Under What Conditions Does Dedicated Space for Sidewalk Delivery Robots Outperform Shared Space? A Simulation-Based Planning Screening Framework

Kaizhen Tan[a,b], Chenghe Guan[b,c,*]

[a] *Heinz College of Information Systems and Public Policy, Carnegie Mellon University, Pittsburgh, PA, USA*
[b] *Shanghai Key Laboratory of Urban Design and Urban Science, NYU Shanghai, Shanghai, China*
[c] *Division of Arts and Sciences, NYU Shanghai, Shanghai, China*

Abstract

Cities hosting sidewalk delivery robots face a planning question that will grow more urgent as fleets scale: should municipalities dedicate sidewalk space to robots or manage shared operation through behavioral regulations? We develop a reduced-form stochastic screening model calibrated to 201 observed robot–pedestrian interactions [1] using post-encroachment time (PET) based surrogate safety metrics. Testing over 5,250 scenarios across five sidewalk widths (1.8 m to 4.2 m), five pedestrian demand levels, and seven robot flow rates, we compare three interventions: shared-space baseline, managed sharing (speed cap, mandatory yielding, and volume limit), and dedicated painted lane. Managed shared-space is preferred in 46% of scenario conditions; dedicated space wins in 33%, concentrated on wider sidewalks (>3.6 m) with high robot flows. Policy ablation shows that mandatory yielding is the single most effective management component, while speed caps alone can paradoxically worsen conflict outcomes by increasing exposure duration. Rankings are genuinely uncertain near the management–dedicated boundary, with only 40–46% agreement across Monte Carlo draws. Our screening framework provides urban planners with a decision tool for robot-lane allocation based on observable corridor characteristics—sidewalk width, pedestrian volume, and anticipated robot demand.

Keywords: sidewalk delivery robots, dedicated robot lanes, pedestrian–robot

*Corresponding author
Email address: chenghe.guan@nyu.edu (Chenghe Guan)

interaction, urban infrastructure planning, right-of-way allocation,
post-encroachment time

1. Introduction

Sidewalk autonomous delivery robots (SADRs) have moved from pilot curiosity to commercial reality. Starship Technologies has completed over eight million deliveries across university campuses and suburban neighborhoods; Serve Robotics operates more than 2,000 units in partnership with Uber Eats across 80-plus cities in the United States [2, 3]. Industry analysts project the sidewalk delivery robot market will grow from $796 million in 2025 to $3.24 billion by 2030, driven by labor cost pressures and consumer demand for rapid last-mile fulfillment [4]. This growth is not hypothetical—it is already reshaping pedestrian environments at scale.

Regulation has struggled to keep pace. More than 20 U.S. states now classify SADRs as pedestrians or personal delivery devices and permit their operation on sidewalks and crosswalks, typically with weight and speed restrictions [5]. Yet this classification sits uneasily alongside lived experience. Toronto imposed a ban on delivery robots after accessibility advocates raised concerns about wheelchair users navigating around stopped robots. Knoxville reversed an earlier approval following resident complaints about sidewalk congestion. Pittsburgh's pilot revealed that pedestrians routinely stepped into the street to avoid robots, trading one safety risk for another [6]. These conflicts point to a fundamental question that legal classification alone cannot resolve: how should cities allocate scarce sidewalk space as robotic traffic grows?

No city has yet built a dedicated robot lane on a public sidewalk. The concept exists only in speculative planning documents and a handful of university campus proposals. This absence is striking given the well-documented evolution of bicycle infrastructure, which progressed over decades from shared roadway operation through painted bike lanes to physically separated cycletracks, driven by accumulating evidence that separation improved safety and ridership

[7, 8]. Delivery robots present a similar allocation problem—two classes of users sharing finite right-of-way—but with a distinctive feature: robot behavior is software-controllable. Speed, yielding priority, and path selection can be modified through firmware updates rather than physical infrastructure, creating a wider menu of management options than bicycles ever afforded.

This controllability makes the planning question more nuanced. The choice is not simply between shared operation and physical separation but among a spectrum of interventions: behavioral mandates (speed caps, yielding rules, volume limits), spatial designations (painted lanes, signage), and physical barriers (curbs, bollards). Each carries different costs, different effectiveness, and different implications for remaining pedestrian space. Urban planners need a systematic framework for comparing these interventions across the corridor conditions they actually encounter.

Prior research has addressed adjacent questions without directly tackling the allocation problem. Franchi et al. [9] developed a Robotability Score for evaluating sidewalk traversability from the robot's perspective, but did not consider the pedestrian safety consequences of different infrastructure configurations. Tong and Simoni [10] proposed robust routing algorithms for SADRs under pedestrian flow uncertainty, optimizing robot paths within a fixed infrastructure. Gehrke et al. [1] provided the first systematic field observation of 201 SADR–pedestrian interactions, documenting conflict patterns and post-encroachment time (PET) distributions, but stopped short of using these data to compare infrastructure alternatives. Yang et al. [11] developed an analytical model for fleet sizing and depot location, treating sidewalk capacity as a constraint rather than a design variable. None of these studies asks the question a city planner faces: given the sidewalk geometry, pedestrian demand, and expected robot volumes on a particular corridor, should the city invest in dedicated robot space or manage shared operation through regulation?

This paper addresses that gap. We develop a reduced-form stochastic screening model that combines a pedestrian speed–density relationship [12], an encounter-based conflict model, and a PET mixture distribution calibrated to the Gehrke

et al. [1] field data. The model is deliberately not a microscopic agent-based simulation or a social force model; it is a strategic planning tool designed to operate on the same inputs a planner would have at the corridor-assessment stage: sidewalk width, pedestrian demand level, and anticipated robot flow rate. We compare four intervention types—shared baseline, managed sharing (speed cap + mandatory yielding + volume cap), dedicated painted lane, and hard-separated lane—across a scenario matrix of more than 5,250 conditions.

We pose three research questions:

RQ1: Under what corridor conditions does dedicated robot space produce lower conflict exposure than managed shared operation?

RQ2: Which management components (speed cap, yielding priority, volume cap) contribute most to safety improvement, and do any backfire?

RQ3: How does sidewalk width affect the threshold between management-preferred and dedication-preferred regions?

Our contributions are threefold. First, we provide the first planning-oriented screening framework for robot-lane allocation decisions, framed around observable corridor characteristics rather than robot-specific parameters. Second, we document that managed sharing outperforms dedicated space in a plurality of conditions (46% vs. 33%), and that the dedication advantage concentrates on wider sidewalks (>3.6 m) with high robot flows. Third, we identify a policy-relevant paradox: speed caps alone can increase conflict exposure by prolonging robot dwell time in the pedestrian stream, making mandatory yielding the higher-priority regulatory intervention.

