses on goods or services. | The firm's primary four-digit SIC code classifies it as goods (e.g., chemicals, primary metal, food and kindred products, apparel) or services (e.g., business services, engineering, accounting, research, management and related services, hotels, travel agents), using the scheme by Srinivasan et al. (2011) and Borah and Tellis (2014). |
| Focal Firm Prior Breaches | The negative effect of breaches for the firm might increase if the firm has had breaches in the past. | Count of the number of breaches of the focal firm |
| Rival Firm Prior Breaches | The negative effect of breaches for the firm might decrease if the closest rival has had breaches in the past. | Count of the number of breaches of the nearest rival of the focal firm |

## APPENDIX D
## Study 3 Manipulation Checks and Experiment Scenarios

**Experiment Scenarios and Customer Data Vulnerability Means**

***Introduction***

*From the list below, please select the [industry name] company whose products and services you use most often.*

Retail: Target, Walmart, Amazon, Costco, Best Buy
Technology: Apple, Microsoft, Google, Facebook, HP
Financial Services: Chase Bank, Wells Fargo, Bank of America, Citibank, American Express

*Please imagine that you receive the following email message from the company.*

**Data Access Vulnerability**

Thank you for being a valued customer of [Selected Company]. We appreciate our relationship with you. We are writing to inform you that the terms of our privacy policy have changed. You may access the full policy on our website. [Selected Company] is committed to protecting our customers against fraudulent activity. Our relationship with you and our other valued customers is our top priority.

Vulnerability Mean = 3.26 out of 7.00

**Data Breach Vulnerability**

Thank you for being a valued customer of [Selected Company]. We appreciate our relationship with you. Unfortunately, it has come to our attention that [Selected Company] has been the victim of a data breach. Through our internal investigation, we have determined that your customer profile was one of those compromised. However, at this time, [Selected Company] investigators have not detected any fraudulent activity in your account. Although we understand this is disappointing news, please know that our relationship with you and our other valued customers remains a top priority.

Vulnerability Mean = 5.57 out of 7.00

**Data Spillover Vulnerability**

Thank you for being a valued customer of [Selected Company]. We appreciate our relationship with you. It has come to our attention that our primary competitor has been the victim of a data breach. Although data breaches are becoming more common, at this time [Selected Company] investigators have not detected any fraudulent activity in your account. Although we understand this is disappointing news for some, please know that our relationship with you and our other valued customers remains a top priority.

Vulnerability Mean = 3.94 out of 7.00

**Data Manifest Vulnerability**

Thank you for being a valued customer of [Selected Company]. We appreciate our relationship with you. Unfortunately, it has come to our attention that [Selected Company] has been the victim of a data breach. Through our internal investigation, we have determined that your customer profile was one of those compromised. [Selected Company] investigators also have detected fraudulent activity in your account. Thus, it appears you have been the victim of identity theft. Although we understand this is disappointing news, please know that our relationship with you and our other valued customers remains a top priority.

Vulnerability Mean = 5.24 out of 7.00

# Data Privacy: Effects on Customer and Firm Performance

Kelly D. Martin, Abhishek Borah, & Robert W. Palmatier

## WEB APPENDIX A

## Details of Coding and Measure Validation for Transparency and Control

### Coding Procedure

We created the focal independent variables of transparency and control by employing a mix of automated and manual coding procedures. First, in order to capture the privacy policy nearest and earlier to a firm's breach date, we capture all the relevant URLs pertaining to the breached firm and its rival's privacy policy by using the Wayback Machine. In order to execute this, we developed a Python code that visited all valid snapshots of the firm's privacy policy. We then selected the policy that was the nearest and earlier to the firm's breach date. Second, in order to ensure that the correct URLs were downloaded and parsed, a manual layer of quality check was performed. Specifically, a random 5% of the URLs were checked to find if there were any errors in the code and the errors were corrected. We then resampled the URLs and found no errors. This process ensured that we correctly retrieved the privacy policy. Third, after obtaining the relevant privacy policy, we employed manual coding to construct the transparency and control variables, which consisted of carefully reading each privacy policy and using a coding schema to create count scores for transparency and control. For the variables that required coding of events, we followed standard procedures for textual coding (Lee 2001).

For the textual coding procedure, we first employed two research assistants who were blind to the study hypothesis. Second, prior to coding the privacy policies, the two research assistants were independently trained on a sample of privacy policies (that were not part of the final sample) to use the coding scheme. One of the co-authors checked to ensure that the research assistants understood the coding scheme. Third, after obtaining all the privacy policies, each

research assistant independently coded the privacy policy. Finally, after coding all the privacy policies, the interrater agreement between the two research assistants was greater than 85%, and all disagreements were resolved through discussion with the first author.

