Received: 28 February 2019 | Revised: 10 June 2021 | Accepted: 19 June 2021
DOI: 10.1002/joom.1154

RESEARCH ARTICLE

WILEY

To err is human: Exploratory multilevel analysis of supply chain delivery delays

Amrou Awaysheh[1] | **Markham T. Frohlich**[1] | **Barbara B. Flynn**[1] | **Patrick J. Flynn**[2]

[1]Kelley School of Business, Indiana University, Indianapolis, Indiana, USA
[2]Poole College of Management, North Carolina State University, Raleigh, North Carolina, USA

Correspondence
Barbara B. Flynn, Kelley School of Business, Indiana University Indianapolis, IN 46202.
Email: bbflynn@indiana.edu

Handling Editor: Gregory Heim

ABSTRACT

We examine the impact of human errors by front-line supply chain employees on delivery delays. We build on normal accident theory (NAT), a multilevel theory describing the relationship between a firm's latent conditions (systemic managerial, technology, and social factors) and human errors. Latent conditions can have the unintended consequence of intensifying the impact of a human error, thus, we hypothesize a moderating effect of latent conditions on the relationship between errors and delivery delays. Hypotheses are tested using archival shipment data provided by a Fortune 500 company and archival carrier violations data. A multilevel design, with 299,399 shipments (level 1) nested within 97 carriers (level 2), was tested using mixed effects regression modeling. The results indicated that both dispatcher and driver errors were related to the probability of a late delivery, and that dispatcher errors were associated with longer delays. The moderating effects of several carrier latent conditions were significant and positive, indicating that both types of errors were more strongly associated with the likelihood of late delivery and that dispatcher errors were associated with longer delays when moderated by carrier latent conditions. The results are discussed from the perspective of NAT and technology management, developing prescriptions, and suggesting opportunities for future research.

KEYWORDS
errors, normal accident theory, supply chain management, technology

1 | INTRODUCTION

Technology has played an important role in developing processes resilient to human error (Adelman, 2019). Important technology-related operational improvements have been achieved across such diverse contexts as retail (Asare et al., 2016), construction (Nnaji et al., 2019), healthcare (Douglas & Larrabee, 2003; Harrington et al., 2011; Knoedler, 2003), retail pharmacies (Elliott et al., 2014), outsourcing to developing country suppliers (Thomas & Ray, 2019), nuclear power generation (Stock & McFadden, 2017), municipal government (Sanford & Bhattacherjee, 2007), logistics (Cantor et al., 2016; Hickman et al., 2015; Scott et al., 2021), aviation (DeFlorio, 2016; Peysakhovich et al., 2018; Rusu et al., 2012), pharmaceutical manufacturing (Markarian, 2019), and automotive manufacturing (Ponnaluri, 2019; Subit et al., 2017). Error reduction technologies include computerized physician order entry (Cowan, 2003; Harrington et al., 2011), vendor-managed inventory (Asare et al., 2016), automated driving technologies (Ponnaluri, 2019; Subit et al., 2017), bar-coded medication administration (Douglas & Larrabee, 2003; Harrington et al., 2011), collaborative decision support systems (Rusu et al., 2012), smart construction helmets (Nnaji et al., 2019), point-of-sale technologies (Asare et al., 2016), eye tracking technology (Peysakhovich et al., 2018), and many other innovative technologies, with a goal of automating human variability out of processes (Granot, 1998).

Why, then, do so many reports of human errors continue to persist, despite investments in technology (Yusuf & Sahroni, 2018)? Explanations include organizational culture (Averett, 2001; Sanford & Bhattacherjee, 2007), unintended consequences that actually cause an increase in errors, such as reduced nurse–physician interaction when computerized physician order entry systems are used (Cowan, 2003), overreliance on technology (Ngwenyama & Nielsen, 2014; Peysakhovich et al., 2018), late employee involvement (Averett, 2001), weak influence skills (Ngwenyama & Nielsen, 2014; O'Conor & Smallman, 1995), inadequate communication (Averett, 2001; Cowan, 2003), lack of perceived benefit (Ngwenyama & Nielsen, 2014), employee resistance (Averett, 2001; Markarian, 2019), technology alert fatigue (Cowan, 2003; Yusuf & Sahroni, 2018), and failure to consider the relationship between a technology, the firm in which it is implemented, and its environment (Gangwa et al., 2014). Scott et al. (2021) describe the case of a monitoring technology that achieved its primary goal of preventing truck drivers from exceeding their maximum daily hours of service, but led to the unintended consequence of causing drivers to speed or drive unsafely. Clearly, social and managerial factors are important to a technology's success, in addition to its unique features (Averett, 2001; Ngwenyama & Nielsen, 2014). In other words, as operations improve through technology, what remains is human error (Sterns & Keller, 1991). For example, in aviation, "technical failure today is the cause of only about 10% of accidents, leaving a significant percentage the implications of human error" (Peysakhovich et al., 2018, p. 1).

This suggests the importance of a socio-technical systems approach, based on two principles (Bednar & Welch, 2020; Leitch & Warren, 2010; Winter et al., 2014): (a) *the interaction between human and technical factors creates the conditions for successful outcomes*, and (b) *optimization of technology or human factors, alone, is suboptimal and will lead to less successful outcomes*. The human element introduces uncertainty, the potential for error, and unintended consequences (Harrington et al., 2011; Scott et al., 2021; Taylor & Robinson, 2015). For example, decision support systems may introduce new opportunities for human errors (Beiger & Kichak, 2004), including over-trusting a technology (Tsai et al., 2003), overriding information not perceived as beneficial (Van der Sijs et al., 2009), and reduced vigilance (Harrington et al., 2011).

We examine variability caused by human error and its impact on performance in a context where technology provides a potential safeguard against adverse consequences. We draw upon normal accident theory (NAT) (Perrow, 1999; Reason, 2000), which describes human errors and their potential for adverse consequences, which NAT defines as disruptions to system outputs. NAT views organizations as complex systems, delineating the interrelationships between human and system characteristics that lead to adverse consequences. We focus on front-line operators in a supply chain: truck drivers conveying shipments and dispatchers interfacing between drivers and carriers. Like all people, drivers and dispatchers sometimes unintentionally commit errors. Because supply chain members are tightly coupled, with little slack, even a minor error can have adverse consequences.

To prevent driver errors, carriers invest in sophisticated technologies such as speed regulators, collision mitigation systems, biometric fatigue sensors, video monitoring, rollover stability systems, geo-fencing, and lane departure warning systems. Since late 2017, the Federal Motor Carrier Safety Administration (FMCSA) has required onboard electronic logging devices (ELDs) that report drivers' speed, idle time, hard braking, vehicle location, engine operating hours, and vehicle miles at frequent intervals. Many carriers have embraced this technology, voluntarily exceeding the mandate to ensure the safety of their drivers, vehicles, cargo, and the public. To prevent dispatcher errors, carriers and shippers invest in transportation management systems (TMS) (see Appendix) to support shipment planning, routing, carrier mix, matching cargo with vehicles, and tracking shipments. Although TMS technology has provided powerful managerial tools for decades (see Moore et al., 1991 for an early application), it is not infallible (Lechler et al., 2019; Mehmood et al., 2017). Supply chains are inherently complex, involving many stakeholders, from customers, shippers, and carriers to government regulators and the public (Bode & Wagner, 2015; Serdarasan, 2013). While information technologies can improve efficiency and reduce uncertainty (Karimi et al., 2004; Melville & Ramirez, 2008; Wu et al., 2013), the data driving them must be accurate and timely (Galbraith, 2014). Further, employees must be able to understand and act on the outputs of information technology (IT). For example, according to a shipping company manager we met with,

> "Having a $10,000 computer that's hooked up to a $60 million GPS system doesn't eliminate all problems. There is still a person behind the keyboard who is acting on the information or relaying that information. I wish I could control or predict that person, but human nature just kicks in."

The human nature he refers to is unexpected outcomes resulting from human errors, which are inevitable in a human–technology interface (Chakravorty, 2011). Thus, we explore human errors by drivers and dispatchers that paradoxically occur despite prevention technology investments.

Specifically, we investigate whether human errors are more likely to have shipment delay implications in the context of various carrier latent conditions (systemic factors). Latent conditions include managerial, technology, and social factors, such as hiring and training practices, policies, priorities, equipment, IT, norms, values, and safety culture. Latent conditions that intensify the negative effects of a human error are known as resident pathogens, while latent conditions that reduce the negative consequences of a human error are known as defensive layers. For example, for a carrier that routinely employs inexperienced drivers or drivers with a poor driving record (resident pathogen), a minor dispatcher error is more likely to lead to adverse consequences. We develop hypotheses about the cross-level moderating effect of carrier latent conditions on the likelihood that a dispatcher or driver error will lead to a late delivery or increase delivery delay.

We contribute to the technology management literature by motivating worrisome questions about the human–technology interface, such as (a) With so much TMS and tractor-trailer IT in place, what are the consequences of human errors? (b) How can the consequences of human errors be lessened through addressing carrier latent conditions, including managerial, technology, and social factors? This exploratory research underlines the importance of latent conditions in reducing or intensifying the adverse consequences of human errors by developing defensive layers through management policies and priorities, investments in technology, and carrier norms and values.

We begin by drawing on NAT's theoretical foundation to describe human errors and latent conditions in supply chain operations and propose hypotheses about the direct effect of errors and the cross-level moderating effects of latent conditions. We test a two-level model using mixed effects regression analyses, with archival data on 299,399 shipments (level 1) by a Fortune 500 company, nested in the 97 carriers (level 2) that conveyed them. The results provide support for the hypothesized main effects and moderating effects of carrier latent conditions. We conclude by suggesting practical actions carriers can take to reduce resident pathogens and strengthen defensive layers to lessen the impact of human errors, as well as posing opportunities for future research.

2 | THEORETICAL FOUNDATION

NAT describes how the tight coupling and complex interactions of high-risk organizations (Adelman, 2019) cause adverse consequences to be perceived as "normal" (Speier et al., 2011; Stock & McFadden, 2017; Weick, 2004). Supply chain operations share many assumptions about high-risk organizations (Bigley & Roberts, 2001), including important decisions made by low-level personnel under minimal supervision (truck drivers on the road) (Caplice, 2007; Short et al., 2007), tight coupling, importance of subsystem interfaces, coordination, and movement (Grabowski & Roberts, 1997), time-dependent processes, minimal slack, and complex interactions (Perrow, 1999). Although supply chain operations do not have the potential for the catastrophic adverse consequences of high-risk organizations such as a nuclear power plant (Stock & McFadden, 2017) or space shuttle (Vaughan, 1996), a relatively minor assignment error by a dispatcher or loading error by a driver can cause delays with significant implications. For example, consider shipper Company A, which was promised delivery of a shipment by a carrier between 9:00 am and noon on Monday. Company A arranged for the carrier to break down the load and last mile deliver it to 40 of Company A's customers on Monday afternoon and Tuesday morning, within 3-h windows. If the carrier's delivery is delayed, Company A will need to make 40 calls to its customers to arrange 40 new delivery windows. If such delays become habitual, Company A will come to be regarded as unreliable, causing its customers to take their business elsewhere.

NAT has three key constructs: adverse consequences, active failures, and latent conditions. We describe each below, provide supply chain examples, and summarize NAT's key propositions.

2.1 | Adverse consequences

NAT's first key construct is adverse consequences, which it operationalizes as accidents. NAT defines an accident as "damage to a defined system that disrupts the ongoing or future outputs of that system" (Perrow, 1999, p. 64). In a supply chain, we focus on late deliveries (rather than crashes) as adverse consequences for several reasons. First, this is consistent with supply chain goals; a reliable supply chain delivers its goods on time. Second, supply chain delivery delays are important system outputs that can cause strained relationships, reduced future business, and levied financial penalties (Bigley & Roberts, 2001; Hubbard, 2003; Miller, Golicic, & Fugate, 2017a). Third, because the number of crashes per opportunity is relatively low, more can be learned (Roberts & Rousseau, 1989; Weick, 1987) about supply chain performance through the study of late deliveries. To avoid confusion with crashes, we use the generic term "adverse consequences," rather than "accidents"; thus, an *adverse consequence in supply chain operations is a delayed delivery*. NAT proposes adverse consequences result from a combination of individual (active failures) and systemic (latent conditions) factors (Reason, 1990).

2.2 | Active failures

NAT's second key construct is active failures. An *active failure is an unintentional human error that has an immediate, often negative, effect* (Howe et al., 2020; Kanki & Hobbs, 2018). People frequently make errors when engaged in the automatic, unconscious cognitive processing known as skill-based processing (Stewart & Grout, 2001). If a situation warrants it, people shift from skill-based processing to rule-based processing (applying cognitive rules developed through past experiences) or to knowledge-based processing, for which their prior knowledge and experience is insufficient (Stewart & Chase, 1999). Different types of errors are associated with each type of cognitive processing (see Table 1; Figure 1).

TABLE 1 Examples of active failures

Type of cognitive processing	Driver errors	Dispatcher errors
Skill-based	• Place losing error: Driver accidentally misses a turn	• Reversal error: Dispatcher reverses digits when entering the delivery address into the TMS
	• Capture error: Driver misreads the map, while thinking about a personal situation	• Description error: Dispatcher enters billing address into the delivery address field
Rule-based	• Misapplication of a good rule: Faced with an obstacle, driver hits the brakes while driving on an icy road	• Context error: Dispatcher assigns a standard 53′ trailer to a shipment to a dense urban destination
	• Application of a bad rule: To avoid a late departure, driver departs without ensuring cargo is properly secured	• Rigidity error: dispatcher consistently fails to look at the "notes" section of a P.O., failing to provide driver with delivery specifics
Knowledge-based	• Symbol manipulation error: Driver does not know how to interpret map correctly	• Knowledge error: Dispatcher does not know which types of loads require safety roil bars in the power units
	• Logical deduction error: Driver fails to investigate engine rattle, assuming it is acceptable because the engine is still running	• Inadequate retrieval error: Dispatcher does not know how to respond to driver information system in the track/trace system

2.2.1 | Skill-based errors

Skill-based cognitive processing uses automatic functioning (Spender, 1996) of hierarchically organized cognitive routines and subroutines (Norman, 1981) as people engage in common activities and make everyday decisions (Stewart & Grout, 2001). Actions governing familiar activities are stored as mental subroutines that operate without conscious intervention, once activated (Stewart & Chase, 1999). For example, a person wanting to go to the kitchen would activate the "walking" routine, which then automatically triggers subroutines for getting up, lifting and placing feet, balance, locomotion, and so on. Skill-based processing errors result from overtaxing cognitive resources (Stewart & Grout, 2001), causing subroutines to become "unruly" (Reason & Mycielska, 1982) through loss of subroutine activation (Hofmann et al., 1995), incorrect subroutine triggering by a similar cue (Sutcliffe & Rugg, 1998), or failure of the correct cue to trigger a subroutine (Norman, 1981). As anyone who has ever forgotten why he or she walked into the kitchen understands, skill-based errors are inevitable and automatic. These "stupid mistakes" (Chakravorty, 2011) are

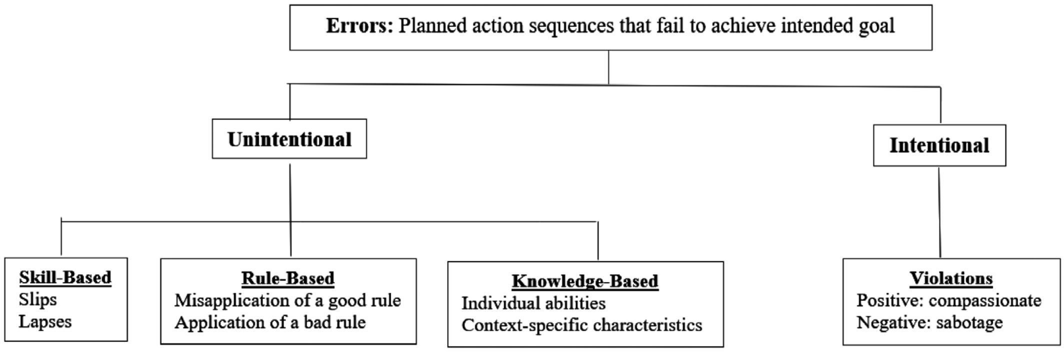

FIGURE 1 Taxonomy of errors (Reprinted with permission Reason, 1990)

the most common type of human error, are unavoidable, and have the most trivial impact (Reason, 2000).

2.2.2 | Rule-based errors

People apply rule-based cognitive processing to less familiar situations as they search for a familiar pattern so they can apply a stored rule (Stewart & Chase, 1999) for determining which subroutines to activate (Rasmussen, 1982; Rasmussen & Jenssen, 1974). This pattern search is virtually automatic (Stewart & Chase, 1999). For example, a professional chess player can quickly determine an effective move after a rapid mental search through numerous patterns of chess plays and their associated benefits and dangers (Simon, 1981). Rule-based errors include application of the wrong rule, incorrect application of the right rule, failing to recognize a situation where a rule should have been applied, rigidity in applying a suboptimal rule, and correct triggering of a malformed rule (Sutcliffe & Rugg, 1998). They are less common than skill-based errors but can have more substantial consequences. Because rule-based errors depend on the extent to which a person has developed effective rules and understands when these rules should be applied, rule-based errors can be prevented by a better search process, nurtured through experience and training (Stewart & Chase, 1999).

2.2.3 | Knowledge-based errors

People apply knowledge-based cognitive processing to novel problems with which they have limited experience (Rasmussen, 1982). Without automatic subroutines or stored rules to draw upon, they must apply internal processing, logical deduction, and symbol manipulation to form a mental picture of a new problem and search for a solution (Stewart & Chase, 1999). Knowledge-based errors are less common than skill-based or rule-based errors, however, their impact is often more serious. Because of the internal processing required, knowledge-based errors are related to individual abilities and context-specific conditions (Rasmussen, 1983), thus they are more preventable through managerial decisions, training, investments in technology, clear procedures, a supportive culture, and other factors.