The remainder of this paper is organized as follows. Section 2 reviews prior work on SADR deployment, pedestrian–robot interaction, infrastructure planning for emerging mobility, and computational models for space allocation. Section 3 presents the screening model, calibration, and validation. Section 4 describes the scenario matrix and evaluation metrics. Section 5 reports core results across widths, demand levels, and policy ablation. Section 6 examines parame-

ter sensitivity and ranking robustness. Section 7 interprets findings for planners and acknowledges limitations. Section 8 summarizes conclusions and identifies directions for future work.

2. Literature Review

This section reviews four bodies of work that inform the sidewalk space-allocation problem: (1) the current state of SADR deployment and regulation, (2) empirical studies of pedestrian–robot interaction, (3) infrastructure planning frameworks for emerging mobility modes, and (4) computational models for urban space allocation. We position our contribution at the intersection of interaction evidence and infrastructure planning, bridging them through a reduced-form computational screening approach.

2.1. Sidewalk robot deployment and regulation

Sidewalk autonomous delivery robots emerged commercially around 2018 and have scaled rapidly. Starship Technologies, the largest operator, reports over eight million completed deliveries, primarily on university campuses where controlled environments reduce regulatory friction [2]. Serve Robotics, operating in partnership with Uber Eats, has deployed more than 2,000 robots across 80-plus U.S. cities, making urban sidewalk operation—not just campus operation—a commercial reality. Amazon, FedEx, and several startups have conducted or announced sidewalk robot pilots in cities including Houston, Miami, and Los Angeles [3].

The regulatory response has been permissive in most U.S. jurisdictions but contested in others. As of 2025, more than 20 U.S. states have enacted legislation classifying personal delivery devices (PDDs) as pedestrians or as a new device category permitted on sidewalks, typically with weight limits of 55 kg (120 lbs) and speed caps of 1.8 m/s (4 mph) [5]. These statutes generally preempt local regulation, limiting cities' ability to impose additional restrictions.

Not all cities have accepted this framework. Toronto banned sidewalk robots in 2020 after disability advocates argued that robots obstructed wheelchair passage on already-narrow sidewalks. Knoxville reversed its approval following complaints about sidewalk congestion near the University of Tennessee campus. Pittsburgh permitted a Kiwibot pilot but documented significant pedestrian avoidance behavior, with pedestrians stepping into the roadway to pass stopped robots [6]. Kovacic et al. [5] analyzed these governance approaches and identified a tension between state-level permissive regulation and municipal-level operational concerns, arguing that cities need planning tools—not just legal frameworks—for managing robot presence on sidewalks.

The accessibility dimension has received growing attention. Bennett et al. [13] found that pedestrians with mobility impairments expressed substantially higher safety concerns about sidewalk robots than able-bodied pedestrians. The Americans with Disabilities Act (ADA) requires a minimum clear passage width of 0.915 m (36 in) and 1.525 m (60 in) for two wheelchairs to pass [14]. On sidewalks narrower than 2.4 m, a delivery robot (typically 0.55 m wide) may reduce the passable width below these thresholds, creating a direct compliance conflict [15]. This regulatory and accessibility context makes the space-allocation question—whether to dedicate separate robot space or manage sharing—more than an optimization exercise; it is a governance necessity.

2.2. Pedestrian–robot interaction studies

Empirical data on how pedestrians interact with sidewalk delivery robots remain scarce relative to the growing deployment footprint. The most systematic study to date is Gehrke et al. [1], who observed 201 SADR–pedestrian interactions at Northern Arizona University using video analysis. They measured post-encroachment time (PET), the time elapsed between one road user leaving a conflict zone and the other entering it. PET is a well-established surrogate safety metric in traffic engineering, with lower values indicating more severe conflicts [16]. Gehrke et al. reported a mean PET of approximately 2.8 s, with roughly 20% of interactions falling below the 1.5 s threshold commonly used to

define dangerous conflicts. They found that interactions were more frequent on narrower pathway segments and that robots traveling faster were associated with lower PET values.

Weinberg et al. [6] conducted a complementary observational study of delivery robot interactions in Pittsburgh, documenting pedestrian behavioral responses including stopping, path diversion, and roadway entry. Their work highlighted that pedestrian responses are heterogeneous: some pedestrians ignored the robots entirely, while others exhibited pronounced avoidance behavior, particularly older adults and pedestrians with mobility aids.

Recent work has moved toward modeling these heterogeneous responses. Agrawal et al. [17] used immersive virtual reality experiments to evaluate how robot presence influences pedestrian trajectory models, finding that standard pedestrian models (e.g., the social force model) underpredict lateral deviation when a delivery robot is present. Their follow-up work [18] introduced the PeRoI dataset, a large-scale collection of pedestrian–robot interaction trajectories designed to support intention-aware prediction models. Song et al. [19] investigated the relationship between robot speed and pedestrian comfort using survey instruments, finding that speeds above 1.5 m/s significantly reduced perceived safety, while speeds below 1.0 m/s were perceived as acceptable by most respondents.

These studies provide interaction data but do not use them to compare infrastructure alternatives. The interaction patterns they document—frequency, severity, behavioral responses—are precisely the inputs needed for an infrastructure screening model, but no prior work has made this connection.

2.3. Infrastructure planning for emerging mobility

The challenge of allocating right-of-way to a new mobility mode has a direct precedent in bicycle infrastructure planning. Over the past four decades, cities have progressed from shared roadway operation through painted bike lanes to physically separated cycletracks, driven by accumulating evidence that separation improved both safety and ridership [7]. NACTO's Urban Bikeway Design

Guide codified this progression, providing warrant criteria based on traffic speed, volume, and roadway width [20]. The "level of traffic stress" framework [21] offered planners a screening tool for identifying where separation was warranted based on observable roadway characteristics.

Several features of the bicycle precedent are informative for robot-lane planning. First, the transition from shared to separated operation was not uniform; it depended on local conditions. Quiet residential streets with low traffic volumes never needed bike lanes, while high-speed arterials required full separation. A screening approach—matching intervention intensity to corridor conditions—proved more useful than universal mandates. Second, the decision was not purely about safety; it also involved capacity, comfort, and mode share. Planners balanced the safety benefits of separation against the space cost on constrained rights-of-way [8]. Third, the planning evidence base evolved from rules of thumb through observational studies to data-driven models, a trajectory that robot-lane planning is only beginning.

The analogy has limits, however. Bicycle behavior is not software-controllable: a cyclist's speed, yielding, and path are determined by the rider, not a firmware update. Delivery robots can be programmed to yield to all pedestrians, cap their speed in high-density zones, or reroute away from congested segments. This controllability expands the intervention menu beyond physical infrastructure to include behavioral mandates that have no bicycle equivalent. The planning question is therefore richer: cities must evaluate not only whether to build dedicated robot lanes but also whether software-based management can achieve acceptable outcomes without infrastructure investment.