For the transparency variable, we used a count of the dummy variables across multiple elements of the privacy policy that signal openness and willingness to provide information to customers. Specifically, we coded whether the firm 1) explains its opt-out policy, 2) explains how it captures data, 3) explains how it uses data, 4) explains its data sharing internally and with third parties, and 5) provides contact information for privacy requests. If a firm's privacy policy has all five characteristics, the policy earns a score of 5 for transparency.

To create the control variable, we counted the number of opt-out choices in the firm's privacy policy, ranging from 0 to 5. Specifically, we coded whether the consumer 1) can opt out of marketing communication, 2) can opt out of saving data usage (e.g., search history), 3) can opt out of storing personal information (e.g., credit card number), 4) can opt out of sharing data with third parties and 5) can opt out of tracking. If a firm's privacy policy has all five characteristics above, the policy earns a score of 5 for control.

**Measure Validation**

Because establishing the causal priority between a latent variable (or construct) and its indicators can be difficult, we test whether the variables transparency and control are causal or effects indicators with the Vanishing Tetrad test for causal indicators (Bollen and Ting 2000). We use the Stata Module *Tetrad* developed by Shawn Bauldry and Kenneth Bollen to test the hypothesis. The null hypothesis for the Vanishing Tetrads test is that the indicators belong to an effect indicator model or are part of a reflective construct.

We first run the test for the transparency construct. Five variables comprise a total of 15 tetrads. The Tetrad module runs a simultaneous test on these tetrads and we obtain a bootstrapped

chi-square value of 132.36 with *p*-value of 0.00. Thus, in this case, a significant test statistic leads to the rejection of the effect indicator model in favor of the causal indicator one or a formative construct. We next run the Vanishing Tetrad test on the control construct. Again, five variables comprise a total of 15 tetrads. If our hypothesis is correct (i.e., the indicators are causal), the test statistic should be significant. We run the simultaneous test on these tetrads and obtain a bootstrapped chi-square value of 21.97 with a p-value of .25. The results are contrary to our expectation. However, as Bollen and Ting (2000) argue, it is important to run two checks.

First, we run a correlational check amongst the observed variables or items. Thus, we examine the covariance matrix to check whether small covariances between the indicators lead to vanishing tetrads. This is indeed the case with the correlations amongst the five variables being very small. Five of the ten correlations are below .10 and the maximum pair-wise correlation is .24. This result is consistent with a causal indicators but not with effect indicators. Second, we re-estimate an effect indicator model. The LR test of the model is not significant. That is we get a chi2(5) = 3.05, Prob > chi2 = 0.69. These results are inconsistent with the effect indicator model, but rather are consistent with the causal indicator specification (Bollen and Ting 2000).

Moreover, the Cronbach's alpha coefficient is 0.39, which shows that the indicators are not as closely related as a group. When we run an exploratory factor analysis, we get two factors with low loadings in each of the factors. These results support that our constructs of transparency and control are formative and not reflective.

## WEB APPENDIX B

### Study 2: Methodology to Calculate Abnormal Returns

We use daily data on stock market returns for each firm in our sample over a period from 270 to 6 trading days prior to the event day. Thus, we use 265 days of data to estimate the parameters. We remove announcements with estimation periods less than 30 days and estimate the following equation:

(B1) $R_{it} = \alpha_i + \beta_i R_{mt} + \varepsilon_{it}$

where *i* denotes the firm, *t* refers to time, $R_{it}$ are the returns for firm *i* on day *t*, $R_{mt}$ denotes the daily returns of the respective market index (e.g., CRSP Equally Weighted Index, German DAX index, Nikkei 225 index), and $\varepsilon_{it}$ denotes the disturbance term, which is an i.i.d. normally distributed error term (Brown and Warner 1985). The abnormal returns to an event on day *t* (Equation B2) is the difference between the normal return that would have occurred on that day, given no event, and the actual return that did occur because of the event, or:

(B2) $AR_{it} = R_{it} - \left\{\hat{\alpha}_i + \hat{\beta}_i R_{mt}\right\}$

where $\hat{\alpha}_i$ and $\hat{\beta}_i$ are the estimated parameters. Moreover, $AR_{it}$ provides an unbiased estimate of the future earnings generated by the event and is a random variable with a zero mean (Fama 1970; Homburg, Vollmayr, and Hahn 2014). With Equation B3, we calculate the cumulative average abnormal returns (CAAR) in an event window as:

(B3) $$CAAR_{t_1,t_2} = \frac{1}{N}\sum_{i=1}^{N}\sum_{t=t_1}^{t_2} AR_{it}$$

In the main text, we use the term “returns” to refer to cumulative average abnormal returns.

# WEB APPENDIX C

## Study 2: Counterfactual Analysis Using Propensity Score Matching

We evaluate the counterfactual returns of firms that offer high transparency and high control (High TC) changing to low transparency and low control (Low TC), and vice versa, using the propensity score matching method. With this method, we can estimate the causal impact of a treatment (High TC in our case) using counterfactuals. We first estimate the counterfactuals using propensity score matching and use those counterfactuals to evaluate the causal impact of firms that have High TC on returns.