Thus, an active failure is an unintentional error that occurs when a person's planned action sequence fails to achieve its intended goal (Reason, 1990). Active failures encompass all three types of cognitive processing: skill-based, rule-based, and knowledge-based. NAT predicts that active failures are associated with adverse consequences. It is important to note that, because of their unconscious and "normal" character, errors are not usually reported or measured (Adelman, 2019), making error-related research challenging. However, errors can be deduced from their immediate effects, such as actions taken to compensate for an error, which are more likely to be recorded. Table 2 lists the arrival codes for deliveries by the shipper we studied, which we call Large Integrated Manufacturer (LIM). The codes describe the immediate effects of unobservable errors; for example, arrival code 5 could be the result of a dispatcher's reversal error, while arrival code 17 could be due to a driver's place-losing error.

2.3 | Latent conditions

NAT's final key construct is latent conditions. *Latent conditions are systemic managerial, technology, and social conditions that define a firm* (Kanki & Hobbs, 2018). *Managerial latent conditions* are embedded in policies and priorities that signal employees about appropriate and inappropriate actions. *Technology latent conditions* encompass investments in technology to encourage or support desirable actions and prevent undesirable

TABLE 2 LIM's arrival codes

Code	Explanation	Code	Explanation
1	Normal delivery, no issues	9	Incorrect address provided on documents
2	Carrier has no drivers available	10	Other carrier-related reason
3	Dispatch mistake	11	Driver related departure delay
4	Insufficient delivery time allowed	12	Driver related arrival delay
5	Dispatcher keying mistake	13	Driver took too long at previous stop
6	Late departure from origin: mechanical breakdown	14	Driver missed de lively window
7	Late arrival at destination: mechanical breakdown	15	Driver missed pickup window
8	Consignee-related reason	16	Driver rest delay
		17	Driver transportation delay

actions. *Social latent conditions* reflect a firm's "… assumptions and values, usually implicit, about how to interpret organizational reality … and how to succeed (Thornton, 2004, p. 76)." They develop through convergent thinking within an industry or firm as processes take on a "rule-like status" (Meyer & Rowan, 1977) over time, causing the development of "scripts" (Brammer et al., 2012), activated during rule-based processing, that enable or constrain actions.

Latent conditions can combine with an active failure to trigger adverse consequences (Kanki & Hobbs, 2018); unintentional errors are exacerbated by latent conditions such as insufficient planning, weak preventive technology, inappropriate priorities, or poor safety culture (Yusuf & Sahroni, 2018). Leape (1994) likens latent conditions to a patient's immune system. A weak immune system does not cause illness, but if a patient is exposed (active failure) to a virus, a weak immune system makes it more likely that illness (adverse consequence) will result. On the other hand, a strong immune system makes it less likely that exposure to a virus will result in illness. Thus, NAT predicts that the effect of an active failure on adverse consequences is moderated by latent conditions. Latent conditions are comprised of resident pathogens and defensive layers (Reason, 2000).

2.3.1 | Resident pathogens

Resident pathogens are managerial, technology, and social factors that increase the likelihood an active failure will have adverse consequences, according to NAT. The more complex and tightly coupled a firm is, the more resident pathogens its latent conditions contain (Reason, 2000). Managerial resident pathogens can inadvertently signal dispatchers and drivers that a carrier places a low priority on driving safety (Detert et al., 2000; Short et al., 2007). For example, a policy that pays drivers a set rate per loaded mile[1] may signal that driving as far as possible each day is more important than safety (Cantor et al., 2013; Scott & Nyaga, 2019). Other managerial resident pathogens include a carrier's tendency to defer vehicle maintenance or tolerate speeding (e.g., a policy of reimbursing drivers for speeding tickets).

Technology resident pathogens include specific technologies, integration with other technologies, and the ability to understand and act upon the information generated by technology (Yusuf & Sahroni, 2018). Minimal TMS investments provide dispatchers limited support in their routing, loading, and tracking decisions, leaving more decisions to their discretion, and setting the stage for rule-based and knowledge-based errors. Poor integration (Rusu et al., 2012) between a carrier's TMS and a shipper's TMS can distort data flows between the two companies. However, even with a sophisticated, well-integrated TMS, a dispatcher may lack the ability to effectively capture, analyze, and act upon the high volumes of information generated.

> Our manual system was slow and we made 10 mistakes every day, and now the automated system is fast—and we make 1000 mistakes every day (Chakravorty, 2011, p. 96).

Social resident pathogens cause perceptions of "how we do things around here" (Hofmann & Stetzer, 1996). Stories about how drivers circumvented rules, policies, and monitoring technologies in the past evolve into "rationalized myths" (Meyer & Rowan, 1977) over time. For example, two rationalized myths discourage error reporting: the perfection myth (if employees try hard enough, they will not commit any mistakes) and the punishment myth (if employees are punished for errors, they will commit fewer of them) (Meyer & Rowan, 1977).

2.3.2 | Defensive layers

Defensive layers are managerial, technology, and social factors that prevent an active failure from having a negative impact, according to NAT. Managerial defensive layers include clear procedures, employee training, control systems, supervisory systems, quality control standards, and priorities (Aird, 2019; Howe et al., 2020; Kanki & Hobbs, 2018). Examples include ecological interface design (Rasmussen, 1983; Rasmussen et al., 1994), training simulators (Stewart & Grout, 2001), cognitive engineering (Rasmussen, 1987; Rasmussen, 1988), compensation policies, priorities, and vehicle maintenance policies.

Technology defensive layers include investments in error monitoring and prevention technologies. TMS supports dispatchers in configuring shipments, putting them out for bid, and managing service contract execution. Driver technology defensive layers include sensors, go/no-go gauges, and switches (Shingo, 1986). For example, if a driver commits a rule-based error by shifting into the wrong gear (correct triggering of the wrong rule) or hitting the brakes on an icy surface (misapplication of a good rule), the potential for adverse consequences will be less if the vehicle is equipped with speed regulating technology. The FMCSA's ELD mandates (Scott et al., 2021) provide a technology defensive layer by monitoring hours-of-service (HOS) compliance by drivers. However, because it does not apply to vehicles manufactured before 2000, carriers can purchase grandfathered trucks or hire owner-operators who own grandfathered trucks, to avoid compliance. Thus, while ELDs are a technology defensive layer, a policy of allowing grandfathered trucks is a managerial resident pathogen.

Social defensive layers are the positive social institutions (Meyer & Rowan, 1977) that combine to form a firm's safety culture, which is "sets of norms, roles,

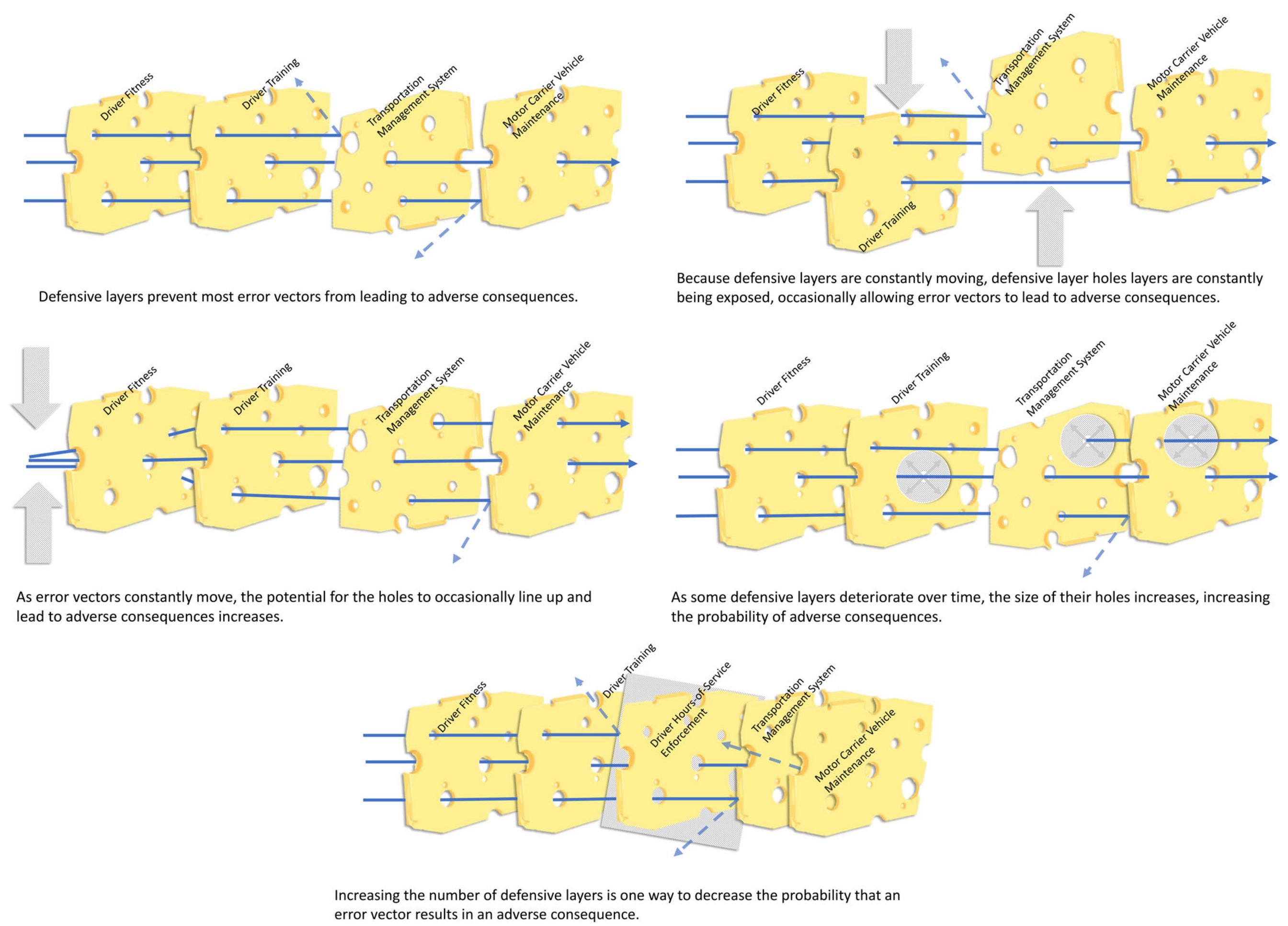

FIGURE 2 Swiss cheese model of error vectors and defensive layers

beliefs, attitudes, and … practices within an organization … concerned with minimizing the exposure of individuals to conditions considered to be dangerous (Taft & Reynolds, 1994, p. 3).” If a firm's safety culture is strong, failing to follow a safety protocol would be unthinkable.

NAT describes defensive layers as being like parallel slices of Swiss cheese (see Figure 2); each defensive layer (slice) has some holes (resident pathogens) through which the impact of an error can potentially pass (Reason, 1990). Since defensive layer holes do not normally line up, an error's impact passing through one defensive layer will be stopped by the subsequent layer. However, defensive layers are constantly shifting, with holes that simultaneously expand and contract, so sometimes the holes in all layers align. For example, the momentary alignment of defensive layer holes caused a skill-based error like a leak of less than one cup of water (Three Mile Island) or a knowledge-based error by an operator unfamiliar with emergency protocols (Chernobyl) to have catastrophic adverse consequences (Reason, 2000).

NAT predicts the strength of a carrier's defensive layers determines the extent to which the negative impact of errors will lead to adverse consequences. Stronger defensive layers are characterized by (a) more layers, (b) fewer holes in each layer, (c) smaller holes, and (d) less shifting of layers.

3 | HYPOTHESES

Figure 3 illustrates our research model. Building on NAT, we propose that dispatcher and driver errors are associated with adverse consequences and that carrier latent conditions moderate this effect. Because errors and their consequences occur at the shipment level, while latent conditions exist at the carrier level, testing this model necessitates a multi-level design.

3.1 | Impact of active failures (Level 1)

Level 1 is comprised of errors by dispatchers and drivers. We hypothesize that dispatcher errors increase the likelihood of adverse consequences. Examples of dispatcher errors include skill-based errors like forgetting to notify a driver about a shipment, entering incorrect address information, or failing to enter “Notes” from the P.O. (“use the second door after 5:00”), rule-based errors like assigning chilled products to a standard trailer (Robinson et al., 2013), assigning hazardous materials to a non-Homeland Security screened trailer, or assigning a standard power unit to a trailer requiring a power unit with a reinforced metal roll bar (Hickman et al., 2015), and knowledge-based errors like misreading the TMS's outputs. Each of these errors increases the likelihood of a delayed delivery as the driver engages in time-consuming tasks to compensate, such as changing vehicles, hunting for the right door, or driving to the correct location. Thus,

> **Hypothesis 1a (H1a).** Dispatcher active failures are positively associated with the likelihood of delayed delivery.

We also hypothesize that driver errors increase the likelihood of adverse consequences. The monotony of

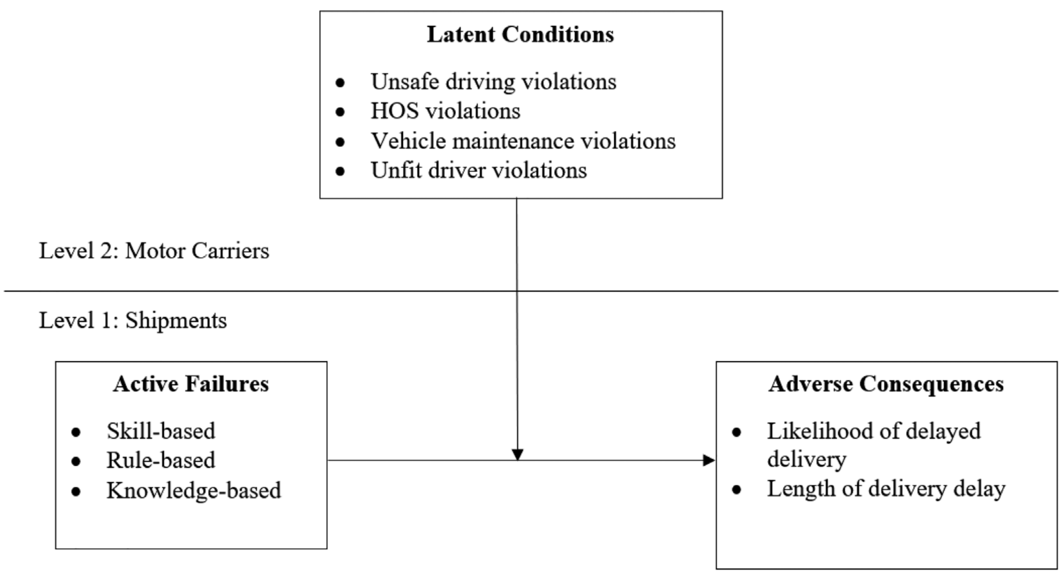

FIGURE 3 Research model

loading, driving 500 or more miles, unloading, then repeating the process the next day (Johnson et al., 2009) can lead to inattentiveness, fatigue, indifference, and slower reaction times that set the stage for skill-based errors like neglecting to oversee the loading of some cargo (omission error), missing a turn (place losing error), or dozing (intrusion error). Weeks and months of travel across thousands of miles can cause a driver to develop an inflated view of his or her ability, causing rule-based errors related to risk taking and poor judgment (Douglas & Swartz, 2016), such as speeding or loading a trailer beyond its weight limit. Driver errors result in an increased likelihood of delayed delivery as the driver compensates, such as reloading the trailer.

> **Hypothesis 1b (H1b).** Driver active failures are positively associated with the likelihood of delayed delivery.

3.2 | Moderating effect of latent conditions on delay likelihood (Level 2)

At the carrier level (Level 2) latent conditions are inherently challenging to measure. Managerial, technology, and social latent conditions are based on deeply held, tacit values that emerge gradually, based on the carrier's unique history, strategy, external environment, management style, and employees (Cameron & Quinn, 2011). Schein (2004) recommends using artifacts (visible structures and processes) to impute values. Consider the analogy of a patient's immune system. Although it is difficult to measure directly, strength of a patient's immune system can be imputed from the artifact of his or her record of illnesses. Because violations are unsafe acts in which employees intentionally bypass policies or procedures (Howe et al., 2020), they reflect a firm's values, resident pathogens, and defensive layers. Thus, we can impute a carrier's latent conditions from its record of violations.

3.2.1 | Unsafe driving violations

A carrier's record of unsafe driving violations is an artifact of latent conditions reflecting its tolerance for speeding, illegal turns, quick lane changes, or driving too fast for conditions. For example, managerial resident pathogens like a policy of no pay for out-of-route miles could incentivize making an illegal U-turn (Swartz et al., 2017) to compensate for an error, and mileage-based pay could encourage driving in severe weather conditions (Cantor et al., 2009). Technology resident pathogens contributing to tolerance for unsafe practices include a TMS with few built-in checks; resulting dispatcher data entry errors could cause a driver to speed, make an illegal turn, or make a quick lane change to compensate. The inability of a carrier to compile and analyze data from the onboard ELD or track/trace system prevents dispatchers from realizing when unsafe driving occurs. Managerial resident pathogens that signal the acceptability of assigning shipments that cannot possibly be delivered on time trigger social resident pathogens through scripts that rationalize the need to speed or tamper with onboard monitoring technology, based on stories glorifying previous drivers who did so.

> … there is empirical evidence that a carrier's … rewards and penalties for drivers create a perceived pressure to 'bend rules,' resulting in scheduling drivers that are still fatigued from previous work and having them rush shipments (Cantor et al., 2013, p. 39).

Managerial defensive layers include policies about the priority of safety versus on-time delivery, hiring only drivers with a safe driving record, punishing or dismissing dispatchers and drivers whose practices have contributed to unsafe driving violations, and training dispatchers to avoid inadvertently encouraging drivers to engage in unsafe driving practices. Technology defensive layers include lane departure warning systems (McNall & Stanton, 2011), video monitoring (Hickman & Hanowski, 2011), geo-fencing notifying dispatchers of off-route shipments, and speed regulators (Cantor et al., 2009).