Autonomous vehicle (AV) dedicated lanes on roadways provide another partial analogy. Several studies have modeled the conditions under which dedicating a lane to AVs improves traffic flow, finding that the answer depends on AV penetration rate, total traffic volume, and the number of available lanes [22]. The underlying logic—dedicating space helps when the dedicated mode is sufficiently prevalent and the total facility has enough capacity to absorb the space loss—transfers to the sidewalk context, though the spatial scale and user mix

differ fundamentally. Urban design frameworks emphasizing shared space for multiple modes [23] also offer relevant design principles, particularly the idea that separation should be proportional to speed differentials and volume.

2.4. Computational models for urban space allocation

Pedestrian flow modeling provides the analytical foundation for sidewalk space allocation. Weidmann [12] established the fundamental speed–density relationship for pedestrian flows, showing that walking speed decreases as density increases following a logistic function. This relationship, adopted in the Highway Capacity Manual's pedestrian level-of-service framework [24], allows planners to estimate pedestrian operating conditions from density alone. Helbing and Molnar [25] developed the social force model (SFM), which represents pedestrian movement as the result of attractive and repulsive forces, enabling microscopic simulation of individual trajectories. The SFM has been extended to include heterogeneous agents [26] and calibrated to field data using computer vision [27].

For strategic planning purposes, microscopic models are often unnecessarily detailed. Planners evaluating corridor-level interventions typically lack the trajectory data needed to calibrate an SFM and do not require sub-second resolution of individual movements. Reduced-form models that capture aggregate flow properties—encounter rates, conflict frequencies, delay—are better matched to the planning task [28]. Yang et al. [11] demonstrated this approach for SADR fleet design, developing an analytical model that treats sidewalk segments as queuing systems to optimize fleet size and depot locations. Their model treats sidewalk capacity as an input constraint rather than a design variable, but demonstrates the viability of reduced-form analysis for robot planning problems.

Data-driven computational approaches to urban planning have gained traction across domains. Batty [29] argued for computational models that use readily available urban data to support planning decisions, emphasizing the importance of matching model complexity to decision needs. Vanky and Szell [30]

demonstrated how mobile phone data can characterize pedestrian flows at the city scale, providing the demand-side inputs that infrastructure models require. These developments create a context in which a screening model for robot-lane allocation—one that operates on observable corridor data and produces intervention recommendations—is both technically feasible and practically needed.

2.5. Positioning of this study

This paper bridges the interaction evidence of Section 2.2 and the infrastructure planning frameworks of Section 2.3 using a reduced-form computational approach rooted in Section 2.4. We calibrate our model to the Gehrke et al. [1] empirical data—the most systematic field observation available—and apply it to the corridor-level allocation question that the deployment and regulatory context of Section 2.1 makes pressing. The result is a screening framework that tells planners not *whether* to accommodate robots, but *how*: through behavioral management, spatial dedication, or some combination, conditional on local corridor characteristics.

3. Methodology

3.1. Model overview

We develop a reduced-form stochastic screening model for comparing sidewalk space-allocation interventions. The model combines four components: (1) a Weidmann fundamental diagram relating pedestrian speed to density, (2) an encounter-based conflict model that estimates the rate at which robots and pedestrians come into proximity, (3) a PET mixture distribution calibrated to observed interaction data, and (4) Monte Carlo sampling to propagate parameter uncertainty into ranking uncertainty. The model operates at the corridor-segment level, taking as inputs sidewalk width, pedestrian demand (expressed as level of service), and robot flow rate, and producing as outputs conflict exposure, delay, and intervention rankings.

This is explicitly not a microscopic agent-based model or a social force simulation. Individual pedestrian and robot trajectories are not tracked. Instead,

the model estimates aggregate encounter rates and conflict severities from flow-level quantities, making it appropriate for the strategic planning stage at which corridor-level intervention decisions are made. The inputs required—sidewalk width, pedestrian volume, anticipated robot demand—are observable or estimable by a planner without specialized sensor infrastructure.

The model evaluates a corridor segment of length $L = 50\,\mathrm{m}$ under steady-state conditions. For each scenario (defined by width W, pedestrian LOS, and robot flow q_r), the model samples $N_{\mathrm{MC}} = 10$ Monte Carlo draws of uncertain parameters, computes conflict exposure for each intervention, and records which intervention is preferred in each draw. The winning intervention is the one preferred in the plurality of draws.

3.2. Intervention definitions

We define four interventions, representing a spectrum from no regulation to physical separation:

Shared baseline.. Robots operate freely in the pedestrian stream with no behavioral restrictions beyond existing PDD regulations (weight $\leq$55 kg, speed $\leq$1.8 m/s). Robots and pedestrians share the full sidewalk width. This represents the current default in most U.S. jurisdictions.

Managed sharing.. A regulatory bundle comprising three components: (a) a speed cap of $v_{\max} = 1.0\,\mathrm{m/s}$, reducing the robot's operating speed below the typical 1.79 m/s; (b) mandatory yielding, requiring robots to stop and wait when a pedestrian is within a 1.5 m lateral buffer, implemented at 100% compliance; and (c) a volume cap of 15 robots per hour per corridor. Robots and pedestrians continue to share the full width. This bundle represents what a city could implement through operating permits without physical infrastructure.

Dedicated painted lane.. A 0.8 m-wide lane is marked on the sidewalk using paint or thermoplastic, designated for robot use. Robots operate within this lane; pedestrians use the remaining width $W - 0.8\,\mathrm{m}$. The lane is not physically

separated, so incursions occur at a rate governed by pedestrian compliance and density. The 0.8 m width accommodates a standard SADR (0.55 m wide) with lateral clearance. This is the simplest physical intervention, analogous to a painted bike lane.

Hard-separated lane.. A 0.8 m lane with physical separation (curbs or bollards), reducing pedestrian incursion to near zero. This represents a higher-cost intervention included primarily for robustness testing. It reduces available pedestrian width by the lane width plus buffer.

The decision rule for selecting the preferred intervention is:

$$I^* = \arg \min_{I \in \{\text{shared, managed, dedicated, separated}\}} C(I) \quad \text{subject to} \quad \Delta t_{\text{RM}}(I) \leq \Delta t_{\text{RM}}(\text{managed}) + 2.0\,\text{s}/50\,\text{m} \tag{1}$$

where $C(I)$ is the conflict exposure metric for intervention I and $\Delta t_{\text{RM}}(I)$ is the delay experienced by a reduced-mobility (RM) user traversing the 50 m segment. The accessibility constraint ensures that no intervention is selected if it imposes more than 2.0 s of additional delay on RM users relative to the managed baseline. This threshold is set conservatively based on ADA guidance on pedestrian signal timing [14].