Let the subscripts 0 indicate firms with Low TC and 1 indicate firms with High TC. Then $Y_{0t}$ is the returns of firm *i* if it has Low TC at time *t*, and $Y_{1t}$ is the returns of firm *i* if it has High TC at time t. In addition, $D_i \in \{0,1\}$ is the indicator of the actual privacy policy strategy (High TC or Low TC) adopted by firm *i*. Each firm earns returns in the state in which it is observed (actual outcome) and potential returns in the state in which it is not observed (counterfactual outcome). Thus, there are four potential returns outcomes:

1. The average returns of firms with **High TC** (observed).
2. The average returns of firms with **Low TC** (observed).
3. The average returns that firms with **High TC** would obtain if they had **Low TC** (estimated).
4. The average returns that firms with **Low TC** would obtain if they had **High TC** (estimated).

Outcomes 3 and 4 represent the focal counterfactuals. Therefore, the causal impact of High TC for firms with High TC (average treatment effect of the treated) can be defined formally as:

(C1) $$E(Y_{1t} - Y_{0t} \mid D = 1) = E(Y_{1t} \mid D = 1) - E(Y_{0t} \mid D = 1)$$

In Equation C1, we do not observe $E(Y_{0t}|D=1)$. This counterfactual represents the returns a firm with High TC would obtain if it had Low TC (outcome 3).

The causal impact of High TC for firms that have Low TC (average treatment effect of the untreated) can be defined formally as:

(C2) $$E(Y_{1t} - Y_{0t}|D=0) = E(Y_{1t}|D=0) - E(Y_{0t}|D=0)$$

In Equation C2, we do not observe $E(Y_{1t}|D=0)$. This counterfactual represents the returns a firm with Low TC would obtain if it had High TC (outcome 4). Thus, the first task is to estimate the counterfactuals and $E(Y_{1t}|D=0)$. We can use $E(Y_{0t}|D=0)$, which is *the observed returns of firms* with Low TC for the counterfactual, and $E(Y_{1t}|D=1)$, defined as *the observed returns of firms* with High TC for the counterfactual $E(Y_{1t}|D=0)$. The challenge is to check whether the observed returns represent the unobserved outcomes. For the counterfactual $E(Y_{0t}|D=1)$, we match firms with High TC with firms that have Low TC, using explanatory variables such as size or competitive intensity, then use the returns for the "matched" firms with Low TC to represent the *counterfactual*. However, it is difficult to find matches with close or exact values on all explanatory variables, and dealing with multiple covariates is challenging in terms of both computational and data issues.

Rosenbaum and Rubin (1983) introduce the propensity score, defined as the probability of receiving the treatment given the observed explanatory variables, to alleviate concerns about matching on all explanatory variables. Propensity score matching methods are widely used (Stuart 2010); they allow researchers to match a treated group with a control group on the basis of the distribution of the propensity score, rather than multiple explanatory variables. For this study, the propensity score is the conditional probability of being a firm with High TC, given the presence of "high transparency and high control" firm characteristics (Rosenbaum and Rubin 1983). There

are two main properties of propensity scores (Stuart 2010). The first is that the propensity score balances the distribution of the explanatory variables in the treated and control groups. The second property states that if treatment assignment is ignorable given the explanatory variables, then treatment assignment is also ignorable given the propensity score. This property justifies matching based on the propensity score rather than on the full set of explanatory variables. Propensity score matching involves three steps: (1) estimate the propensity score, (2) match firms on the basis of their propensity scores, and (3) evaluate the difference in outcomes (returns.in our case). We estimate propensity scores using a probit model. The dependent variable is an indicator variable, such that 1 denotes firms with High TC and 0 denotes firms with Low TC. We use the median split on the values of the interaction of transparency and control to create the binary dependent variable. As explanatory variables that might affect a firm's decision to be High TC, we use firm size, prior number of breaches, and industry fixed effects. We rely on nearest neighbor matching, such that each observation is matched with at least the specified number of observations from the other treatment level. Using one match per observation, we evaluate three effects using the estimated counterfactuals: average treatment effect of the treated (ATT), average treatment effect of the untreated (ATU), and average treatment effect (ATE), or the effect on all firms (treatment and control). The ATE provides the measure to evaluate whether the transparency × control interaction has a significant "causal" impact on returns.

We find that the counterfactual returns that a firm with High TC would obtain if it had Low TC are –.51%, with an ATT of .36%. The counterfactual returns that a firm with Low TC would obtain if it had High TC also are negative but less so (–.01%), with an ATU of .27%. The ATE is .31% (Abadie–Imbens robust standard error .0017599, $p = .08$) (Imbens 2000).