When a dispatcher or driver makes an error, the likelihood of adverse consequences is greater if latent conditions encourage driver fatigue, stress, or overconfidence (Brown, 1996; Cantor et al., 2009; Ray et al., 1993). On the other hand, if latent conditions encourage drivers to be alert, well-rested, and level-headed, an error is less likely to have adverse consequences. Thus,

> **Hypothesis 2 (H2).** A carrier's record of unsafe driving violations moderates the likelihood that an active failure by a (a) dispatcher or (b) driver results in delayed delivery, such that the likelihood of delayed delivery will be enhanced for shipments by carriers with more unsafe driving violations.

3.2.2 | Hours of service violations

A carrier's record of hours of service (HOS) violations is an artifact of its tolerance for exceeding HOS limits.[2]

Managerial resident pathogens include dispatchers routinely assigning long, irregular driving hours (Saltzman & Belzer, 2002), a policy of using grandfathered trucks without onboard ELD systems, penalties for failing to deliver the last shipment at the end of the day (Swartz et al., 2017), mileage-based pay (Miller & Saldanha, 2016), and rewards for on-time delivery (Hofmann & Stetzer, 1998). Technology resident pathogens include TMS that allows assignments that could not possibly be delivered on time, carriers failing to act upon ELD reports of drivers exceeding HOS limits (Miller & Saldanha, 2016), manual logbooks, and non-tamperproof ELD systems. Because people interpret and attach meaning to their work environment (Hofmann & Stetzer, 1998), drivers' perceptions of behavior-outcome contingencies embedded in carrier latent conditions create social resident pathogens such as the rationalized myth that "taking risks is simply part of [the] job" (Hofmann & Stetzer, 1996, p. 310).

Managerial defensive layers include a policy of only hiring drivers with low HOS violations, well-communicated policies about mandatory rest periods, HOS compliance incentives, and fines or termination for HOS violations. Defensive technology layers against HOS violations include control poka-yokes (Shingo, 1986) that shut down a vehicle when its HOS limit is reached, biometric fatigue sensors (Cantor et al., 2008), regular monitoring of ELD data, and tamper-proof ELD devices. Over time, these work together to develop the social defensive layer norm that exceeding HOS limits is unthinkable.

If an error occurs in a system that tolerates exceeding HOS limits, the likelihood of a delivery delay is greater as the driver has no disincentive to avoid exceeding HOS limits and fatigue increases the likelihood of further skill-based errors, selection of the wrong rules, and activation of rules by the wrong triggers. Additional errors increase the likelihood of a delivery delay due to the time needed to compensate. Thus, when there is a dispatcher or driver error, the likelihood of adverse consequences increases when the carrier has latent conditions that enable or tolerate exceeding HOS limits. On the other hand, if latent conditions support adherence to HOS limits, an error is less likely to have adverse consequences.

> **Hypothesis 3 (H3).** A carrier's record of hour-of-service violations moderates the likelihood that an active failure by a (a) dispatcher or (b) driver results in delayed delivery, such that the likelihood of delayed delivery will be enhanced for shipments by carriers with more hours-of-service violations.

3.2.3 | Vehicle maintenance violations

A carrier's record of vehicle maintenance violations reflects its tolerance for keeping vehicles needing maintenance in service. Managerial resident pathogens include investments in poorly designed or aging vehicles prone to maintenance issues (Britto et al., 2010), a policy of deferring maintenance (Hubbard, 2003), limited spending on vehicle maintenance (Miller & Saldanha, 2016), and financial incentives for drivers to keep vehicles on the road. Technology resident pathogens encouraging vehicle maintenance violations include poor TMS maintenance scheduling modules, lack of onboard maintenance indicators, dispatchers encouraging drivers to ignore onboard maintenance indicators, and dispatchers' inability to collate, analyze, and act on ELD maintenance data. Social resident pathogens include the rationalized myth that staying in service is preferable to shutting down for scheduled maintenance.

Managerial defensive layers include driver incentives for following vehicle maintenance schedules, priority on keeping vehicles well maintained, driver penalties for deferring required maintenance, and a policy of purchasing only new or well-maintained used vehicles. Technology defensive layers include ELDs providing real time data on engine conditions, sensing devices, devices that lock the ignition of a vehicle with a maintenance issue, and tamper-proof warning signals. Social defensive layers are based on norms related to the importance of preventive maintenance.

If a driver commits an error while driving a poorly maintained vehicle, delayed delivery is more likely because of the increased likelihood of a breakdown while driving extra miles to remedy it. On the other hand, a carrier that exhibits pristine vehicle maintenance compliance has defensive layers that lessen the impact of unexpected dispatcher and driver errors. Thus, when a dispatcher or driver makes an error, the likelihood of adverse consequences increases when the carrier has latent conditions that enable vehicle maintenance violations.

> **Hypothesis 4 (H4).** A carrier's record of vehicle maintenance violations moderates the likelihood that an active failure by a (a) dispatcher or (b) driver results in delayed delivery, such that the likelihood of delayed delivery will be enhanced for shipments by carriers with more vehicle maintenance violations.

3.2.4 | Unfit driver violations

A carrier's record of unfit driver violations indicates its tolerance of driving by ill, unlicensed (Hagberg et al., 1995), or insufficiently rested drivers. Managerial resident pathogens include incentives encouraging drivers to drive or face no pay, and a policy of hiring drivers with questionable driving records (Granot, 1998; Miller, Saldanha, et al., 2017b) or less qualified drivers who are unable to find jobs with higher-paying carriers (Swartz et al., 2017) to reduce costs (Miller, Saldanha, et al., 2018a). Social resident pathogens include the rationalized myth that an available truck with no driver is wasted capacity (Hubbard, 2003), and scripts justifying driving while ill, unlicensed, or fatigued.

Unfit driving increases driver stress, due to fear of losing their livelihood or qualifying credentials. If an error occurs for a carrier with stressed, fatigued, or poorly qualified drivers, adverse consequences are more likely to result. On the other hand, policies and norms ensuring a carrier's drivers are qualified, licensed, and experienced form defensive layers that prevent resident pathogens from intensifying the relationship between an error and adverse consequences. Thus, when a dispatcher or driver commits an error, the likelihood of adverse consequences is greater if the carrier's latent conditions enable or tolerate unfit drivers.

Hypothesis 5 (H5). A carrier's record of driver fitness violations moderates the likelihood that an active failure by a (a) dispatcher or (b) driver results in delayed delivery, such that the likelihood of delayed delivery will be enhanced for shipments by carriers with more driver fitness violations.

3.3 | Moderating effect of latent conditions on delay length

The length of delivery delay is an important supply chain outcome in the same way a patient's length of stay is an important healthcare outcome. In a tightly coupled supply chain, every minute of delivery delay translates into revenue loss. When shifting defensive layers contain more holes, the holes align more often, and the severity of adverse consequences increases. By analogy, if a patient with weak defensive layers is exposed to a virus, the patient is not only more likely to contract an illness (analogous to likelihood of delay), he or she is also likely to experience more severe symptoms (analogous to length of delay). In addition to the patient's immune system, other defensive layers include the patient's overall health, access to medical care, socioeconomic status, and population density. The weaker the defensive layers, the more holes each layer has (Reason, 2000) and the sicker the patient will become if exposed to a virus. Similarly, we expect that a carrier's latent conditions moderate the relationship between errors and delay time.

We hypothesize that dispatcher errors are associated with longer delivery delays than driver errors. Dispatcher errors occur earlier in the process and can have a cumulative effect; a driver may already be dealing with a delay when pulling away from the loading dock. Like the bullwhip effect, the magnitude of the delay grows as its source is farther from the impact.

Hypothesis 6a (H6a). A carrier's record of unsafe driving violations moderates the relationship between an active failure and delivery delay, such that the difference in delay between dispatcher and driver active failures will be enhanced for shipments transported by carriers with greater unsafe driving violations.

Hypothesis 6b (H6b). A carrier's record of hours-of-service violations moderates the relationship between an active failure and delivery delay, such that the difference in delay between dispatcher and driver active failures will be enhanced for shipments transported by carriers with greater hours-of-service violations.

Hypothesis 6c (H6c). A carrier's record of vehicle maintenance violations moderates the relationship between an active failure and delivery delay, such that the difference in delay between dispatcher and driver active failures will be enhanced for shipments transported by carriers with greater vehicle maintenance violations.

Hypothesis 6d (H6d). A carrier's record of driver fitness violations moderates the relationship between an active failure and delivery delay, such that the difference in delay between dispatcher and driver active failures will be enhanced for shipments transported by carriers with greater driver fitness violations.

4 | METHOD

4.1 | Data sources

4.1.1 | Level 1

Shipment-level data were collected from the TMS of a Fortune 500 company that we call LIM. It produces home furnishings, appliances, consumer electronics, and housewares, with annual revenues of more than $20 billion; LIM's freight and warehousing expenses comprise 16% of revenue. Prior to our data collection, LIM had invested over $1 billion in advanced systems and technology from leading vendors and installed next-generation GPS tracking technology on every truck, updated every 30 s. However, despite substantial investments in technology defensive layers, some shipments continued to be lost, delayed, or routed to incorrect locations. Ten to fifteen LIM employees support the ~600 shipments active at any given time; resolving a problem can require 3–4 h, and our discussions with LIM supervisors indicated that dispatcher interventions sometimes exacerbate problems. We define a shipment as an entire 53-foot trailer filled with LIM products, packed full cube. We worked closely with 11 LIM managers, highly familiar with LIM's flow of products, information, customers, and logistics, to obtain data on its shipments from 2014 to 2016, accurate to the second. We restricted the data to carriers that transported 100 or more of LIM's shipments during that time, to avoid the potential for skewed violations data (Miller, Saldanha, et al., 2018a). This reduced the sample size by only 399 shipments (~0.1% of the sample), indicating LIM used very few spot market carriers. The final sample consisted of 299,399 shipments conveyed by 97 carriers. The data was checked for downloading errors by LIM's IT personnel. To further ensure data integrity, a member of the research team and a LIM employee individually checked the downloaded data, then jointly examined it.

4.1.2 | Level 2

Data on the artifacts of carrier latent conditions (unsafe driving, HOS, vehicle maintenance, and driver fitness violations) was collected for LIM's carriers from the Federal Motor Carrier Safety Administration's (FMCSA) Compliance, Safety, and Accountability (CSA) program dataset, which rates carriers in safety categories known as BASICs (Behavior Analysis and Safety Improvement Categories) (Miller, Schwieterman, & Bolumole, 2018b). A carrier's CSA score for a BASIC is the sum of its violations, as ticketed or reported from roadside inspections, weighted by severity and recency, yielding a 0–30 score. A CSA score of 0 = 0th percentile (best possible score), 2.31 = 25th percentile, 4.72 = 50th percentile, 8.17 = 75th percentile, and 30 = 100th percentile (worst possible score). CSA scores are specific to a BASIC, since they are calculated as percentiles, thus they cannot be compared or combined with scores from other BASICs (Federal Motor Carrier Safety Administration, 2019).

4.2 | Dependent variables

The dependent variables were based on LIM's definition of on-time delivery. A third-party data analytics firm calculates, for LIM, the scheduled delivery time for each shipment as the driving time between the origin–destination city pair, based on speed limits, mandated break times, and a buffer. The scheduled delivery times are agreed to by LIM and the selected carrier. Upon delivery, actual arrival times are recorded in LIM's TMS. Shipments arriving more than 15 min past the scheduled delivery time are designated as late. *PROB[LATE]* is a binary variable where 0 = on time and 1 = late. *PROB[LATE]* was the dependent variable for testing H1–H5 (n = 299,399), which assess the likelihood a shipment will be delivered late. We calculated *DELAY* for the subset of the analytic dataset comprised of only the late deliveries (n = 16,293) (H6a–H6d) as the difference between the scheduled delivery time and the actual delivery time.

4.3 | Independent variables

Table 3 summarizes the variable operationalization, and Table 4 provides descriptive statistics.

4.3.1 | Active failures

Upon arrival at the destination, the driver, dock manager, LIM's arrival manager, and dispatcher assign an arrival code (previously listed in Table 2) to indicate issues with the shipment or "normal delivery—no issues." Because it is impossible to measure most errors directly, since they are not reported and the driver or dispatcher may not even be aware of them, we operationalized errors as their immediate effects, which are reported as shipment arrival codes. For H1–H5, *Dispatcher Error* is a binary variable with a value of 1 if an arrival code is between 2 and 10 and a value of 0 otherwise (see Table 2). *Driver Error* is a binary variable with a value of 1 if an arrival code of 11–17 was selected and a value of 0 otherwise. *Error Type* (H6a–H6d) is a binary variable with a value of 1 for arrival codes 2–10 (dispatcher errors) and 0 for arrival codes 11–17 (driver errors).

4.3.2 | Artifacts of carrier latent conditions

The artifacts of latent conditions were operationalized as the carrier's percentile rank for each BASIC during the time period of the shipment data. *Unsafe driving violations* are defined by the CSA program as operating a

TABLE 3 Operationalization of variables

Variable	Operationalization	Full sample (H1–H5)			Late shipments subsample (H6a–H6d)		
		Mean	SD	n	Mean	SD	n
Level one: Shipments							
Active failures							
Dispatcher error	0 = no dispatcher error 1 = dispatcher error	0.04[a]		299,399			
Driver error	0 = no driver error 1 = driver error	0.04		299,399			
Error type	0 = driver error 1 = dispatcher error				0.58		16,293
Dependent variables							
PROB[LATE]	0 = on time 1 = late	0.05		299,399			
DELAY	Difference between the scheduled and actual delivery time (minutes)				754.93	1791.53	16,293
Level two: Carriers							
Latent conditions							
Unsafe driving violations	BASIC percentile rank (0–30)	1.80[b]	1.78	97	1.80	1.78	97
Hours-of-service violations	BASIC percentile rank (0–30)	0.44	0.68	97	0.44	0.68	97
Vehicle maintenance violations	BASIC percentile rank (0–30)	3.54	2.90	97	3.54	2.90	97
Driver fitness violations	BASIC percentile rank (0–30)	0.07	0.11	97	0.07	0.11	97

[a]Because standard deviations are not meaningful for binary variables, they are not reported.
[b]The latent conditions values are the same in the full sample and late shipments subsample because they are measured at the carrier level, not the shipments level.

TABLE 4 Descriptive statistics

a. Shipments

	Shipments with errors		Shipments delivered late		% Shipments with errors delivered late
	n	%	n	%	
Dispatcher errors	11,349	3.8%	9,482	3.2%	83.5%
Driver errors	12,845	4.2%	6,811	2.3%	53.0%
All shipments	24,194	8.1%	16,293	5.4%	67.3%

b. Carriers

Measure	Mean	SD
Number of tractors owned	812.13	2,306.19
Number of trailers owned	1,999.39	5,161.18
Number of tractors leased	256.38	739.03
Number of trailers leased	1,594.47	12,914.49
Number of vehicles	1,246.00	3,558.67
Annual vehicle miles	80,815,333.00	182,293,731
Number of drivers	1,175.41	2,789.53

vehicle in a dangerous or careless manner, including speeding, reckless driving, improper lane changes, texting while driving, not wearing seatbelts, and operating in adverse weather conditions. *HOS violations* are defined as operation of a vehicle by a driver who is fatigued, ill, in noncompliance with HOS regulations, or who does not retain records of duty status for 6 months. *Vehicle maintenance violations* include inoperative brakes, lights, and other mechanical defects, improper cargo securement, failure to make required repairs, failure to prevent shifting cargo, spilled or dropped cargo, and overloading. *Driver fitness violations* are defined as truck operation by drivers who are unfit due to lack of training, experience, or medical qualifications, including failing to have a valid commercial driver's license, and failing to maintain driver qualification files (FMCSA, 2019).

4.4 | Control variables

The standardized values of three control variables were included. *Carrier ownership ratio* is the ratio of vehicles a carrier owns versus leases, which can impact how well the carrier maintains its fleet; a carrier concerned about resale value may invest more in maintenance (Scott & Nyaga, 2019). *Carrier annual vehicle miles* is the number of miles a carrier's vehicles are driven annually, to control for differences between carriers in the distance shipments are transported. *Carrier number of drivers* is the number of drivers per carrier; it controls for carrier size, which is related to resources reflecting a carrier's ability to invest in technology (Cantor et al., 2009; Cantor et al., 2016; Manrodt et al., 2003), visibility to monitoring agencies (Miller, Golicic, & Fugate, 2017a), and frequency of violations (Scott et al., 2021). Additional controls employed in the robustness checks are described below.

4.5 | Models and analysis

The hypotheses were tested using maximum likelihood two-level mixed effects regression analysis, where H1–H5 used generalized mixed effects models and H6 used linear mixed effects models. Mixed effects modeling accounts for the clustered nature of the data (Bliese et al., 2018; Raudenbush & Bryk, 2002). Because it is particularly well-suited to cross-level interactions between imbalanced clusters (McNeish & Kelley, 2018), mixed effects modeling is ideal for our sample and hypotheses. Multilevel samples should have a minimum of 30 observations at each level to ensure statistical power. Level 1 in our analysis is shipments (n = 299,299 in the full sample, n = 16,293 in the late shipments subset), which are nested within carriers (Level 2, n = 97 in the late shipments subset). The *lme4* (Bates et al., 2015) and *multilevel* (Bliese, 2016) packages in *R for Windows 3.6.0* were used for analysis.