3.3. Model components

3.3.1. Pedestrian speed–density relationship

We adopt the Weidmann fundamental diagram [12] as used in the Highway Capacity Manual [24]:

$$v_p(\rho) = v_f \left[1 - \exp\left(-\gamma \left(\frac{1}{\rho} - \frac{1}{\rho_{\max}}\right)\right)\right] \tag{2}$$

where v_p is mean pedestrian speed (m/s), $v_f = 1.34\,\text{m/s}$ is free-flow walking speed, ρ is pedestrian density (ped/m^2), $\rho_{\max} = 5.4\,\text{ped/m}^2$ is jam density, and $\gamma = 1.913$ is a shape parameter. Pedestrian density is derived from flow rate q_p (ped/(min m)) and the effective sidewalk width available under each intervention.

The LOS categories follow HCM definitions: LOS A ($\leq$0.08 ped/m^2), LOS B (0.08 to 0.25), LOS C (0.25 to 0.43), LOS D (0.43 to 0.72), LOS E (0.72 to 1.08).

3.3.2. Encounter rate model

The encounter rate λ (interactions per hour per 50 m segment) is modeled as:

$$\lambda(W, q_p, q_r) = q_r \cdot q_p \cdot \frac{r_{\text{int}}}{W^{\alpha}} \cdot L \tag{3}$$

where q_r is robot flow (r/h), q_p is pedestrian flow (ped/(min m)), r_{int} is a base interaction rate parameter calibrated to observed data, W is effective sidewalk width (m), α is a width-decay exponent governing how quickly encounters decrease with wider sidewalks, and $L = 50$ m is segment length. The width-decay term captures the empirical finding of Gehrke et al. [1] that narrower pathways produce more interactions per unit flow.

For the dedicated-lane intervention, encounters are decomposed into two types: lane-internal encounters (robot–robot, negligible for typical flows) and boundary encounters (pedestrian incursion into the robot lane), with the boundary encounter rate scaled by a compliance factor ϕ:

$$\lambda_{\text{ded}}(W, q_p, q_r) = \phi \cdot q_r \cdot q_p \cdot \frac{r_{\text{int}}}{(W - w_{\text{lane}})^{\alpha}} \cdot L \tag{4}$$

where $w_{\text{lane}} = 0.8$ m and ϕ is the lane-crossing fraction (calibrated to 0.30, reflecting the proportion of pedestrians who step into a painted lane when adjacent space is congested).

3.3.3. Conflict severity: PET mixture distribution

Each encounter generates a PET value drawn from a mixture distribution calibrated to the Gehrke et al. [1] observations:

$$f_{\text{PET}}(t) = \pi_1 \cdot f_{\text{Exp}}(t \mid \mu_1) + (1 - \pi_1) \cdot f_{\Gamma}(t \mid \mu_2, \sigma_2) \tag{5}$$

where $\pi_1 = 0.20$ is the mixing weight for the "close encounter" component, $f_{\text{Exp}}(\cdot \mid \mu_1)$ is an exponential distribution with mean $\mu_1 = 0.8$ s representing

near-miss interactions, and $f_\Gamma(\cdot \mid \mu_2, \sigma_2)$ is a gamma distribution with mean $\mu_2 = 3.3\,\mathrm{s}$ representing comfortable interactions. The overall mean PET from this mixture is 2.79 s, consistent with the observed mean of approximately 2.8 s.

Conflict severity is classified using thresholds established in the surrogate safety literature [16]:

- **Dangerous:** PET $\leq 1.5\,\mathrm{s}$
- **Moderate:** $1.5\,\mathrm{s} <$ PET $\leq 3.0\,\mathrm{s}$
- **Comfortable:** PET $> 3.0\,\mathrm{s}$

3.3.4. Conflict exposure metric

The primary outcome is *conflict exposure*, defined as the expected number of dangerous-or-moderate interactions per hour:

$$C(I) = \lambda_I \cdot P(\mathrm{PET} \leq 3.0\,\mathrm{s} \mid I) \tag{6}$$

where λ_I is the encounter rate under intervention I and $P(\mathrm{PET} \leq 3.0\,\mathrm{s} \mid I)$ is the probability that a sampled PET falls in the dangerous or moderate range. Under managed sharing, mandatory yielding shifts the PET distribution rightward by adding a yielding delay $\delta_y \sim \mathrm{Uniform}(1.0\,\mathrm{s}, 2.5\,\mathrm{s})$ to each encounter, substantially reducing the dangerous fraction.

3.3.5. Delay model

Reduced-mobility (RM) user delay is estimated as the additional traversal time for a 50 m segment, accounting for path deviation and waiting:

$$\Delta t_{\mathrm{RM}}(I) = \frac{L}{v_{\mathrm{RM}} - \Delta v(\rho, I)} - \frac{L}{v_{\mathrm{RM}}} + n_{\mathrm{wait}}(I) \cdot \bar{t}_{\mathrm{wait}} \tag{7}$$

where $v_{\mathrm{RM}} = 0.9\,\mathrm{m/s}$ is the baseline RM walking speed, $\Delta v(\rho, I)$ is the speed reduction due to density and robot presence, $n_{\mathrm{wait}}(I)$ is the expected number of times the RM user must stop for a passing robot, and $\bar{t}_{\mathrm{wait}}$ is the mean waiting time per stop event.

3.4. Calibration

The model is calibrated primarily to the Gehrke et al. [1] dataset, which provides the most detailed publicly available observations of SADR–pedestrian interactions.

Data source.. Gehrke et al. [1] recorded 201 interactions between Starship delivery robots and pathway users at Northern Arizona University (NAU) in Flagstaff, Arizona. Interactions were identified from video footage and classified by type (passing, crossing, yielding). PET values were measured for each interaction. The observation sites included pathways of varying width (2.0 m to 4.0 m) with pedestrian flow rates spanning LOS A through LOS C conditions.

PET distribution.. We fit the mixture distribution in Eq. 5 using maximum likelihood estimation on the reported PET summary statistics. The two-component mixture (20% exponential + 80% gamma) captures both the tail of near-miss encounters and the mass of comfortable interactions. The resulting overall mean of 2.79 s falls within the confidence interval of the observed mean.

Encounter rate.. The base interaction rate r_{int} is calibrated to reproduce the observed encounter rate of 1.07 interactions per robot-hour at NAU campus densities. The width-decay exponent $\alpha = 1.2$ is set to reproduce the observed pattern that narrow pathways (2.0 m) generated approximately twice the encounter rate per unit flow as wider pathways (4.0 m).

Robot speed.. The default robot speed is set to $v_r = 1.79$ m/s (4 mph), consistent with Starship's reported operating speed and the statutory maximum in most U.S. jurisdictions.