Prior to hypothesis testing, we ensured that model specification was correct (Bliese et al., 2018). To establish the two-level models, we calculated the intra-class correlation coefficients (ICC[1]) for each dependent variable (0.14 for *PROB[LATE]*, 0.12 for *DELAY*), which indicate 14% of the variance in *PROB[LATE]* in the full dataset is attributable to carrier-level effects and 12% of the variance in *DELAY* in the late shipments subset is attributable to carrier-level effects. These values are acceptable, following conventions for multilevel empirical models (Bliese et al., 2018). Next, we examined whether the Level 1 intercepts and slopes varied randomly by estimating a fixed effects, random intercept-fixed slope, and random intercept-random slope model for each dependent and independent variable combination (three sets of models). In each case, the random intercept-random slope model was the best fitting model,[3] thus, all subsequent models include both random effects. As an illustration, using Raudenbush and Bryk's (2002) notation, the equations for *PROB[LATE]* predicted by dispatcher error[4] are:

$$PROB[LATE]_{ij} = \beta_{0j} + \beta_{1j} Dispatcher\ Error_{ij} + r_{ij} \tag{1a}$$

$$\beta_{0j} = \gamma_{00} + \gamma_{01} Violations_j + u_{0j} \tag{1b}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11} Violations_j + u_{1j} \tag{1c}$$

$$\begin{aligned} PROB[LATE]_{ij} = \gamma_{00} + \gamma_{01} Dispatcher\ Error_{ij} + \gamma_{10} \\ + \gamma_{11}\left(Violations_j \times Dispatcher\ Error_{ij}\right) \\ + \mu_{1j} Dispatcher\ Error_{ij} + u_{0j+} r_{ij} \end{aligned} \tag{2}$$

Similar equations were used for *PROB[LATE]* predicted by driver error and for *DELAY* predicted by error type. H1–H5 were tested using generalized linear mixed effects regression (logit) analysis because of the binary dependent variable (*PROB[LATE]*), while H6 was tested using linear mixed effects regression analysis. Table 5 contains the bivariate correlation coefficients. Predicted interaction plots were used to interpret significant cross-level interactions.

5 | RESULTS

5.1 | Main effects

The mixed-effects regression models examined the main effects of active failures by dispatchers and drivers on the shipper's loads, nested within carriers, moderated by carrier records of the four types of violations. The models of the main effects for the likelihood that a shipment with a dispatcher error or driver error will be delivered late are presented in Table 6. H1a, which predicted the presence of a dispatcher error will increase the likelihood of late delivery, was tested as the main effect of dispatcher error on *PROB[LATE]*. As reported in Model 1 (Table 6), this effect was positive and significant ($\gamma_{10} = 6.06$, $p < .001$). Deliveries associated with a dispatcher error are significantly more likely to be late than those not associated with a dispatcher error. Similarly, H1b predicted that the presence of a driver error would increase *PROB[LATE]*. Model 2 indicates this effect was also positive and significant ($\gamma_{20} = 4.29$, $p < .001$). Deliveries associated with driver errors are significantly more likely to be late. Thus, H1a and H1b were supported.

5.2 | *PROB[LATE]* interactions

H2–H5 tested the cross-level interaction effects, predicting the carrier's record of violations for unsafe driving, HOS, vehicle maintenance, and driver fitness violations would exacerbate the likelihood that a dispatcher error will result in a late delivery, and the likelihood that a driver error will result in a late delivery. Tables 7 and 8 contain the cross-level interaction models.

TABLE 5 Correlation coefficients

a. Level 1, Full sample (n = 299,399)

	1	2	3	4	5	6	7	8	9	10
1. PROB[LATE]										
2. Dispatcher Error	0.68[a]									
3. Driver Error	0.44[a]	−0.04[a]								
4. Unsafe Driving Violations	0.01[a]	−0.01[a]	0.00[a]							
5. HOS Violations	0.00	−0.03[a]	0.04[a]	0.73[a]						
6. Vehicle Maintenance Violations	0.03[a]	0.00	0.11[a]	0.22[a]	0.58[a]					
7. Driver Fitness Violations	−0.01[a]	−0.02[a]	−0.01[a]	0.41[a]	0.29[a]	0.07[a]				
8. Carrier Vehicle Ownership Ratio	0.02[a]	0.04[a]	−0.04[a]	0.20[a]	0.03[a]	−0.06[a]	0.16[a]			
9. Carrier Annual Vehicle Miles	0.00	0.07[a]	−0.04[a]	−0.13[a]	−0.24[a]	−0.18[a]	−0.18[a]	0.20[a]		
10. Carrier Number of Drivers	−0.01[a]	0.06[a]	−0.03[a]	−0.18[a]	−0.23[a]	−0.21[a]	−0.15[a]	0.10[a]	0.87[a]	

b: Level 1, Subsample of late shipments (n = 16,923)

	1	2	3	4	5	6	7	8	9
1. DELAY									
2. Error Type	0.13[a]								
3. Unsafe Driving Violations	0.02[a]	−0.07[a]							
4. HOS Violations	0.04[a]	−0.13[a]	0.50[a]						
5. Vehicle Maintenance Violations	0.02[a]	−0.11[a]	0.11[a]	0.67[a]					
6. Driver Fitness Violations	−0.02[a]	−0.02[a]	0.36[a]	0.18[a]	0.04[a]				
7. Carrier Vehicle Ownership Ratio	−0.04[a]	0.07[a]	0.13[a]	−0.18[a]	−0.19[a]	0.17[a]			
8. Carrier Annual Vehicle Miles	0.08[a]	0.11[a]	−0.20[a]	−0.29[a]	−0.22[a]	−0.19[a]	0.17[a]		
9. Carrier Number of Drivers	0.05[a]	0.07[a]	−0.27[a]	−0.27[a]	−0.23[a]	−0.15[a]	0.08[a]	0.87[a]	

c. Level 2 (n = 97)

	1	2	3	4
1. Unsafe Driving Violations				
2. HOS Violations	0.67[a]			
3. Vehicle Maintenance Violations	0.24[a]	0.58[a]		
4. Driver Fitness Violations	0.34[a]	0.42[a]	0.13	

[a]$p < .05$, *Error Type* = 1 for dispatcher error, 0 for driver error.

TABLE 6 *PROB[LATE]* due to dispatcher or driver error: Mixed effects logistic regression results

Level and variable		Model 1	Model 2
Level 1	Intercept (γ_{00})	3.69*** (0.10)	−3.22*** (0.07)
H1a	Dispatcher Error (γ_{10})	6.06*** (0.11)	
H1b	Driver Error (γ_{20})		4.29*** (0.11)
Level 2 Controls			
	Motor Carrier Vehicle Ownership Ratio (γ_{01})	0.05 (0.08)	0.00 (0.07)
	Motor Carrier Annual Vehicle Miles (γ_{02})	−0.04 (0.15)	0.33** (0.10)
	Motor Carrier Number of Drivers (γ_{03})	−0.23 (0.15)	−0.31** (0.10)
Variance components			
	Within-carrier ($L1$) variance (σ^2)	3.29	3.29
	Intercept ($L2$) variance (τ_{00})	0.78	0.58
	Slope ($L2$) variance (τ_{11})	0.84	2.01
	Intercept-slope ($L2$) covariance (τ_{01})	−0.56	−0.38
Additional information			
	−2 log likelihood (ML)	−34,253.30	−46,829.10
	Conditional R^2	0.33	0.15

Note: All models used motor carriers with over 100 observations only; n = 299,399, k = 97. *Dispatcher Error* coded 1 for error and 0 for no error. *PROB[LATE]* ICC(1) = 0.14, random effect: *Dispatcher Error Centered* (γ_{10}), dependent variable; 0 = on time, 1 = late. $^{\dagger}p < .10$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

In each table, Model 1 contains the full model for main effects only, including un-hypothesized main effects for the fixed effects variables, and subsequent models test the hypothesized cross-level interaction effects. Model 8 contains the full model, including both main effects and cross-level interaction effects. Table 7 summarizes the results of the tests for moderating effects of artifacts of carrier latent conditions on the relationship between dispatcher errors and likelihood of late delivery. The cross-level interaction terms for unsafe driving (Model 4) and vehicle maintenance (Model 6) violations were positive, but not significant, thus, H2a and H4a were not supported. There were significant, positive cross-level interactions between dispatcher errors and the carrier's record of HOS (Model 2, $\gamma_{12} = 0.42$, $p < .01$) and driver fitness violations (Model 3, $\gamma_{14} = 0.87$, $p < .001$). Figure 4 illustrates these effects, such that the relationship between a dispatcher error and the probability of a late delivery is greater for carriers with a record of more HOS violations. While unfit driver violations has a similar positive interaction effect, there is also a strong negative level 2 main effect. Thus, H3a was supported and H5a was marginally supported.

Table 8 contains the results of the tests for moderating effects of artifacts of carrier latent conditions on the relationship between driver errors and the likelihood of late delivery. The cross-level interactions between a driver error and the carrier's record of unsafe driving (Model 2) and vehicle maintenance violations (Model 4) were not significant. Thus, H2b and H4b were not supported. The cross-level interaction between a driver error and the carrier's record of HOS violations was in the expected direction and was significant (Model 3, $\gamma_{12} = 0.24$, $p < .10$), supporting H3b. There was a significant cross-level interaction between a driver error and driver fitness violations (Model 5, $\gamma_{14} = 1.89$, $p < .001$), supporting H5b. Figure 5 illustrates these effects.

Our results also reveal some intriguing additional effects that were not hypothesized. When considering errors by either dispatchers (Model 1, Table 7) or drivers (Model 1, Table 8) as main effects, driver fitness violations are a significant negative predictor of the carrier's average *PROB[LATE]*; none of the other violations main effects was significant. Thus, only driver fitness violations were related to reduced carrier average likelihood of a delivery delay. Further, while the test of H4b is not significant (Model 4, Table 8), the cross-level interaction between driver active failures and carrier vehicle maintenance violations is significant and negative when considering all the cross-level interactions simultaneously (Model 8, Table 8).

5.3 | *DELAY* interactions

Table 9 contains the results of the moderation tests for artifacts of a carrier's latent conditions on the relationship between dispatcher vs. driver errors (*Error Type*) and

TABLE 7 *PROB[LATE]* due to dispatcher error: Generalized linear mixed effects regression results

Level and variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Level 1								
Intercept (γ_{00})	−3.82*** (0.13)	−3.76*** (0.13)	−3.77*** (0.13)	−3.79*** (0.13)	−3.80*** (0.13)	−3.71*** (0.15)	−3.73*** (0.015)	−3.82*** (0.13)
Dispatcher Error (γ_{10})	6.06*** (0.11)	5.88*** (0.14)	5.89*** (0.11)	5.99*** (0.14)	6.00*** (0.10)	5.89*** (0.12)	5.99*** (0.11)	6.00*** (0.21)
Level 2								
Unsafe Driving Violations (γ_{01})	0.06 (0.06)	0.03 (0.06)	0.07 (0.05)	0.06 (0.0.6)	0.06 (0.06)	0.07 (0.06)	0.06 (0.06)	0.08 (0.07)
HOS Violations (γ_{02})	−0.11 (0.16)	−0.10 (0.14)	−0.241 (0.14)	−0.11 (0.15)	−0.10 (0.16)	−0.26 (0.17)	−0.12 (0.19)	−0.29 (0.22)
Vehicle Maintenance Violations (γ_{03})	0.05 (0.03)	0.04 (0.03)	0.04 (0.03)	0.04 (0.03)	0.04 (0.03)	0.05 (0.03)	0.05 (0.03)	0.06† (0.03)
Driver Fitness Violations (γ_{04})	−1.22*** (0.33)	−1.27*** (0.35)	−1.22*** (0.31)	−1.23*** (0.29)	−1.59*** (0.21)	−1.25*** (0.31)	−1.63*** (0.45)	−1.21** (0.39)
Cross-level interactions								
H2a—Dispatcher Error × Unsafe Driver Violations (γ_{11})		0.11 (0.07)						−0.01 (0.10)
H3a—Dispatcher Error × HOS Violations (γ_{12})			0.42** (0.15)			0.42** (0.16)		0.58 (0.36)
H4a—Dispatcher Error × Vehicle Maintenance Violations (γ_{13})				0.02 (0.03)				−0.05 (0.05)
H5a—Dispatcher Error × Driver Fitness Violations (γ_{14})					0.87*** (0.21)		0.94* (0.39)	0.03 (0.57)
Level 1 controls								
Quarter Dummy $1_{\text{Q2 vs. Q1}}$ (γ_{20})						0.17*** (0.03)	0.17*** (0.03)	
Quarter Dummy $2_{\text{Q3 vs. Q1}}$ (γ_{30})						0.14*** (0.03)	0.14*** (0.03)	
Quarter Dummy $3_{\text{Q4 vs. Q1}}$ (γ_{40})						−0.01 (0.03)	−0.01 (0.03)	
Year Dummy $1_{\text{2015 vs. 2014}}$ (γ_{50})						−0.05† (0.03)	−0.05† (0.03)	
Year Dummy $2_{\text{2016 vs. 2014}}$ (γ_{60})						−0.40*** (0.03)	−0.40*** (0.03)	
Level 2 controls								
Motor Carrier Vehicle Ownership Ratio (γ_{05})	0.07 (0.08)	0.07 (0.08)	0.07 (0.08)	0.07 (0.08)	0.07 (0.08)	0.07 (0.08)	0.07 (0.08)	0.07 (0.08)
Motor Carrier Annual Vehicle Miles (γ_{06})	−0.11 (0.14)	−0.11 (0.14)	−0.11 (0.14)	−0.11 (0.13)	−0.10 (0.14)	−0.10 (0.16)	−0.10 (0.16)	−0.10 (0.14)
Motor Carrier Number of Drivers (γ_{07})	−0.16 (0 .14)	−0.15 (0.14)	−0.15 (0.14)	−0.15 (0.13)	−0.15 (0.14)	−0.16 (0.17)	−0.16 (0.18)	−0.15 (0.15)

TABLE 7 (Continued)

Level and variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Variance Components								
Within-carrier ($L1$) variance (σ^2)	3.29	3.29	3.29	3.29	3.29	3.29	3.29	3.29
Intercept ($L2$) variance (τ_{00})	0.74	0.73	0.73	0.74	0.74	0.70	0.71	0.73
Slope ($L2$) variance (τ_{11})	0.85	0.82	0.80	0.85	0.84	0.82	0.85	0.79
Intercept-slope ($L2$) covariance (τ_{01})	−0.55	−0.55	−0.55	−0.55	−0.55	−0.54	−0.54	−0.55
Additional information								
−2 log likelihood (ML)	−34,251.00	−34,250.00	−34,248.60	−34,250.80	−34,250.60	−34,132.70	−34,134.70	−34,248.10
Conditional R^2	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33

Note: All models used motor carriers with over 100 observations only; n = 299,399, k = 97. *Dispatcher Error* coded 1 for Error and 0 for no Error. *PROB[LATE]* ICC (1) = 0.14, random effect: *Dispatcher Error* (γ_{10}), dependent variable; 0 = on time, 1 = late. $^{\dagger}p < .10$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

the length of time for which a late delivery is delayed, H6a–d. *Error Type* was significant and positive (Model 1, $\gamma_{10} = 517.42$, $p < .001$), thus, the average length of delay for late deliveries was greater for dispatcher errors. The cross-level interactions between *Error Type* and the carrier's record of unsafe driving (Model 2) and driver fitness violations (Model 5) were not significant, thus, H6a and H6d were not supported. There were significant cross-level interactions between *Error Type* and the carrier's record of HOS (Model 3, $\gamma_{12} = 197.50$, $p < .05$) and vehicle maintenance violations (Model 4, $\gamma_{13} = 37.92$, $p < .05$), supporting H6b and H6c. Figure 6 illustrates that the difference in length of delay for dispatcher vs. driver errors was greater for shipments transported by carriers with a record of more HOS and vehicle maintenance violations, respectively.

5.4 | Robustness checks

We conducted additional analyses to test the robustness of our findings. In addition to errors and carriers' latent conditions, it is possible that factors such as weather and seasonal traffic patterns could contribute to the probability of a late delivery. Therefore, our first set of robustness checks added Level 1 fixed effects to control for time of year and calendar year in all significant cross-level interaction models. We created dummy-coded fixed effects by quarter (where January–March represents Q1) and year (2014–2016). The Level 1 calendar effects indicate differences across both seasons and calendar years in the likelihood of late delivery. Models 6 and 7 (Table 7) replicate the significant results for H3a and H5a, above and beyond the calendar effects. Likewise, Model 7 (Table 7) replicates the significant result for H5b. Finally, Models 6 and 7 (Table 8) replicate the significant results for H6b and H6c above and beyond the calendar effects. Model 7 (Table 8) failed to replicate the significant result for H3b, thus, the moderation effect of HOS carrier latent conditions in exacerbating the likelihood of a delivery delay following a driver error was not robust to calendar effects.