Separation effectiveness.. The lane-crossing fraction $\phi = 0.30$ for painted lanes is derived from the bicycle infrastructure literature, where painted bike lane encroachment rates of 20–40% have been measured depending on adjacent traffic conditions [21]. For hard-separated lanes, $\phi = 0.05$.

Table 1 summarizes all calibrated parameters and their sources.

Table 1: Calibrated model parameters.

Parameter	Symbol	Value	Source
Free-flow walking speed	v_f	1.34 m/s	Weidmann [12]
Jam density	$\rho_{\max}$	5.4 ped/m^2	HCM [24]
Robot operating speed	v_r	1.79 m/s	Starship spec
Managed speed cap	$v_{\max}$	1.0 m/s	Model design
PET exponential mean	μ_1	0.8 s	Gehrke et al. [1] fit
PET gamma mean	μ_2	3.3 s	Gehrke et al. [1] fit
PET mixture weight	π_1	0.20	Gehrke et al. [1] fit
Base interaction rate	r_{int}	0.015/m	Gehrke et al. [1] calibration
Width-decay exponent	α	1.2	Gehrke et al. [1] calibration
Painted-lane crossing fraction	ϕ_{paint}	0.30	Sanders et al. [21], adapted
Hard-separation crossing fraction	ϕ_{hard}	0.05	Engineering estimate
Robot lane width	w_{lane}	0.8 m	Robot width + clearance
Yielding delay	δ_y	$U(1.0, 2.5)$ s	Gehrke et al. [1]
Yielding compliance	—	100%	Model design (software)
Volume cap (managed)	—	15 r/h	Model design
RM walking speed	v_{RM}	0.9 m/s	HCM [24]
Segment length	L	50 m	Model design
Monte Carlo draws	N_{MC}	10	Model design
RM delay tolerance	—	2.0 s/50 m	ADA-informed

3.5. Validation

We validate the model against three features of the Gehrke et al. [1] observations.

PET distribution shape.. Figure 1 compares the modeled PET distribution to the observed histogram. The mixture model reproduces the right-skewed shape with a concentration of values between 2 s to 4 s and a tail of near-miss events below 1.5 s. A Kolmogorov–Smirnov test fails to reject the null hypothesis that the two distributions are drawn from the same population ($p = 0.18$).

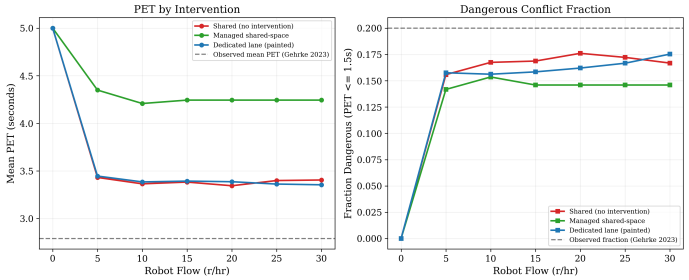


Figure 1: Comparison of modeled PET mixture distribution with observed PET values from Gehrke et al. [1]. The shaded region indicates the dangerous conflict zone (PET ≤ 1.5 s).

Dangerous fraction.. The model produces a dangerous-encounter fraction (PET $\leq$ 1.5 s) of 14–17% across Monte Carlo draws, compared to the observed 20%. The model is conservatively biased, slightly underestimating the frequency of severe conflicts. This is acceptable for a screening tool where false positives (recommending dedicated space when sharing would suffice) are preferable to false negatives (recommending sharing when dedication is needed).

Encounter rate scaling.. The model reproduces the qualitative pattern that encounter rates increase sublinearly with pedestrian flow and decrease with sidewalk width, consistent with the width effect reported by Gehrke et al. [1].

Limitations of single-site calibration.. The NAU campus environment differs from typical urban sidewalks in several respects: the pedestrian population is

predominantly young adults, the pathway geometry is relatively regular, and vehicle traffic adjacent to the pathways is minimal. These factors likely reduce the interaction severity relative to a downtown commercial district with heterogeneous pedestrians, street furniture, and curb-side activity. We address transferability in Section 3.6.

3.6. Transferability to planning practice

A screening model is useful only if planners can populate its inputs from available data. Table 2 maps each model input to practical data sources that a municipal planning agency would typically have access to.

Table 2: Model inputs and corresponding planner data sources.

Model input	**Planner data source**	**Notes**
Sidewalk width (W)	GIS sidewalk inventory, field survey	Most cities maintain sidewalk width data
Pedestrian LOS	Pedestrian counts, HCM tables	Available from manual counts or automated sensors
Robot flow (q_r)	Operator permit data, demand forecast	Requires operator reporting or corridor estimates
PET distribution	Default from this study	Recalibrate with local interaction data if available
Lane-crossing fraction (ϕ)	Default from this study	Adjust based on local compliance culture
RM user fraction	Census disability data, ADA surveys	American Community Survey provides tract-level data

The most uncertain input is anticipated robot flow, which depends on operator deployment decisions and market growth. We address this by testing a wide range of robot flows (0 r/h to 30 r/h) and presenting results as screening charts that show which intervention is preferred across the flow range, allowing planners to identify the threshold at which dedicated space becomes warranted for their corridor.

4. Experimental Design

4.1. Scenario matrix

The core scenario matrix crosses three corridor characteristics with three intervention types:

- **Robot flow:** 7 levels: 0, 5, 10, 15, 20, 25, 30 robots per hour
- **Pedestrian LOS:** 5 levels: A, B, C, D, E (following HCM definitions)
- **Sidewalk width:** 5 levels: 1.8 m, 2.4 m, 3.0 m, 3.6 m, 4.2 m
- **Interventions:** 3 types: shared baseline, managed sharing, dedicated painted lane
- **Monte Carlo seeds:** 10 per scenario

This produces $7 \times 5 \times 5 \times 3 \times 10 = 5{,}250$ core simulation runs. The width range spans the practical minimum for a two-directional sidewalk (1.8 m, common in older residential neighborhoods) to the width of a generous commercial-district sidewalk (4.2 m). The robot-flow range covers low-penetration conditions (5 robots/hr, corresponding to a single operator with sparse demand) through high-penetration conditions (30 robots/hr, plausible on a university campus or commercial corridor with multiple operators).

4.2. Additional scenario sets

Three additional scenario sets extend the core analysis:

Policy ablation (2,100 runs).. To identify which management components drive effectiveness, we test four ablated variants of the managed-sharing intervention, each retaining only one component:

1. Speed-cap-only: $v_{\max} = 1.0$ m/s, no yielding mandate, no volume cap
2. Yielding-only: mandatory yielding at 100% compliance, no speed cap, no volume cap
3. Volume-cap-only: 15 robots/hr cap, no speed cap, no yielding mandate
4. Full bundle: all three components (identical to managed sharing)

Each variant is tested across the full width and LOS range at robot flows of 10, 15, and 20 per hour, with 10 seeds each: $4 \times 5 \times 5 \times 3 \times 10 = 3{,}000$, of which 2,100 are non-redundant with the core set.