Next, to test for Level 2 (carrier level) endogeneity we replicated all the hypothesis tests using group mean-centered independent variables (Bliese et al., 2018) in Tables 10–12. While mixed effects models are the most appropriate approach for the nesting of our data, it is also important to test for Level 2 (carrier level) endogeneity. Thus, as a robustness check, we replicated all analyses using group mean-centered independent variables while controlling for the predictors' group means (Bliese et al., 2018), combining the benefits of mixed effects models with protection against the endogeneity of fixed

TABLE 8 *PROB[LATE]* due to driver error: Generalized linear mixed effects regression results

Level and variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Level 1								
Intercept (γ_{00})	−3.14*** (0.11)	−3.12*** (0.10)	−3.12*** (0.11)	−3.16*** (0.13)	−3.12*** (0.14)	−3.07*** (0.13)	−3.06*** (0.13)	−3.16*** (0.11)
Driver Error (γ_{10})	4.30*** (0.12)	4.18*** (0.13)	4.20*** (0.13)	4.43*** (0.17)	4.16*** (0.14)	4.19*** (0.16)	4.15*** (0.14)	4.44*** (0.14)
Level 2								
Unsafe Driving Violations (γ_{01})	−0.03 (0.05)	−0.05 (0.05)	−0.04 (0.05)	−0.03 (0.05)	−0.04 (0.05)	−0.04 (0.05)	−0.04 (0.05)	−0.03 (0.05)
HOS Violations (γ_{02})	−0.09 (0.12)	−0.09 (0.12)	−0.13 (0.12)	−0.09 (0.13)	−0.09 (0.13)	−0.14 (0.16)	−0.10 (0.16)	−0.17 (0.12)
Vehicle Maintenance Violations (γ_{03})	0.02 (0.03)	0.02 (0.03)	0.02 (0.03)	0.03 (0.03)	0.02 (0.03)	0.03 (0.03)	0.03 (0.03)	0.04 (0.03)
Driver Fitness Violations (γ_{04})	−0.94*** (0.18)	−0.94*** (0.16)	−0.94*** (0.119)	−0.94*** (0.28)	−1.27*** (0.32)	−0.99*** (0.28)	−1.33*** (0.35)	−1.16*** (0.19)
Cross-level interactions								
H2b—Driver Error × Unsafe Driver Violations (γ_{11})		0.07 (0.06)						−0.06 (0.08)
H3b—Driver Error × HOS Violations (γ_{12})			0.24† (0.13)			0.24 (0.22)		0.51*** (0.14)
H4b—Driver Error × Vehicle Maintenance Violations (γ_{13})				−0.04 (0.04)				−0.10* (0.04)
H5b—Driver Error × Driver Fitness Violations (γ_{14})					1.89*** (0.25)		2.00*** (0.27)	1.25*** (0.18)
Level 1 controls								
Quarter Dummy $1_{\mathrm{Q2\ vs.\ Q1}}$ (γ_{20})						0.09** (0.03)	0.09** (0.03)	
Quarter Dummy $2_{\mathrm{Q3\ vs.\ Q1}}$ (γ_{30})						0.08* (0.03)	0.08* (0.03)	
Quarter Dummy $3_{\mathrm{Q4\ vs.\ Q1}}$ (γ_{40})						0.05* (0.03)	0.05* (0.03)	
Year Dummy $1_{2015\ \mathrm{vs.}\ 2014}$ (γ_{50})						0.03 (0.02)	0.03 (0.02)	
Year Dummy $2_{2016\ \mathrm{vs.}\ 2014}$ (γ_{60})						−0.38*** (0.03)	−0.38*** (0.03)	
Level 2 controls								
Motor Carrier Vehicle Ownership Ratio (γ_{05})	0.03 (0.07)	0.03 (0.07)	0.03 (0.07)	0.03 (0.07)	0.03 (0.07)	0.03 (0.07)	0.03 (0.07)	0.03 (0.07)
Motor Carrier Annual Vehicle Miles (γ_{06})	0.29* (0.11)	0.29** (0.11)	0.29** (0.11)	0.29* (0.12)	0.29* (0.14)	0.30* (0.15)	0.30* (0.15)	0.29** (0.11)
Motor Carrier Number of Drivers (γ_{07})	−0.30* (0.12)	−0.30** (0.10)	−0.30** (0.11)	−0.30* (0.13)	−0.30* (0.14)	−0.32* (0.15)	−0.32* (0.15)	−0.30** (0.11)

TABLE 8 (Continued)

Level and variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Variance components								
Within-carrier (L1) variance (σ^2)	3.29	3.29	3.29	3.29	3.29	3.29	3.29	3.29
Intercept (L2) variance (τ_{00})	0.53	0.53	0.53	0.53	0.53	0.52	0.52	0.53
Slope (L2) variance (τ_{11})	2.01	1.99	1.98	2.01	1.96	2.01	1.99	1.91
Intercept-slope (L2) covariance (τ_{01})	−0.35	−0.34	−0.34	−0.35	−0.34	−0.34	−0.34	−0.34
Additional information								
−2 log likelihood (ML)	−46,826.70	−46,826.40	−46,826.10	−46,826.40	−42,825.80	−46,669.40	−34,134.70	−46,824.50
Conditional R^2	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15

Note: All models used motor carriers with over 100 observations only; n = 299,399, k = 97. *Dispatcher Error* coded 1 for Error and 0 for no Error. *PROB[LATE]* ICC (1) = 0.14, random effect: *Driver Error* (γ_{10}), dependent variable; 0 = on time, 1 = late. $^{\dagger}p < .10$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

effects modeling. This approach represents the "gold standard" for estimating cross-level interaction effects (Aguinis et al., 2003; Antonakis et al., 2019; Bell & Jones, 2015; Bliese et al., 2020; Certo et al., 2017; McNeish & Kelley, 2018). We centered (demeaned) the independent variables based on the group mean (e.g., $Dispatcher\ Error_{ij} - \overline{Dispatcher\ Error_j}$) for each observation while also adding an additional fixed effect control for the group mean (e.g., $\overline{Dispatcher\ Error_j}$). Therefore, in the robustness checks, the independent variable for the *PROB[LATE]* models was the presence of a dispatcher error relative to the carrier's proportion of shipments with dispatcher errors. Similarly, the independent variable for the driver error models was the presence of a driver error relative to the carrier's proportion of shipments with driver errors. The independent variable for the *DELAY* models was the presence of a dispatcher error relative to the carrier's ratio of dispatcher to driver errors. These models used the group mean centered predictors and added a fixed effect for the group mean of the original predictor from the hypothesis test (i.e., group average rate of dispatcher error). Tables 10–12 correspond to Table 7–9, where all hypothesized interactions were examined, and the significant interactions were tested again using the calendar fixed effects. In Table 10, Models 3 and 6 and Models 5 and 7 provide robust support for the moderating effects of HOS violations and driver fitness violations in H3a and H5a, respectively. Likewise, in Table 11 Model 3 and Models 5 and 7 provide robust support for the moderating effects of HOS violations and driver fitness violations in H3b and H5b, respectively. Similar to the results using time-based fixed effects above, Model 6 (Table 11) failed to support the moderating effect of HOS violations above and beyond the calendar effects. Finally, in Table 12, Models 4 and 7 and Models 5 and 8 provide robust support for the moderating effects of HOS violations and vehicle maintenance violations in H6b and H6c, respectively. In sum, using the "gold standard" of group mean centering provided robust support for the results of our hypothesis tests, while also accounting for potential Level 2 endogeneity.

The online supplement contains the results of further robustness analyses of our hypothesized models. Tables S1–S3 show corresponding cross-level interaction models that only control for the specific Level 2 moderator, to address the potential for multicollinearity between the artifacts of carrier latent conditions. Tables S4–S6 show corresponding cross-level interaction models for our significant hypothesis tests that add a control for standardized geospatial distance. Tables S7–S9 contain estimated models for our significant results using a penalized maximum

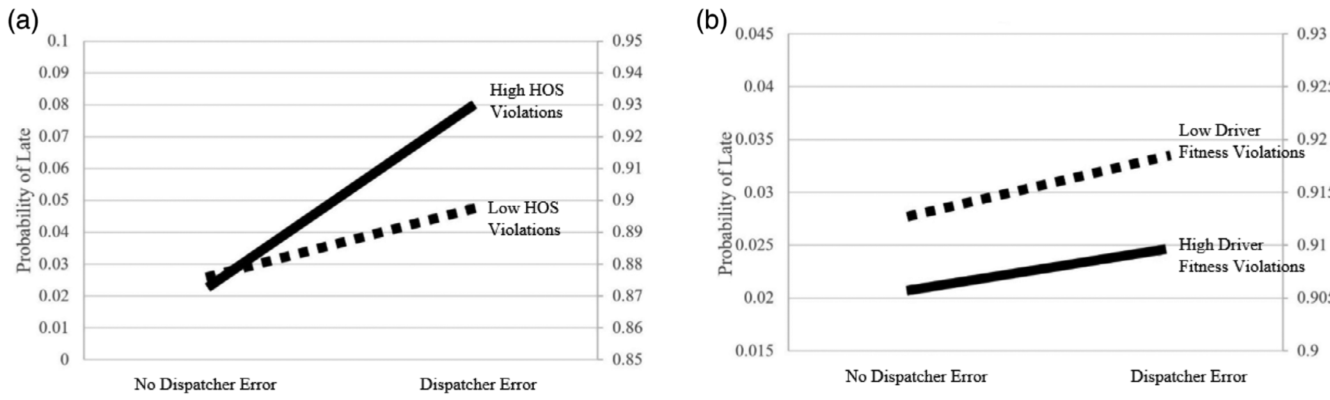

FIGURE 4 Predicted plots for cross-level interactions of carrier violations with dispatcher errors, with HOS and driver fitness violations predicting the probability of late delivery

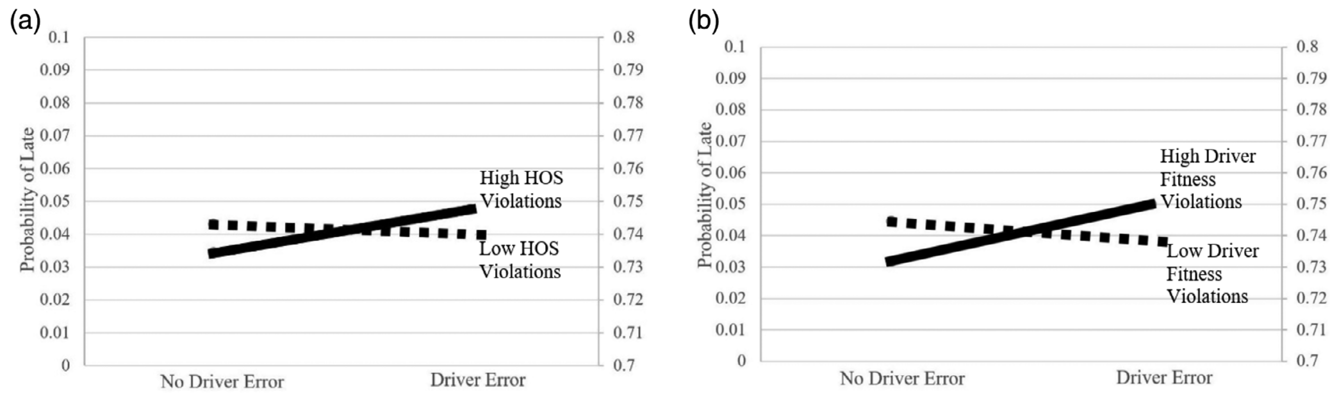

FIGURE 5 Predicted plots for cross-level interactions of carrier violations with driver errors, with HOS and driver fitness violations predicting the probability of late delivery

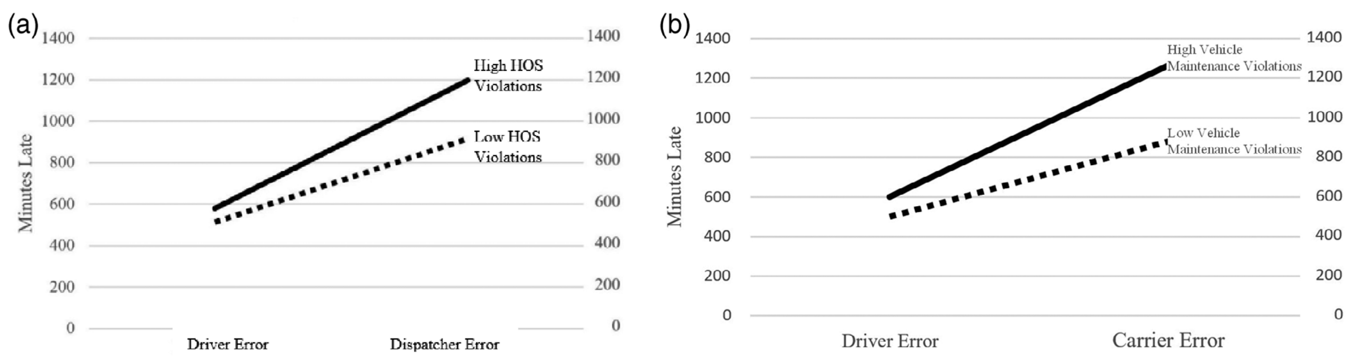

FIGURE 6 Predicted plot for cross-level interaction of shipment-level error type with carrier violations predicting the degree of DELAY

likelihood Bayesian estimation, which Browne and Draper (2006) suggest has advantages over likelihood-based methods, in terms of unbiasedness and efficiency. Generally, each set of robustness analysis findings mirror the results of our hypothesized models, further supporting their robustness.

6 | DISCUSSION

6.1 | Key findings

Errors should be expected in complex systems (Reason, 2000); indeed, 5.44% of LIM's deliveries (over

TABLE 9 *DELAY* due to error type: Linear mixed effects regression results

Level and variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Level 1								
Intercept (γ_{00})	510.54*** (82.18)	510.92*** (82.10)	497.87*** (83.01)	470.60*** (84.32)	510.13 (82.61)	439.74*** (89.10)	411.95*** (90.38)	487.49*** (84.63)
Error Type (Dispatcher $_{vs.}$ Driver Error) (γ_{10})	517.42*** (59.47)	530.87*** (91.15)	433.76*** (70.73)	380.86*** (89.30)	518.68*** (72.35)	436.86*** (71.15)	383.74*** (89.88)	504.51*** (118.84)
Level 2								
Unsafe Driving Violations (γ_{01})	28.02 (38.63)	27.81 (38.60)	22.27 (38.81)	24.05 (38.52)	28.11 (38.61)	21.93 (38.97)	23.69 (38.67)	17.53 (38.54)
HOS Violations (γ_{02})	−0.73 (113.09)	−1.34 (112.85)	21.04 (114.45)	−7.46 (113.03)	−1.41 (113.65)	17.84 (115.04)	−11.69 (113.57)	30.59 (115.38)
Vehicle Maintenance Violations (γ_{03})	−0.86 (15.36)	−0.75 (15.32)	3.13 (15.76)	13.45 (16.93)	−0.80 (15.35)	3.98 (15.84)	14.51 (17.06)	9.25 (17.11)
Driver Fitness Violations (γ_{04})	−67.81 (411.36)	−73.02 (410.76)	−99.82 (416.56)	−67.91 (411.04)	−71.16 (428.80)	−86.57 (418.25)	−54.17 (412.68)	−197.06 (429.72)
Cross-level interactions								
H6a—Error Type × Unsafe Driver Violations (γ_{11})		−7.41 (39.96)						−78.09 (52.53)
H6b—Error Type × HOS Violations (γ_{12})			197.50* (95.24)			198.91* (95.71)		289.69† (156.95)
H6c—Error Type × Vehicle Maintenance Violations (□$_{13}$)				37.92* (18.79)			38.19* (18.92)	11.61 (24.56)
H6d—Error Type × Driver Fitness Violations (γ_{14})					−3.97 (565.42)			−219.98 (601.44)
Level 1 controls								
Quarter Dummy $1_{Q2\ vs.\ Q1}$ (γ_{20})						98.70* (40.01)	98.60* (40.00)	
Quarter Dummy $2_{Q3\ vs.\ Q1}$ (γ_{30})						73.52† (39.18)	74.50† (39.18)	
Quarter Dummy $3_{Q4\ vs.\ Q1}$ (γ_{40})						8.87* (40.73)	80.35* (40.73)	
Year Dummy $1_{2015\ vs.\ 2014}$ (γ_{50})						21.57 (33.03)	21.62 (33.02)	
Year Dummy $2_{2016\ vs.\ 2014}$ (γ_{60})						−63.16† (36.71)	−63.03† (36.69)	
Level 2 Controls								
Motor Carrier Vehicle Ownership Ratio (γ_{05})	−8.32 (42.38)	−8.47 (42.29)	−9.97 (42.93)	−10.44 (42.37)	−8.28 (42.28)	−6.80 (43.10)	−7.29 (42.52)	−12.30 (42.35)
Motor Carrier Annual Vehicle Miles (γ_{06})	−7.04 (94.78)	−7.56 (94.59)	1.19 (96.03)	−1.07 (94.76)	−6.50 (94.56)	−2.96 (96.36)	−4.98 (95.09)	−4.52 (94.90)
Motor Carrier Number of Drivers (γ_{07})	36.38 (88.41)	37.23 (88.40)	28.90 (89.91)	31.30 (88.51)	35.64 (88.15)	33.64 (90.23)	35.88 (88.81)	38.82 (88.93)

(Continues)

TABLE 9 (Continued)

Level and variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Variance components								
Within-carrier (L1) variance (σ^2)	2,947,611.40	2,947,660.20	2,947,474.50	2,947,545.50	2,947,598.50	2,945,793.60	2,945,866.00	2,947,522.80
Intercept (L2) variance (τ_{00})	106,720.8	106,050.59	105,398.3	104,229.67	105,924.21	107,280.90	106,080.16	101,343.67
Slope (L2) variance (τ_{11})	221,161.49	225,379.4	209,298.2	212,881.2	227,016.61	212,911.18	216,654.00	213,745.33
Intercept-slope (L2) covariance (τ_{01})	0.80	0.80	0.77	0.79	0.80	0.77	0.79	0.77
Additional information								
−2 log likelihood (ML)	−144,558.60	−14,454.00	−144,551.10	−144,552.70	−144,551.40	−144,522.60	−1,445,524.20	−144,533.10
Conditional R^2	0.13	0.13	0.13	0.13	0.13	0.13	0.13	0.13

Note: All models used shipments that were late and motor carriers with over 100 observations only; n = 16,293, k = 96. *Error Type* coded 1 for dispatcher Error and 0 for driver Error. *DELAY* ICC (1) = 0.12, random effect: *Error Type* (γ_{10}), dependent variable: minutes late. $^{\dagger}p < .10$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

14 per day) were delayed due to errors by dispatchers or drivers. Thus, every few hours, LIM must address a disruption in its 24/7 operations. This is a shockingly large number, particularly for a company like LIM, which has made substantial investments in error-prevention technology. This research led to important findings about human errors and delivery delays. Both dispatcher and driver errors were associated with increased likelihood of late delivery, and dispatcher errors were associated with longer delays than driver errors. Most artifacts of carrier latent conditions (HOS, driver fitness, and vehicle maintenance violations) exacerbated this effect. Only unsafe driving violations was not a significant moderator in any of the models.

Further, although avoiding late deliveries is the implied reason for practices resulting in unsafe driving, driver fitness, HOS, and vehicle maintenance violations, our main effects analysis reveals that only driver fitness violations were associated with reduced likelihood of a late delivery and none of the violations main effects was related to significant reductions in delay time. Thus, the findings suggest very little benefit to cutting safety corners in the interest of on-time delivery.

6.2 | Implications for research

This research makes important contributions to the technology management literature by demonstrating error prevention technologies are not necessarily a turnkey solution. Humans provide an important interface with technology through the content and timeliness of their inputs, their ability to collate and analyze the data generated by technology, and their development of insightful policies and priorities based on the technology's outputs. This research highlights the importance of this technology-human interface. Technology cannot prevent all possible human errors. However, understanding latent conditions are more actionable than errors leads to concrete recommendations for improving defensive layers and minimizing resident pathogens through investments in technology, managerial priorities, and policies. For example, the health information systems-induced errors described by Yusuf and Sahroni (2018) could be addressed by improving defensive layers through improved process design and training. Peysakhovich et al. (2018) cited inadequate monitoring and cross-checking by pilots as responsible for more than 80% of aircraft accidents; improving defensive layers by investing in eye tracking technology could provide a non-invasive, relatively inexpensive means for reducing monitoring errors and resulting accidents. Scott et al.'s (2021) description of the unintended consequence of drivers

TABLE 10 *PROB[LATE]* due to dispatcher error: Group mean centered mixed effects logistic regression results

Level and variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Level 1								
Intercept (γ_{00})	−3.99*** (0.13)	−3.95*** (0.13)	−3.95*** (0.12)	−3.97*** (0.13)	−3.97*** (0.23)	−3.88*** (0.15)	−3.89*** (0.15)	−3.99*** (0.13)
Dispatcher Error Centered (γ_{10})	6.06*** (0.10)	5.89*** (0.14)	5.89*** (0.11)	5.99*** (0.14)	6.00*** (0.13)	5.90*** (0.12)	5.99*** (0.11)	6.01*** (0.15)
Level 2								
Unsafe Driving Violations (γ_{01})	0.06 (0.06)	0.04 (0.06)	0.07 (0.06)	0.07 (0.06)	0.06 (0.06)	0.06 (0.06)	0.06 (0.06)	0.07 (0.06)
HOS Violations (γ_{02})	−0.11 (0.18)	−0.10 (0.16)	−0.20 (0.16)	−0.10 (0.16)	−0.09 (0.17)	−0.22 (0.18)	−0.12 (0.19)	−0.25 (0.16)
Vehicle Maintenance Violations (γ_{03})	0.04 (0.03)	0.04 (0.03)	0.04 (0.03)	0.04 (0.03)	0.04 (0.03)	0.05 (0.03)	0.05 (0.03)	0.05† (0.03)
Driver Fitness Violations (γ_{04})	−1.21*** (0.31)	−1.25*** (0.30)	−1.20*** (0.26)	−1.21* (0.36)	−1.50† (0.86)	−1.25*** (0.37)	−1.54** (0.59)	−1.21*** (0.19)
Group Mean Dispatcher Error (γ_{05})	9.37*** (0.27)	9.36** * (0.27)	9.40*** (0.22)	3.38*** (0.27)	9.33*** (1.36)	9.12*** (0.47)	8.96*** (0.54)	9.46*** (0.25)
Cross-level interactions								
H2a—Dispatcher Error Centered × Unsafe Driver Violations (γ_{11})		0.10 (0.07)						−0.02 (0.08)
H3a—Dispatcher Error Centered × HOS Violations (γ_{12})			0.41* (0.16)			0.41* (0.17)		0.57** (0.18)
H4a—Dispatcher Error Centered × Vehicle Maintenance Violations (γ_{13})				0.02 (0.03)				−0.04 (0.04)
H5a—Dispatcher Error Centered × Driver Fitness Violations (γ_{14})					0.86** (0.29)		0.93* (0.39)	0.06 (0.21)
Level 1 controls								
Quarter Dummy $1_{Q2\ vs.\ Q1}$ (γ_{20})						0.17*** (0.03)	0.17*** (0.03)	
Quarter Dummy $2_{Q3\ vs.\ Q1}$ (γ_{30})						0.14*** (0.03)	0.14*** (0.03)	
Quarter Dummy $3_{Q4\ vs.\ Q1}$ (γ_{40})						−0.01 (0.03)	−0.01 (0.03)	
Year Dummy $1_{2015\ vs.\ 2014}$ (γ_{50})						−0.05† (0.03)	−0.05† (0.03)	
Year Dummy $2_{2016\ vs.\ 2014}$ (γ_{60})						−0.40*** (0.03)	−0.40*** (0.03)	
Level 2 controls								
Motor Carrier Vehicle Ownership Ratio (γ_{06})	0.08 (0.08)	0.08 (0.08)	0.08 (0.08)	0.08 (0.08)	0.08 (0.08)	0.08 (0.08)	0.09 (0.08)	0.08 (0.08)
Motor Carrier Annual Vehicle Miles (γ_{07})	−0.18 (0.14)	−0.18 (0.13)	−0.17 (0.15)	−0.18 (0.14)	−0.17 (0.22)	−0.17 (0.19)	−0.16 (0.16)	−0.18 (0.14)
Motor Carrier Number of Drivers (γ_{08})	−0.11 (0.14)	−0.10 (0.13)	−0.11 (0.16)	−0.11 (0.14)	−0.11 (0.21)	−0.13 (0.19)	−0.14 (0.16)	−0.11 (0.14)

(Continues)

TABLE 10 (Continued)

Level and variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Variance components								
Within-carrier (L1) variance (σ^2)	3.29	3.29	3.29	3.29	319	3.29	3.29	3.29
Intercept (L2) variance (τ_{00})	0.67	0.67	0.67	0.67	0.67	0.65	0.65	0.67
Slope (L2) variance (τ_{11})	0.83	0.80	0.79	0.83	0.82	0.80	0.83	0.78
Intercept-slope (L2) covariance (τ_{01})	−0.46	−0.46	−0.46	−0.46	−0.47	−0.44	−0.45	−0.46
Additional information								
−2 log likelihood (ML)	−34,251.10	−34,250.20	−34,248.70	−34,250.90	−34,250.70	−34,133.20	−34,153.10	−34,248.30
Conditional R^2	0.34	0.34	0.34	0.34	0.34	0.34	0.34	0.34

Note: All models used motor carriers with over 100 observations only; n = 299,399, k = 97. *Dispatcher Error* coded 1 for error and 0 for no error. *PROB[LATE]* ICC (1) = 0.14, random effect: *Dispatcher Error Centered* (γ_{10}), dependent variable; 0 = on time, 1 = late. $^{\dagger}p < .10$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

whose HOS are monitored by ELDs subsequently increasing other undesired behaviors such as speeding, could be addressed through eliminating the resident pathogen of mileage-based compensation and replacing it with a defensive layer of behavior-based compensation.

This research contributes to the supply chain literature by emphasizing the importance of a carrier's latent conditions in intensifying the relationship between errors and delivery delays. Management practices often focus on directly preventing errors, which Reason (2000) likens to swatting at individual mosquitoes. Reason (2000) emphasizes the importance of focusing on latent conditions that intensify the adverse consequences of active failures, rather than on active failures, per se. Thus, the greater benefit comes from "draining the swamp" (Reason, 2000), removing or blocking resident pathogens. Analogous to the relationship between assignable and common causes of variability in statistical process control, while front-line supply chain operators frequently commit errors, the outcomes are exacerbated by latent conditions that only management can address.

This research further contributes to supply chain research by considering human error as a source of delivery delays and moving beyond the assumption that technology can automate human variability out of processes. There are many factors that can cause a delayed delivery but developing a better understanding of the role of active failures and latent conditions helps advance knowledge on preventing and reducing delivery delays.

6.3 | Implications for practice

Although every supply chain journey involves a dynamic mix of unavoidable factors like weather or traffic conditions, overlaid on these factors is the potential for dispatcher or driver errors to delay delivery, exacerbated by carrier latent conditions. Regardless of whether errors occur at a carrier's headquarters or on the road, the result is the need to reschedule delayed deliveries and appease downstream customers.

How can latent conditions be addressed, to minimize the impact of errors? According to NAT, adverse consequences occur when defensive layers are inadequate (Reason, 2000). Prior research indicates that some high-risk organizations have achieved very low adverse consequence rates through careful design and management of technology, people, and processes (Gaba, 2000), becoming high-reliability organizations by developing defensive layers that proactively neutralize resident pathogens (Bigley & Roberts, 2001; Ruchlin et al., 2004; Waller & Roberts, 2003).

TABLE 11 *PROB[LATE]* due to driver error: Group mean centered mixed effects logistic regression results

Level and variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Level 1								
Intercept (γ_{00})	−3.04*** (0.12)	−3.03*** (0.11)	−3.03*** (0.12)	−3.05*** (0.10)	−3.03*** (0.12)	−2.96*** (0.14)	−2.96*** (0.14)	−3.05*** (0.10)
Driver Error Centered (γ_{10})	4.32*** (0.12)	4.18*** (0.12)	4.21*** (0.13)	4.43*** (0.14)	4.18*** (0.12)	4.20*** (0.16)	4.17*** (0.15)	4.43*** (0.13)
Level 2								
Unsafe Driving Violations (γ_{01})	−0.03 (0.05)	−0.04 (0.05)	−0.03 (0.05)	−0.03 (0.05)	−0.03 (0.05)	−0.04 (0.06)	−0.04 (0.06)	−0.03 (0.05)
HOS Violations (γ_{02})	−0.10*0.12)	−0.10 (0.12)	−0.12 (0.14)	−0.10 (0.12)	−0.10 (0.13)	−0.13 (0.18)	−0.11 (0.16)	−0.14 (0.11)
Vehicle Maintenance Violations (γ_{03})	0.03 (0.03)	0.03 (0.03)	0.03 (0.03)	0.04 (0.03)	0.03 (0.03)	0.04 (0.03)	0.04 (0.03)	0.04 (0.03)
Driver Fitness Violations (γ_{04})	−1.06*** (0.16)	1.06*** (0.16)	−1.06*** (0.26)	−1.06*** (0.20)	−1.26*** (0.17)	−1.12** (0.40)	−1.36*** (0.33)	−1.19*** (0.15)
Group Mean Driver Error (γ_{05})	1.01*** (0.17)	1.01*** (0.19)	1.06*** (0.28)	0.99*** (0.14)	1.02*** (0.17)	0.85* (0.37)	0.81* (0.35)	1.02*** (0.13)
Cross-level interactions								
H2b—Driver Error Centered × Unsafe Driver Violations (γ_{11})		0.08 (0.06)						−0.05 (0.07)
H3b—Driver Error Centered × HOS Violations (γ_{12})			0.27* (0.13)			0.27 (0.20)		0.51*** (0.14)
H4b—Driver Error Centered × Vehicle Maintenance Violations (γ_{13})				−0.03 (0.04)				−0.10* (0.04)
H5b—Driver Error Centered × Driver Fitness Violations (γ_{14})					1.98*** (0.18)		2.08*** (0.37)	1.27*** (0.14)
Level 1 controls								
Quarter Dummy $1_{\text{Q2 vs. Q1}}$ (γ_{20})						0.09** (0.03)	0.09** (0.03)	
Quarter Dummy $2_{\text{Q3 vs. Q1}}$ (γ_{30})						0.08** (0.03)	0.08** (0.03)	
Quarter Dummy $3_{\text{Q4 vs. Q1}}$ (γ_{40})						0.05† (0.03)	0.05† (0.03)	
Year Dummy $1_{\text{2015 vs. 2014}}$ (γ_{50})						0.02 (0.02)	0.02 (0.02)	
Year Dummy $2_{\text{2016 vs. 2014}}$ (γ_{60})						−0.38*** (0.03)	−0.38*** (0.03)	
Level 2 controls								
Motor Carrier Vehicle Ownership Ratio (γ_{06})	0.01 (0.07)	0.01 (0.07)	0.01 (0.08)	0.01 (0.08)	0.01 (0.08)	0.01 (0.08)	0.02 (0.08)	0.01 (0.07)
Motor Carrier Annual Vehicle Miles (γ_{07})	0.27* (0.12)	0.27* (0.114)	0.27* (0.12)	0.27* (0.11)	0.27* (0.11)	0.28 (0.18)	0.28† (0.16)	0.27* (0.11)
Motor Carrier Number of Drivers (γ_{08})	−0.27* (0.13)	−0.27* (0.11)	−0.27* (0.13)	−0.27* (0.10)	−0.27* (0.12)	−0.29 (0.18)	−0.30† (0.17)	−0.27* (0.11)

(Continues)

TABLE 11 (Continued)

Level and variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Variance components								
Within-carrier (L1) variance (σ^2)	3.29	3.29	3.29	3.29	3.29	3.29	3.29	3.29
Intercept (L2) variance (τ_{00})	0.57	0.57	0.57	0.57	0.57	0.56	0.56	0.57
Slope (L2) variance (τ_{11})	1.96	1.92	1.92	1.95	1.91	1.95	1.92	1.85
Intercept–slope (L2) covariance (τ_{01})	−0.19	−0.19	−0.18	−0.2	−0.19	−0.2	−0.21	−0.19
Additional information								
−2 log likelihood (ML)	46,832.40	−46,831.90	46,831.70	46,832.20	46,831.40	−46,674.80	−46,674.40	46,830.10
Conditional R^2	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16

Note: All models used motor carriers with over 100 observations only; n = 299,399, k = 97. *Dispatcher Error* coded 1 for error and 0 for no error. *PROB[LATE]* ICC (1) = 0.14, random effect: *Driver Error Centered* (γ_{10}), dependent variable: 0 = on time, 1 = late. $^{\dagger}p < .10$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

Managerial defensive layers, such as a policy of platooning four trucks together, provide four drivers to double-check directions and proctor each other's driving behavior. Similarly, two massive mega-trucks carrying the same amount of cargo as three standard trucks have proportionally fewer opportunities for human errors. Other managerial defensive layers include reduction of the approved carrier list by removing those whose violations record indicates resident pathogens. A shipper could contract only with high-compliance carriers having few CAS violations, dropping those with a history of unsafe driving from consideration or limiting carriers to only those with well-integrated technology-enabled systems that facilitate 360° visibility. For a high-volume supply chain connection, such as from a factory to a major distribution center, a shipper could contract for dedicated services with a limited set of trusted carriers. New technologies improve truck visibility and provide updated ETA information, allowing a dispatcher to adjust recipients' expectations, while other technologies allow validation of vehicle type, licenses and permits prior to dispatching.

Defensive layers can also be strengthened through investments in technology to minimize resident pathogens through error proofing (Norman, 1981). AI and autonomous vehicle technologies reduce the adverse consequences of errors, as automation replaces drivers and dispatchers. Predictive analytics, including AI systems, can reduce the impact of dispatcher errors. For example, ClearMetal analyzes supply chain lanes to place containers more cost efficiently (Banker, 2016), and IBM's Watson Research Center focuses on real-time analytics for traffic prediction and multimodal dynamic routing (IBM, 2017).

6.4 | Limitations

Like all research, our design had several limitations. It was based on a deep dive into the operations of a single, large shipper, potentially limiting generalizability. However, the use of a single shipper also contributes to the strength of our findings by controlling for shipper-level effects; LIM's investments in technologies are constant across all shipments. LIM had made very substantial investments in technology designed to prevent dispatcher and driver errors, thus we expect that shippers with less investment in technology will have a greater likelihood of adverse consequences resulting from an error, as well as longer delivery delays. Further, LIM draws its drivers from the same pool of carriers used by many other large shippers, thus, we expect carrier effects to be similar across shippers.