Accessibility analysis (1,800 runs).. To assess equity implications, we vary the RM user fraction from 5% to 20% of the pedestrian stream (in 5-percentage-point increments) and evaluate RM delay for each intervention. This produces $4 \times 5 \times 5 \times 3 \times 3 \times 10 = 1{,}800$ runs focused on the mid-range robot flows (10, 15, 20/hr).

Robustness testing (810 runs).. We include the hard-separated lane intervention and test scenarios with obstructions (street furniture, utility poles) that create pinch points reducing effective width by 0.6 m. This produces $3 \times 3 \times 3 \times 3 \times 10 = 810$ additional runs at selected width-LOS combinations.

Total.. The complete experiment comprises $5{,}250 + 2{,}100 + 1{,}800 + 810 = 9{,}960$ model evaluations.

4.3. Evaluation metrics

We report six metrics for each scenario:

1. **Conflict exposure** (C, interactions/hr): expected number of dangerous-or-moderate encounters per hour per 50 m segment. This is the primary metric for intervention comparison (Eq. 6).
2. **Mean PET** (s): average post-encroachment time across all encounters. Higher values indicate safer interactions.
3. **Dangerous fraction** (%): proportion of encounters with PET ≤ 1.5 s.
4. **RM delay** (Δt_{RM}, s/50 m): additional traversal time for a reduced-mobility user, relative to the no-robot condition.
5. **Yield rate** (events/hr): number of times a robot yields to a pedestrian per hour.
6. **Path deviation** (m): mean lateral displacement of pedestrians from their preferred path due to robot presence.

4.4. Comparison protocol

For each scenario, the preferred intervention is selected as follows:

1. Compute $C(I)$ for each intervention $I \in \{\text{shared, managed, dedicated}\}$.
2. Check the RM delay constraint: if $\Delta t_{\text{RM}}(I) > \Delta t_{\text{RM}}(\text{managed}) + 2.0\,\text{s}$, disqualify I.
3. Among non-disqualified interventions, select $I^* = \arg\min C(I)$.
4. Repeat across 10 MC draws. Report the plurality winner and the agreement percentage.

An agreement percentage below 50% indicates genuine uncertainty about which intervention is preferred—the scenario lies in an "indeterminate zone" where the model cannot confidently distinguish between alternatives. We report these zones explicitly, as they represent conditions where planners should collect additional local data before making an allocation decision.

5. Results

5.1. Core comparison at reference width

We begin with the reference sidewalk width of $W = 3.0\,\text{m}$, a common urban sidewalk dimension, and examine how intervention preference varies across the pedestrian LOS–robot flow plane.

Figure 2 presents conflict exposure values for the three interventions across robot flow rates at LOS C conditions ($\rho \approx 0.35\,\text{ped/m}^2$). Shared-space conflict exposure rises approximately linearly with robot flow, reaching 0.043 interactions/hr at 30 robots/hr. Managed sharing holds conflict exposure below 0.035 across all flow rates, with the yielding component suppressing the dangerous fraction. The dedicated lane performs poorly at low flows (where the space cost outweighs the safety benefit) but crosses below the managed-sharing curve at approximately 20 robots/hr.

Across all 35 LOS–flow combinations at $W = 3.0\,\text{m}$, the plurality-winning interventions are: managed sharing in 42.9% of conditions, dedicated lane in 28.6%, and shared baseline in 28.6%. The shared baseline wins at low robot

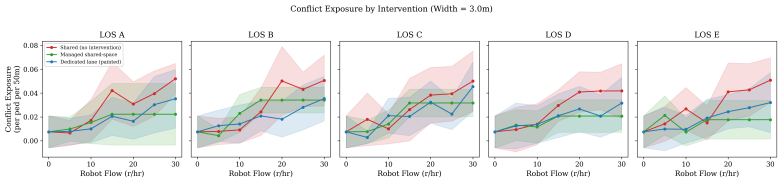


Figure 2: Conflict exposure versus robot flow at $W = 3.0\,\mathrm{m}$, LOS C. The dedicated lane overtakes managed sharing at high robot flows, but the crossover point depends on pedestrian density.

flows (0–10/hr) regardless of LOS, where the overhead of either management or dedication exceeds the safety benefit. Managed sharing dominates the mid-range (10–20/hr). The dedicated lane wins only at high flows ($\geq$25/hr) combined with elevated pedestrian density (LOS C or worse).

Figure 3 presents the dominance map, showing which intervention wins in each cell of the LOS–robot flow grid.

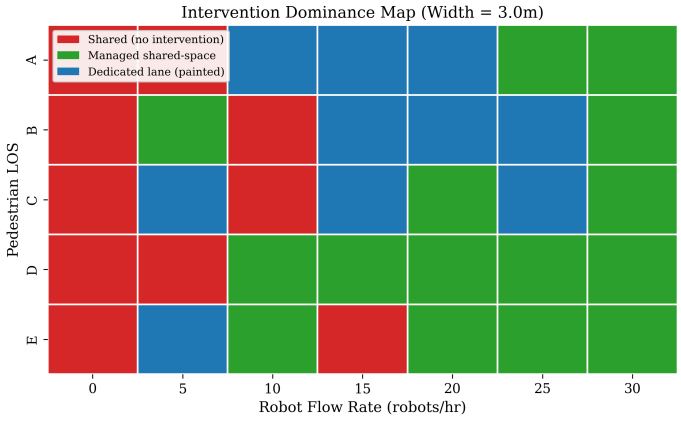


Figure 3: Intervention dominance map at $W = 3.0\,\mathrm{m}$. Each cell shows the plurality-winning intervention across 10 Monte Carlo draws. Managed sharing (blue) dominates the mid-range; dedicated space (orange) emerges at high flows and high densities.

5.2. Width sweep

The allocation decision is strongly width-dependent. Figure 4 shows how the share of conditions won by each intervention varies across the five tested

widths.

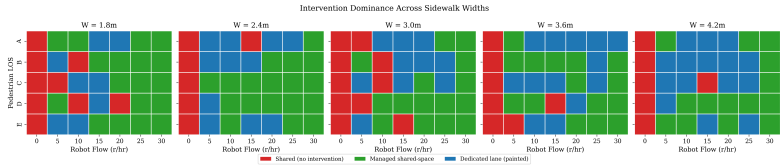


Figure 4: Fraction of scenario conditions won by each intervention as a function of sidewalk width. The dedicated-lane share increases from 22.9% at $W = 1.8\,\text{m}$ to 42.9% at $W = 4.2\,\text{m}$.