TABLE 12 *DELAY* due to error type: Group mean centered mixed effects regression results

Level and variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9
Level 1									
Intercept (γ_{00})	547.95*** (41.97)	281.77** (103.36)	283.48** (103.34)	274.29** (103.75)	255.67* (103.72)	283.23** (109.85)	213.08† (0.08)	194.19† (110.12)	278.00** (1.04.58)
Error Type (Dispatcher vs. Driver Error) (γ_{10})	520.35*** (59.99)								
Error Type (Dispatcher vs. Driver Error) Centered (γ_{20})		493.78*** (55.57)	504.24*** (85.79)	419.07*** (66.35)	375.16*** (83.33)	495.00*** (67.76)	421.85*** (66.72)	377.79*** (83.85)	483.71*** (112.58)
Level 2									
Unsafe Driving Violations (γ_{01})		20.64 (37.56)	20.62 (37.54)	15.38 (37.73)	17.48 (37.46)	20.79 (37.57)	15.00 (37.97)	17.10 (37.69)	12.23 (37.52)
HOS Violations (γ_{02})		61.33 (110.27)	64.48 (110.11)	77.24 (111.33)	52.23 (110.11)	60.17 (110.84)	74.86 (112.23)	48.75 (110.98)	83.06 (112.24)
Vehicle Maintenance Violations (γ_{03})		−3.91 (15.18)	−3.80 (15.15)	0.03 (15.56)	8.98 (16.61)	−3.82 (15.17)	0.80 (15.70)	9.97 (16.79)	5.60 (16.78)
Driver Fitness Violations (γ_{04})		48.19 (404.85)	42.97 (404.53)	16.54 (409.62)	84.46 (403.84)	44.75 (419.36)	27.43 (412.50)	54.95 (406.61)	−75.50 (421.39)
Group Mean Error Type (γ_{05})		5116.26*** (1288.52)	5085.33*** (1288.79)	5041.94*** (1289.20)	4929.74*** (1287.78)	5080.64*** (1289.36)	5037.32*** (1301.07)	4921.06*** (1299.49)	4760.11*** (1289.03)
Cross-level interactions									
H6a—Error Type Centered × Unsafe Driver Violations (γ_{21})			−5.60 (37.87)						−69.17 (50.35)
H6b—Error Type Centered × HOS Violations (γ_{22})				179.09* (90.59)			180.85* (91.02)		260.51† (149.72)
H6c—Error Type Centered × Vehicle Maintenance Violations (γ_{23})					33.30† (17.56)			33.60† (17.67)	10.00 (23.08)
H6d—Error Type Centered × Driver Fitness Violations (γ_{24})						−0.47 (526.92)			−181.39 (567.43)

(Continues)

TABLE 12 (Continued)

Level and variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9
Level 1 controls									
Quarter Dummy $1_{Q2\ vs.\ Q1}$ (γ_{30})							98.63* (40.01)	98.46* (40.01)	
Quarter Dummy $2_{Q3\ vs.\ Q1}$ (γ_{40})							25.14 (33.05)	25.12 (33.03)	
Year Dummy $1_{2015\ vs.\ 2014}$ (γ_{60})							25.14 (33.05)	25.12 (33.03)	
Year Dummy $2_{2016\ vs.\ 2014}$ (γ_{70})							−57.52 (36.73)	−57.50 (36.71)	
Level 2 controls									
Motor Carrier Vehicle Ownership Ratio (γ_{06})	−3.49 (39.48)	−24.18 (41.54)	−24.22 (41.47)	−25.41 (42.00)	−25.52 (41.43)	−24.05 (41.44)	−22.33 (42.29)	−22.44 (41.72)	−26.76 (41.45)
Motor Carrier Annual Vehicle Miles (γ_{07})	2.59 (90.20)	−30.72 (94.05)	−30.95 (93.89)	−22.62 (95.15)	−24.80 (93.84)	−30.23 (93.85)	−27.01 (95.80)	−28.98 (94.46)	−26.84 (94.00)
Motor Carrier Number of Drivers (γ_{08})	19.69 (83.34)	62.91 (88.22)	63.37 (88.23)	54.95 (89.53)	57.41 (88.07)	62.13 (87.97)	60.14 (90.16)	62.45 (88.68)	62.90 (88.47)
Variance components									
Within-carrier (L1) variance (σ^2)	2,947,347.80	2,949,740.40	2,949,757.00	2,949,498.70	2,949,518.00	2,949,688.1	2,947,800.7	2,947,821.4	2,949,319.5
Intercept (L2) variance (τ_{00})	99,700.29	101,092.1	100,686.08	100,207.19	99,281.25	100,589.32	102,792.00	177,106.10	97,470.588
Slope (L2) variance (τ_{11})	227,568.4	176,625.5	180,648.12	169,698.85	17,401.17	182,256.9	172,567.40	177,106.10	178,865.05
Intercept-slope (L2) covariance (τ_{01})	0.80	0.79	0.79	0.75	0.78	0.79	0.75	0.78	0.76
Additional information									
−2 log likelihood (ML)	−144,579.20	−144,547	−144,542.50	−144,539.80	−144,541.50	−144,539.90	−144,511.40	−144,513.10	−144,522.20
Conditional R^2	0.13	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12

Note: All models used shipments that were late and motor carriers with over 100 observations only; n = 16,923, k = 96. *Error Type* coded 1 for dispatcher error and 0 for driver error. *DELAY* ICC(1) = 0.12, random effect: *Error Type Centered* (γ_{10}), dependent variable: minutes late. $^{\dagger}p < .10$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

Our operationalization of active failures was based on LIM's arrival codes, which are biased toward rule-based and knowledge-based errors, whose immediate effect is more substantive than skill-based errors. Tapping into the "stupid mistakes" that constitute many skill-based errors is not possible using archival data sources, since most would not be reported. We encourage future research to study errors using alternative data sources, such as blogs, surveys, and interviews with drivers and dispatchers. Further, by operationalizing errors using arrival codes, we acknowledge that there are likely additional errors not included in the analysis because they did not result in late delivery. Similarly, we operationalized carrier latent conditions based on violations as a proxy for their tacit values. More direct measurement of technology investments and surveys related to carrier policies, priorities, and practices should yield interesting insights into resident pathogens and defensive layers.

6.5 | Opportunities for future research

The descriptive statistics in Table 4 revealed some intriguing findings. First, errors do not necessarily result in delivery delays; a third of the errors, overall, were not associated with any delivery delay. Second, although more shipments had driver errors than dispatcher errors (12,845 vs. 11,349), dispatcher errors were much more likely to be associated with a delivery delay (83.5% of dispatcher errors vs. 53% of driver errors). This may be due to dispatcher errors being more serious, drivers developing strategies to compensate for their own errors, or other reasons that pose interesting opportunities for future research.

The timing of the data provides an opportunity to examine important relationships before and after the FMCSA's ELD mandate went into effect. Our data is for shipments conveyed between 2014 and 2016. ELD vendors have been able to self-certify compliance of their ELDs with the FMCSA mandate since February 16, 2016. Although we do not have information about which of LIM's carriers had ELDs installed when the shipments in our dataset took place, very few leading carriers were self-certified during this period. Since December 18, 2017, all long-haul carriers have been subject to the ELD mandate, unless their vehicles are grandfathered, thus the future may be very different as ELDs are required to collect more data. This provides an interesting opportunity for a natural experiment, examining delivery delays and other important outcomes before and after the ELD mandate went into effect.

Perhaps the most intriguing finding is that, although many technologies are marketed as turnkey error prevention solutions, human errors still occur, even in a shipper using state-of-the-art technologies. Although many minor errors can be considered "stupid mistakes," even minor errors can have adverse consequences. Like LIM, many shippers and carriers have made substantial investments in defensive layers in the form of technologies intended to prevent errors, yet our findings indicate that these investments are not paying off as well as anticipated. Why is this the case? Does the problem lie in the technologies, the interface between human operators and the technologies, or elsewhere? How is this related to latent conditions? Do error rates vary between shippers and across carriers? Researchers need to dig into the underlying drivers of these findings.

Similarly, there are many interesting opportunities for future research related to carrier latent conditions. For example, mileage-based pay uses an outcome-based contract to align driver compensation with the carrier's interests. Alternatively, carriers could consider various forms of behavior-based contracts to incentivize behaviors expected to lead to desired outcomes (Zu & Kaynak, 2012), potentially converting a managerial resident pathogen into a defensive layer. For example, a carrier could give drivers with no safety violations an early opportunity to select their shipments or pay the safest drivers a salary (guaranteed income). Research focused on specific policies and priorities can help clarify how specific resident pathogens and defensive layers interact. Since our research constructs touch on regulators and legal regulations, there are also many opportunities for future research to dig more deeply into public policy implications.

NAT offers a foundation for understanding human errors and their potential for adverse consequences at all levels in supply chains and in many other operations management applications. We do not view our findings as limited to the trucking context and look forward to seeing broader applications of NAT and testing of the generalizability of our findings to other settings.

7 | CONCLUSIONS

Although, "to err is human, to forgive divine," supply chains are not forgiving. Even a minor error by a truck driver or dispatcher can translate into a delivery delay, whose adverse consequences are intensified by carrier latent conditions. Although we examined cross-level interaction effects on individual shipments, the larger implications are for the extended supply chain, where a single delivery delay can have a cascading effect. Consider the analogy of a flight whose arrival is delayed; departing flights are delayed as they await transfer of the late flight's passengers, causing flights from other airports to be subsequently delayed. The same cascading effect occurs in supply chains, particularly those that are tightly coupled with minimal slack, as downstream customers wait for their materials, forcing their customers to wait, and so forth.

Researchers and managers should not like what the data about supply chain errors and carrier latent conditions reveals. Resident pathogens in carrier latent conditions foretell even more late deliveries in the future. It is bad enough that so many violations occur, but what makes it even worse is that cutting safety corners appears to compound the adverse consequences of errors, rather than leading to faster deliveries. This emphasizes the importance of investments in technology and reassessment of policies and priorities by carriers.

More broadly, this exploratory research suggests that errors and latent conditions have a more pervasive effect than most managers realize. This raises important concerns for operations management, in general. For every minute of time saved through process improvements within a factory, there are hours of delay yet to be reduced in its supply chains. After many decades of evolution of supply chain practice, why have carriers with worrisome latent conditions been able to continue to operate in such a manner? It is our hope that future research further extends understanding of how to address latent conditions to minimize the impact of errors in various contexts and move toward supply chains that are high reliability, rather than high-risk, organizations.

ENDNOTES

[1] The standard mileage-based pay is about $0.35 per mile for company drivers and $1.00 per mile for owner-operators. (https://www.truckdriverssalary.com/) Drivers are expected to cover most of their variable expenses, including food, tolls, fines for maintenance issues, and speeding tickets; owner-operators must also cover the fixed capital cost of vehicles.

[2] In the United States, truck drivers are allowed to drive 11 h in a 14-h period (https://www.fmcsa.dot.gov/regulations/hours-service/summary-hours-service-regulations). Canada allows truck drivers to drive 13 h during a 14-h period (https://laws-lois.justice.gc.ca/eng/regulations/SOR-2005-313/page-1.html), while Mexico limits drivers to 8 h of daylight driving per day or 7 h of nighttime driving (https://loadtrek.net/2018/07/14/operating-in-mexico-hours-of-service-for-commercial-vehicles).

[3] For *PROB[LATE]* predicted by dispatcher error, the −2 log likelihood difference was 357.93, $p < .001$; for *PROB[LATE]* predicted by driver error, the −2 log likelihood difference was 1833.40, $p < .001$; for *DELAY* predicted by error type, the −2 log likelihood difference was 89.46, $p < .001$.

[4] $PROB[LATE]_{ij}$ is the likelihood of late delivery for the ith shipment for the jth carrier, β_{0j} is the carrier-specific intercept for the jth carrier, β_{1j} is the carrier-specific slope for dispatcher error in the jth carrier, and r_{ij} is the residual for the ith shipment in the jth carrier. γ_{00} is the overall intercept for the likelihood of late delivery across all carriers, γ_{01} captures how much the intercept of likelihood of late delivery is expected to change for a one-unit change in violations for carrier j, and u_{0j} is a random effect for the jth carrier that captures how much the intercept for that carrier differs from the overall intercept after accounting for violations. γ_{10} is the overall slope for the likelihood of late delivery across all carriers, γ_{11} captures how much the slope of likelihood of late delivery is expected to change for a one-unit change in violations for carrier j, and u_{1j} is a random effect for the jth carrier that captures how much the slope for that carrier differs from the overall slope after accounting for violations.

REFERENCES

Adelman, J. (2019). High-reliability healthcare: Building safer systems through just culture and technology. *Managing Risk*, *64*(3), 137–141.

Aguinis, H., Gottfredsen, R. K., & Culpepper, S. A. (2003). Best-practice recommendations for estimating cross-level interaction effects using multilevel modeling. *Journal of Management*, *39*, 1490–1528.

Aird, P. (2019). Deepwater programs, safety, and loss control. In P. Aird (Ed.), *Deepwater drilling: Well planning, design, engineering, operations, and technology applications*. Elsevier.

Antonakis, J., Bastardoz, N., & Rönkkö, M. (2019). On ignoring the random effects assumption in multilevel models: Review, critique, and recommendations. *Organizational Research Methods*, *24*(2), 443–483.

Asare, A. K., Brashear-Alejandro, T. G., & Kang, J. (2016). B2B technology adoption in customer-driven supply chains. *Journal of Business & Industrial Marketing*, *31*(1), 1–12.

Averett, P. (2001). People: The human side of systems technology. *The Journal for Quality and Participation*, *24*(2), 34–37.

Banker, S. (2016). Predictive analytics comes to the logistics industry. https://logisticsviewpoints.com/2016/02/29/predictive-analytics-comes-to-the-logistics-industry/

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

Bednar, P. M., & Welch, C. (2020). Socio-technical perspectives on smart working: Creating meaningful and sustainable systems. *Information Systems Frontiers*, *22*(2), 281–298.

Beiger, R. G., & Kichak, J. P. (2004). Computerized physician order entry: Helpful or harmful? *Journal of the American Medical Informatics Association*, *11*(2), 100–103.

Bell, A., & Jones, K. (2015). Explaining fixed effects: Random effects modeling of time-series cross-sectional and panel data. *Political Science Research and Methods*, *3*(1), 133–153.

Bigley, G., & Roberts, K. (2001). The incident command system: High-reliability organizing for complex and volatile task events. *Academy of Management Journal*, *44*(6), 1281–1299.

Bliese, P. (2016). multilevel: Multilevel Functions. R package version 2.6. https://CRAN.R-project.org/package=multilevel

Bliese, P., Maltarich, M., & Hendricks, J. (2018). Back to basics with mixed-effects models: Nine take-away points. *Journal of Business and Psychology*, *33*(1), 1–23.

Bliese, P., Schepker, D. J., Essman, S. M., & Ployhart, R. E. (2020). Bridging methodological divides between macro- and microresearch: Endogeneity and methods for panel data. *Journal of Management*, *46*(1), 70–99.

Bode, C., & Wagner, S. M. (2015). Structural drivers of upstream supply chain complexity and the frequency of supply chain disruptions. *Journal of Operations Management*, *36*, 215–228.

Brammer, S., Jackson, G., & Matten, D. (2012). Corporate social responsibility and institutional theory: New perspectives on private governance. *Socio-Economic Review*, *10*, 3–28.

Britto, R., Corsi, T., & Grimm, C. (2010). The relationship between carrier financial performance and safety performance (Master's thesis). Concordia University, Montreal, Quebec, Canada.

Brown, K. (1996). Workplace safety: A call for research. *Journal of Operations Management*, *14*(1), 157–171.

Browne, W. J., & Draper, D. (2006). A comparison of Bayesian and likelihood-based methods for fitting multilevel models. *Bayesian Analysis*, *1*(3), 473–574.

Cameron, K., & Quinn, R. (2011). *Designing and changing organizational culture: Based on the competing values framework*. Jossey-Bass.

Cantor, D., Celebi, H., Corsi, T., & Grimm, C. (2013). Do owner-operators pose a safety risk on the nation's highways? *Transportation Research Part E: Logistics and Transportation Review*, *59*, 34–47.

Cantor, D., Corsi, T., & Grimm, C. (2008). Deterrents of carrier safety technology adoption. *Transportation Research Part E: Logistics and Transportation Review*, *44*, 932–947.

Cantor, D., Corsi, T., & Grimm, C. (2009). Do electronic logbooks contribute to carrier safety performance? *Journal of Business Logistics*, *30*(1), 203–222.

Cantor, D., Corsi, T., Grimm, C., & Singh, P. (2016). Technology, firm size, and safety: Theory and empirical evidence from the U.S. carrier industry. *Journal of Business Logistics*, *55*(2), 149–167.

Caplice, C. (2007). Electronic markets for truckload transportation. *Production and Operations Management*, *16*(4), 423–436.

Certo, S. T., Withers, M. C., & Semadeni, M. (2017). A tale of two effects: Using longitudinal data to compare within-and between-firm effects. *Strategic Management Journal*, *38*(7), 1536–1556.

Chakravorty, S. S. (2011). The trouble with too much information. *MIT Sloan Management Review*, *53*(1), 96.

Cowan, L. (2003). Literature review and risk mitigation strategy for unintended consequences of computerized physician order entry. *Nursing Management*, *31*(1), 27–32.

DeFlorio, F. (2016). Continuing airworthiness and air operator's certification. In F. DeFlorio (Ed.), *Airworthiness: An introduction to aircraft certification and operation*. Elsevier.

Detert, J. R., Schroeder, R. S., & Mauriel, J. J. (2000). A framework for linking culture and improvement initiatives in firms. *Academy of Management Review*, *25*(4), 850–863.

Douglas, J., & Larrabee, S. (2003). Bring barcoding to the bedside. *Nursing Management*, *34*, 37–40.

Douglas, M., & Swartz, S. (2016). Truck driver safety: A evolutionary research approach. *Transportation Journal*, *55*(3), 258–281.

Elliott, R. A., Putman, K. D., Franklin, M., Annemons, L., Verhaeghem, N., Eden, M., Hayre, J., Rogers, S., Sheikh, A., & Avery, A. J. (2014). Cost-effectiveness of a pharmacist-led information technology intervention for reducing rates of clinically important errors in medicines management in general practices. *PharmoEconomics*, *32*, 573–590.

Federal Motor Carrier Safety Administration. (2019). *Safety measurement system (SMS) methodology: Behavior analysis and safety improvement category (BASIC) prioritization status*. U.S. Department of Transportation.

Gaba, D. M. (2000). Structural and organizational issues in patient safety: A comparison of health care to other high hazard industries. *California Management Review*, *43*(1), 83–102.

Galbraith, J. R. (2014). Organizational design challenges resulting from big data. *Journal of Organization Design*, *3*(1), 2–13.

Gangwa, H., Date, H., & Raoot, A. D. (2014). Review on IT adoption: Insights from recent technologies. *Journal of Enterprise Information Management*, *27*(4), 488–502.

Grabowski, M., & Roberts, K. (1997). Risk mitigation in large-scale systems: Lessons from high reliability firms. *California Management Review*, *39*(4), 152–162.

Granot, H. (1998). The human factor in industrial disasters. *Disaster Prevention and Management*, *7*(2), 92–102.

Hagberg, M., Silverstein, B., Wells, R., Smith, M., Hendrick, H. W., Carayon, P., & Pérusse, M. (1995). *Work related musculoskeletal disorders (WMSDs): A reference book for prevention*. Taylor & Francis.

Harrington, L., Kennerly, D., & Johnson, C. (2011). Safety issues related to the electronic medical record (EMR): Synthesis of the literature from the last decade, 2000-2009. *Journal of Healthcare Management*, *56*(1), 31–44.

Hickman, J., Buo, F., Camden, M., Hanowski, R., Medina, A., & Mabry, J. (2015). Efficiency of roll stability control and lane departure warning systems using carrier-collected data. *Journal of Safety Research*, *52*, 59–63.