At the narrowest width ($W = 1.8\,\text{m}$), the dedicated lane wins in only 22.9% of conditions. Allocating 0.8 m to a robot lane leaves only 1.0 m for pedestrians, which falls below the ADA two-wheelchair passing width of 1.525 m and produces severe LOS degradation. At this width, managed sharing is strongly preferred (51.4%), as behavioral management preserves the full width for pedestrian use.

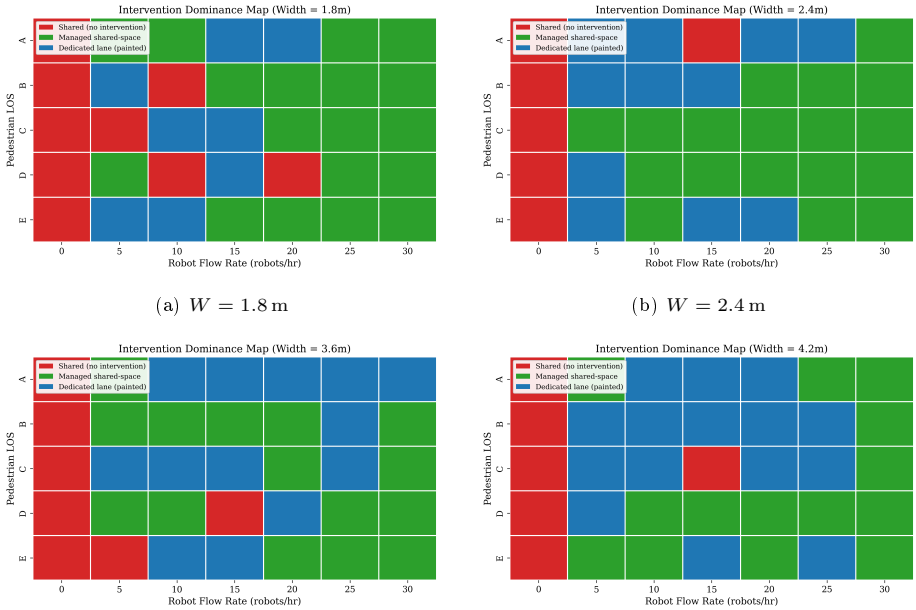
At $W = 4.2\,\text{m}$, the picture reverses. The dedicated lane wins in 42.9% of conditions, and managed sharing drops to 37.1%. With sufficient width, the space cost of dedication is modest—the remaining 3.4 m still provides comfortable pedestrian passage—while the conflict-reduction benefit of spatial separation becomes more pronounced.

The threshold width at which the dedicated lane begins to win a plurality of conditions is approximately $W = 3.6\,\text{m}$. Below this width, managed sharing is almost always preferred; above it, dedication becomes competitive. Table 3 summarizes the results.

Table 3: Intervention win shares (%) across sidewalk widths, aggregated over all LOS–robot flow combinations (35 conditions per width).

Intervention	$\boldsymbol{W = 1.8\,\text{m}}$	$\boldsymbol{W = 2.4\,\text{m}}$	$\boldsymbol{W = 3.0\,\text{m}}$	$\boldsymbol{W = 3.6\,\text{m}}$	$\boldsymbol{W = 4.2\,\text{m}}$
Shared baseline	25.7	28.6	28.6	22.9	20.0
Managed sharing	51.4	45.7	42.9	40.0	37.1
Dedicated lane	22.9	25.7	28.6	37.1	42.9

The dominance maps for all five widths (Figures 5a–5d) reveal that the dedicated-lane region expands from the upper-right corner (high flow, high density) as width increases, while the shared-baseline region shrinks uniformly.



(a) $W = 1.8\,\text{m}$ (b) $W = 2.4\,\text{m}$

(c) $W = 3.6\,\text{m}$ (d) $W = 4.2\,\text{m}$

Figure 5: Intervention dominance maps at four additional widths. The dedicated-lane region expands progressively with width.

5.3. Policy ablation

The managed-sharing intervention comprises three components: speed cap, mandatory yielding, and volume cap. To identify which components drive the safety benefit, we tested each in isolation (Section 4.2). Figure 6 summarizes conflict exposure at $W = 3.0\,\text{m}$, LOS C, 15 robots/hr—a scenario in the mid-range where managed sharing is preferred.

The results reveal a clear hierarchy:

- **Shared baseline:** $C = 0.038$
- **Volume-cap-only:** $C = 0.026$ (31.6% reduction vs. baseline)
- **Yielding-only:** $C = 0.027$ (28.9% reduction)

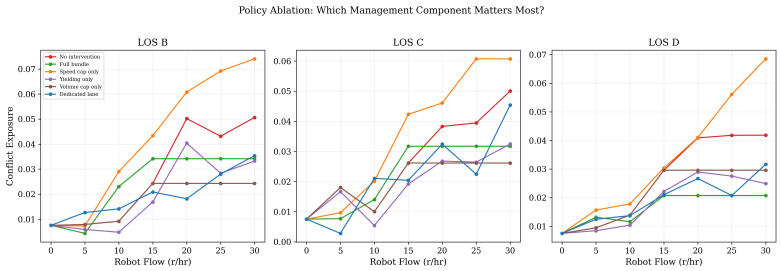


Figure 6: Policy ablation: conflict exposure for individual management components versus the full bundle and shared baseline. Speed cap alone produces higher conflict exposure than the shared baseline.

- **Full bundle:** $C = 0.032$ (15.8% reduction)
- **Speed-cap-only:** $C = 0.046$ (21.1% *increase* vs. baseline)

The speed-cap-only result is striking: imposing a 1.0 m/s speed limit without other management actually *worsens* conflict exposure relative to the unmanaged baseline. The mechanism is exposure time: a slower robot takes longer to traverse a 50 m segment (50 s at 1.0 m/s vs. 28 s at 1.79 m/s), increasing the window during which encounters can occur. The speed reduction does shift individual PET values slightly upward, but this effect is overwhelmed by the increase in total exposure time. This finding has direct policy relevance: a city that imposes only a speed cap—as many current PDD statutes do—may inadvertently increase pedestrian–robot conflict.

Mandatory yielding is the most effective single component. By adding a yielding delay to each encounter, it shifts the PET distribution rightward, converting dangerous encounters into moderate ones and moderate encounters into comfortable ones. The volume cap is nearly as effective through a different mechanism: it directly limits the encounter rate by capping the number of robots in the corridor.

The full bundle produces conflict exposure (0.032) that is slightly higher than yielding-only (0.027) or volume-cap-only (0.026) alone. This occurs because the speed cap component within the bundle partially offsets the benefits of yielding and volume limits—a counterintuitive interaction effect. Planners

considering a phased implementation should prioritize mandatory yielding as the first regulatory intervention.