Hickman, J., & Hanowski, R. (2011). Use of a video monitoring approach to reduce at-risk behaviors in commercial vehicle operations. *Transportation Research Part F: Traffic Psychology and Behaviour*, *14*, 189–198.

Hofmann, D., Jacobs, R., & Landy, F. (1995). High reliability process industries: Individual, micro and macro organizational safety influences on safety performance. *Journal of Safety Research*, *26*(3), 131–149.

Hofmann, D., & Stetzer, A. (1996). A cross-level investigation of factors influencing unsafe behaviors and accidents. *Personnel Psychology*, *49*(2), 307–339.

Hofmann, D., & Stetzer, A. (1998). The role of safety culture and communication in accident interpretation: Implications for learning from negative events. *Academy of Management Journal*, *41*(6), 644–657.

Howe, J. L., Butler, R. L., Kim, T. C., & Kellogg, K. M. (2020). Addressing adverse events in health care through a safety science lens. In E. Iadanza (Ed.), *Clinical engineering handbook*. Elsevier.

Hubbard, T. (2003). Information, decisions and productivity: Onboard computers and capacity utilization in trucking. *The American Economic Review*, *93*(4), 1328–1353.

IBM. (2017). Smarter transportation. http://researcher.watson.ibm.com/researcher/view-group.php?id-3944.

Johnson, J., Bristow, J., McClure, D., Schneider, D., & Kenneth, C. (2009). Long distance drivers: Their joys and frustrations. *Journal of Transport Management*, *20*(1), 1–20.

Kanki, B. G., & Hobbs, A. (2018). Organizational factors and safety culture. In T. Sgobba, B. Kanki, J. F. Clervoy, & G. M. Sanders (Eds.), *Space safety and human performance*. Elsevier.

Karimi, J., Somers, T. M., & Gupta, Y. P. (2004). Impact of environmental uncertainty and task characteristics on user satisfaction with data. *Information Systems Research*, *15*(2), 175–193.

Knoedler, J. (2003). *IT integration transforms the care continuum* (pp. 8–9). IT Solutions.

Leape, L. (1994). Error in medicine. *JAMA*, *272*(23), 1851–1857.

Lechler, S., Canzaniello, A., Robmann, B., Heiko, A., & Hartmann, E. (2019). Real-time data processing in supply chain management: Revealing the uncertainty dilemma. *International Journal of Physical Distribution & Logistics Management*, *49*(10), 1003–1019.

Leitch, S., & Warren, M. J. (2010). ETHICS: The past, present and future of sociotechnical systems design. *IFIP Advances in Information and Communication Technology*, *325*, 189–197.

Markarian, J. (2019). Embracing the digital factory for bio/pharma manufacturing. *Pharmaceutical Technology*, *43*, 16–21.

Manrodt, K., Kent, J., & Parker, R. (2003). Operational implications of mobile communications in the carrier industry. *Transportation Journal*, *32*(3), 50–58.

McNall, L., & Stanton, J. (2011). Private eyes are watching you: Reactions to location sensing technologies. *Journal of Business Logistics*, *26*, 299–399.

McNeish, D., & Kelley, K. (2018). Fixed effects models versus mixed effects models for clustered data: Reviewing the approaches, disentangling the differences, and making recommendations. *Psychological Methods*, *24*(1), 20–35.

Melville, N., & Ramirez, R. (2008). Information technology innovation diffusion: An information requirements paradigm. *Information Systems Journal*, *18*(3), 247–273.

Mehmood, R., Meriton, R., Graham, G., Hennelly, P., & Kumar, M. (2017). Exploring the influence of big data on city transport operations: A Markovian approach. *International Journal of Operations and Production Management*, *37*(1), 75–104.

Meyer, J. W., & Rowan, B. (1977). Institutionalized firms: Formal structure vs. myth and ceremony. *American Journal of Sociology*, *83*, 340–363.

Miller, J., Golicic, S., & Fugate, B. (2017a). Developing and testing a dynamic theory of carrier safety. *Journal of Business Logistics*, *38*(2), 96–114.

Miller, J., & Saldanha, J. (2016). A new look at the longitudinal relationship between carrier financial performance and safety. *Journal of Business Logistics*, *37*(3), 284–306.

Miller, J., Saldanha, J., Rungtusanatham, M., & Knemeyer, M. (2017b). How does driver turnover affect carrier performance, and what can managers do about it? *Journal of Business Logistics*, *38*(3), 197–216.

Miller, J., Saldanha, J., Rungtusanatham, M., Knemeyer, A., & Goldsby, T. (2018a). How does electronic monitoring affect hours-of-service compliance? *Transportation Journal*, *57*(4), 329–364.

Miller, J., Schwieterman, M., & Bolumole, Y. (2018b). Effects of carriers' growth or contraction on safety: A multiyear panel analysis. *Journal of Business Logistics*, *39*(2), 138–156.

Moore, E. W., Warmke, J. M., & Gorban, L. R. (1991). The indispensable role of management science in centralizing freight operations at Reynolds metals company. *Interfaces*, *21*(1), 107–129.

Ngwenyama, O., & Nielsen, P. A. (2014). Using organizational influence processes to overcome information systems implementation barriers: Lessons from a longitudinal case study of SPI implementation. *European Journal of Information Systems*, *23*, 205–222.

Nnaji, C., Gambayese, J., Karakhan, A., & Eseonu, C. (2019). Influential safety technology adoption predictors in construction. *Engineering, Construction & Architectural Management*, *26*(11), 2655–2681.

Norman, D. (1981). Categorization of action slips. *Psychological Review*, *88*(1), 1–15.

O'Conor, G., & Smallman, C. (1995). The hybrid manager: A review. *Management Decision*, *33*(7), 19–28.

Perrow, C. (1999). *Normal accidents: Living with high-risk technologies*. Princeton University Press.

Peysakhovich, V., Lefrancois, O., Debais, F., & Causse, M. (2018). The neuroergonomics of aircraft cockpits: The four stages of eye-tracing integration to enhance flight safety. *Safety*, *4*(8), 1–15.

Ponnaluri, R. (2019). *ITE council update* (pp. 16–17). ITE J.

Rasmussen, J. (1982). Human errors – A taxonomy for describing malfunctions in industrial installations. *Journal of Occupational Accidents*, *4*(2–4), 311–333.

Rasmussen, J. (1983). Skills, rules and knowledge: Signals, signs, and symbols, and other distinctions in human performance models. *IEEE Transactions on Systems, Man, and Cybernetics*, *3*, 257–266.

Rasmussen, J. (1987). Cognitive control and human error mechanisms. In J. Rasmussen, K. Duncan, & J. Leplat (Eds.), *New technology and human error* (pp. 53–62). John Wiley & Sons.

Rasmussen, J. (1988). Cognitive engineering: A new profession? In L. P. Goodstein, H./. B. Anderson, & S. E. Olsen (Eds.), *Cognitive systems engineering*. John Wiley & Sons.

Rasmussen, J., & Jenssen, A. (1974). Mental procedures in real-life tasks: A case study of electronic trouble-shooting. *Ergonomics*, *17*, 293–307.

Rasmussen, J., Pejterson, A., & Goodstein, L. (1994). *Cognitive systems engineering*. John Wiley & Sons.

Raudenbush, S., & Bryk, A. (2002). *Hierarchical linear models: Application and data analysis methods*. Sage.

Ray, P., Purswell, J., & Bowen, D. (1993). Behavioral safety program: Creating a new corporate culture. *International Journal of Industrial Ergonomics*, *12*, 193–198.

Reason, J. (1990). *Human error*. Cambridge University Press.

Reason, J. (2000). Human error: models and management. *BMJ*, *320*, 768–770.

Reason, J., & Mycielska, K. (1982). *Absent minded? The psychology of mental lapses and everyday errors*. Prentice Hall.

Rexhausen, D., Pibernik, R., & Kaiser, G. (2012). Customer-facing supply chain practices—The impact of demand and distribution management on supply chain success. *J. Ops. Mgt.*, *30*(4), 269–281.

Roberts, K., & Rousseau, D. (1989). Research in nearly failure-free, high-reliability firms: Having the bubble. *IEEE Transactions on Engineering Management*, *36*(2), 132–139.

Robinson, J., Thomas, R., & Manrodt, K. (2013). Food for thought in the carrier selection decision. *Transportation Journal*, *52*(2), 277–296.

Ruchlin, H., Dubbs, N., & Callahan, M. (2004). The role of leadership in instilling a culture of safety: Lessons from the literature. *Journal of Healthcare Management*, *49*(1), 47–58.

Rusu, L. I., Rahayu, W., Torabi, T., Puersch, T., Coronafo, W., Harris, A. T., & Reed, K. (2012). Moving towards a collaborative decision support system for aeronautical data. *Journal of Intelligent Manufacturing*, *23*, 2085–2100.

Saltzman, G. M., & Belzer, M. H. (2002). The case for strengthened carrier hours-of-service regulations. *Transportation Journal*, *41* (4), 51–71.

Sanford, C., & Bhattacherjee, A. (2007). IT implementation in a developing country municipality: A sociocognitive analysis. *Journal of Global Information Management*, *15*(3), 20–42.

Schein, E. (2004). *Organizational culture and leadership* (3rd ed.). Jossey-Bass.

Scott, A., Balthrop, A., & Miller, J. W. (2021). Unintended responses to IT-enabled monitoring: The case of the electronic logging device mandate. *Journal of Operations Management*, *67*(2), 152–181.

Scott, A., & Nyaga, G. N. (2019). The effect of firm size, asset ownership and market prices on regulatory violations. *Journal of Management*, *65*(7), 685–709.

Serdarasan, S. (2013). A review of supply chain complexity drivers. *Computers & Industrial Engineering*, *66*(3), 533–540.

Shingo, S. (1986). *Zero quality control: Source inspection and the Poka yoke system*. Productivity Press.

Short, J., Boyle, L., Shackelford, S., Inderbelzen, B., & Bergoffer, G. (2007). The role of safety culture in preventing commercial motor vehicle crashes. In *Commercial truck and bus safety*. Transportation Research Board.

Simon, H. (1981). *The sciences of the artificial*. MIT Press.

Spender, J. C. (1996). Competitive advantage from tacit knowledge? Unpacking the concept and its strategic implications. In B. Mosingeon & A. Edmondson (Eds.), *Organizational learning and competitive advantage*. Sage Publications.

Speier, C., Whipple, J., Closs, D., & Voss, M. (2011). Global supply chain design considerations: Mitigating product safety and security risks. *Journal of Operations Management*, *29*, 721–736.

Sterns, A., & Keller, R. R. (1991). Human error and equipment design in the chemical industry. *Professional Safety*, *36*(5), 37–41.

Stewart, D. M., & Chase, R. B. (1999). The impact of human error on delivering service quality. *Production & Operations Management*, *8*(3), 240–263.

Stewart, D. M., & Grout, J. R. (2001). The human side of mistake-proofing. *Production & Operations Management*, *10*(4), 440–459.

Stock, G., & McFadden, K. (2017). Improving service operations: Linking safety culture to hospital performance. *Journal of Service Management*, *28*(1), 52–84.

Subit, D., Laporte, S., & Sandoz, B. (2017). Will automated driving technologies make today's effective restraint systems obsolete? *AJPH Perspectives*, *107*(10), 1590–1592.

Sutcliffe, A., & Rugg, G. (1998). A taxonomy of error types of failure analysis and risk assessment. *International Journal of Human-Computer Interaction*, *10*(4), 381–405.

Swartz, S., Douglas, M., Roberts, M., & Overstreet, R. (2017). Leavin' on my mind: Influence of safety culture on truck drivers' attitudes and intention to leave. *Transportation Journal*, *56*(2), 184–209.

Taft, B., & Reynolds, S. (1994). *Learning from disasters*. Butterworth-Heinemann.

Taylor, R. G., & Robinson, S. L. (2015). An information system security breach at first freedom credit union: What goes in must come out. *Journal of International Academy for Case Studies*, *21*(1), 131–138.

Thomas, T. A., & Ray, K. (2019). Online outsourcing and the future of work. *Journal of Global Responsibility*, *10*(3), 226–238.

Thornton, R. H. (2004). *Markets from cultures: Institutional logics and organizational development in higher education publishing*. Stanford University Press.

Tsai, T. L., Fridsma, D. B., & Gatti, G. (2003). Computer decision support as a source of interpretation error: The case of electrocardiograms. *Journal of the American Medical Informatics Association*, *10*(5), 478–483.

Van der Sijs, H., Lammeis, L., van den Tweel, A. J., Aarts, J., Berg, M., Vulto, A., & van Gelder, T. (2009). Time-dependent drug-drug interaction alerts in care provider order entry: Software may inhibit medication error reductions. *Journal of the American Medical Informatics Association*, *16*(6), 854–868.

Vaughan, D. (1996). *The challenger launch decision: Risky technology, culture, and deviance at NASA*. University of Chicago Press.

Waller, M., & Roberts, K. (2003). High reliability and organizational behavior: Finally, the twains must meet. *Journal of Organizational Behavior*, *24*(7), 813–814.

Weick, K. (1987). Organizational culture as a source of high reliability. *California Management Review*, *29*(2), 112–127.

Weick, K. (2004). Normal accident theory as a frame, link and provocation. *Organization & Environment*, *17*(1), 27–31.

Winter, S., Berente, N., Howison, J., & Butler, B. (2014). Beyond the organizational "container:" conceptualizing 21st century sociotechnical work. *Information and Organization*, *24*(4), 250–269.

Wu, Y., Cegielski, C. G., Hazen, B. T., & Hall, D. J. (2013). Cloud computing in support of supply chain information system infrastructure: Understanding when to go to the cloud. *Journal of Supply Chain Management*, *49*(3), 25–41.

Yusuf, M., & Sahroni, M. N. (2018). Investigating health information systems-induced errors. *International Journal of Health Care Quality Assurance*, *31*(8), 1014–1029.

Zu, X., & Kaynak, H. (2012). An agency theory perspective on supply chain management. *International Journal of Operations and Production Management*, *32*(4), 423–446.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Awaysheh, A., Frohlich, M. T., Flynn, B. B., & Flynn, P. J. (2021). To err is human: Exploratory multilevel analysis of supply chain delivery delays. *Journal of Operations Management*, *67*(7), 882–916. https://doi.org/10.1002/joom.1154

APPENDIX

Transportation management systems

Transportation management systems (TMS) are used by shipping companies and carriers for planning shipments, dispatching vehicles, and measuring performance (Caplice, 2007; Rexhausen et al., 2012). They are the interface between an ERP and warehouse/distribution module. TMS manage planning (route planning, load optimization, shipment batching, provider selection), execution (tracking, monitoring, reception, documentation), and follow-up (tracing, customs, invoicing). Figure A1 illustrates the high-level process. When a shipping company requests a shipment, in conjunction with GPS tracking, the carrier's TMS flags the closest available vehicle(s) to that shipment. If that operator is legally within his or her allowed hours of service (HOS) driving and the vehicle also has all the necessary equipment and required permits, then it is routed to the customer's point of origin to pick up the shipment. If all nearby vehicles are unavailable, then the customer's order is periodically re-planned by the TMS until a vehicle becomes available and is assigned by a dispatcher to pick up the shipment.

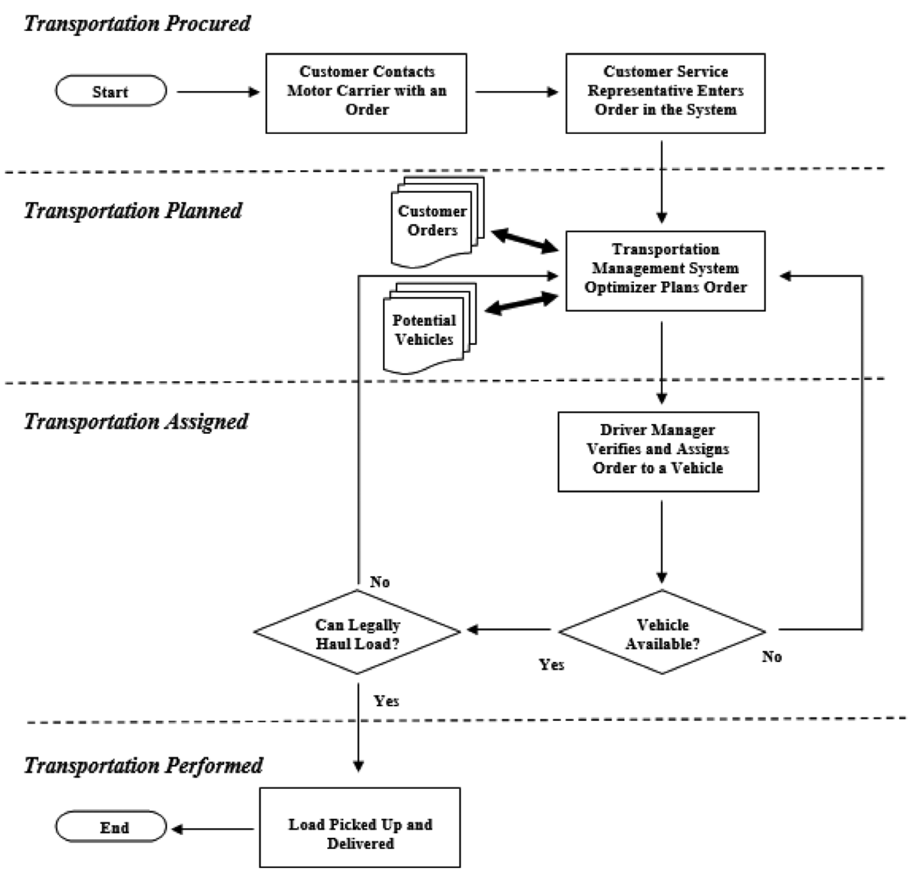

FIGURE A1 Flowchart of a generic transportation management system (TMS)