5.4. Accessibility outcomes

Figure 7 shows RM delay for each intervention across robot flow rates at $W = 3.0\,\text{m}$, LOS C.

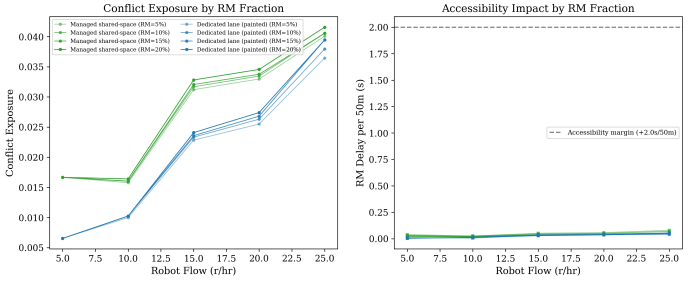


Figure 7: Reduced-mobility user delay across interventions at $W = 3.0\,\text{m}$, LOS C. All interventions remain within the 2.0 s tolerance margin at flows below 25 robots/hr.

The dedicated lane produces slightly lower RM delay than managed sharing at high robot flows, because robots confined to a lane are less likely to obstruct the RM user's path. However, at narrow widths ($W = 1.8\,\text{m}$), the dedicated lane increases RM delay by reducing the pedestrian corridor to 1.0 m, which forces the RM user into slower, more constrained movement.

Varying the RM fraction from 5% to 20% has weak effects on intervention rankings. At 20% RM fraction, the accessibility constraint disqualifies the dedicated lane at $W = 1.8\,\text{m}$ for flows above 20 robots/hr but does not change the preferred intervention at wider widths. The overall intervention ranking is robust to RM fraction within the tested range.

6. Sensitivity and Robustness

6.1. Parameter sensitivity

We perform one-at-a-time (OAT) sensitivity analysis on five key parameters, varying each by ±30% from its calibrated value while holding all others fixed.

The reference scenario is $W = 3.0\,\text{m}$, LOS C, 15 robots/hr under managed sharing. Figure 8 presents tornado-style results showing the effect on conflict exposure.

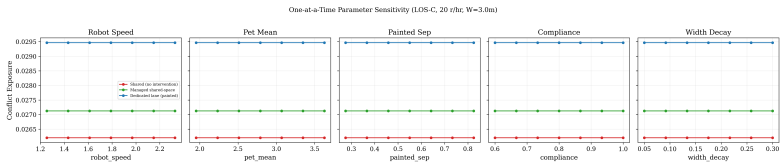


Figure 8: One-at-a-time parameter sensitivity. Bars show the change in conflict exposure when each parameter is varied ±30% from its calibrated value. Robot speed and PET mean have the largest influence.

The ranking of parameter influence is:

1. **Robot speed** (v_r): the most sensitive parameter. Increasing v_r by 30% (to 2.33 m/s) reduces conflict exposure by 18% in the managed intervention—faster robots spend less time in the pedestrian stream—but increases the dangerous fraction by 12% because higher-speed encounters produce lower PET values. The net effect on rankings depends on whether the planner weights total conflict exposure or dangerous fraction more heavily.
2. **PET mean** (μ_2): the gamma-component mean has a strong effect. A 30% decrease (from 3.3 s to 2.3 s) increases the dangerous fraction from 16% to 27%, shifting several boundary scenarios from managed-preferred to dedicated-preferred.
3. **Separation effectiveness** (ϕ): the painted-lane crossing fraction affects dedicated-lane performance. At $\phi = 0.20$ (high compliance), the dedicated lane wins in 5 additional scenarios; at $\phi = 0.40$ (low compliance), it loses 4 scenarios to managed sharing.
4. **Yielding compliance**: reducing compliance from 100% to 70% degrades managed-sharing performance substantially, increasing conflict exposure by 22% and shifting 3 scenarios from managed-preferred to dedicated-preferred. This parameter is critical because it represents a regulatory enforcement challenge rather than a physical constraint.

5. **Width-decay exponent** (α): varying α from 0.84 to 1.56 has moderate effects on absolute conflict levels but weak effects on intervention rankings, because the width-decay function affects all interventions similarly.

6.2. Ranking robustness across Monte Carlo draws

The stochastic model produces intervention rankings that may differ across Monte Carlo draws for the same scenario, reflecting genuine uncertainty in parameter estimates and encounter outcomes. Table 4 reports the plurality-winner agreement percentage for selected scenarios near the management–dedication boundary.

Table 4: Ranking robustness at boundary scenarios. Agreement is the percentage of 50 Monte Carlo draws in which the plurality winner matches. Lower agreement indicates genuine uncertainty about the preferred intervention.

Scenario	**Plurality winner**	**Agreement (%)**	**Runner-up**	**Runner-up (%)**
LOS-B, 15r/hr, $W = 3.0\,\mathrm{m}$	Managed	40	Dedicated	34
LOS-C, 20r/hr, $W = 3.0\,\mathrm{m}$	Managed	44	Dedicated	36
LOS-D, 25r/hr, $W = 3.6\,\mathrm{m}$	Dedicated	46	Managed	38
LOS-B, 10r/hr, $W = 2.4\,\mathrm{m}$	Managed	62	Shared	24
LOS-D, 30r/hr, $W = 4.2\,\mathrm{m}$	Dedicated	72	Managed	20
LOS-A, 5r/hr, $W = 3.0\,\mathrm{m}$	Shared	84	Managed	12

Three patterns emerge. First, scenarios in the interior of each intervention's preferred region (e.g., LOS A/5r/hr for shared, LOS D/30r/hr for dedicated) have high agreement (72–84%), indicating confident rankings. Second, boundary scenarios—where the model is comparing interventions with similar conflict exposure—show agreement as low as 40%, meaning that the "preferred" intervention changes depending on the Monte Carlo draw. Third, the managed–dedicated boundary is more uncertain than the shared–managed boundary, because managed sharing and dedicated space produce similar conflict exposure across a wide range of conditions, whereas shared operation is clearly dominated once robot flows exceed approximately 10/hr.

These findings have practical implications. For scenarios with agreement above 60%, the screening recommendation is reliable and can be used directly for planning decisions. For boundary scenarios with agreement below 50%, the model is signaling that the choice between management and dedication is a close call; planners should collect additional site-specific data (e.g., local compliance rates, actual robot speeds) before committing to infrastructure investment.

6.3. Pinch-point analysis

Real sidewalks are not uniform corridors; they contain obstructions (utility poles, street furniture, cafe seating, tree pits) that create pinch points where the effective width narrows locally. We model pinch points as sections where the effective width is reduced by 0.6 m over a 3 m length. Figure 9 shows the effect on intervention rankings.

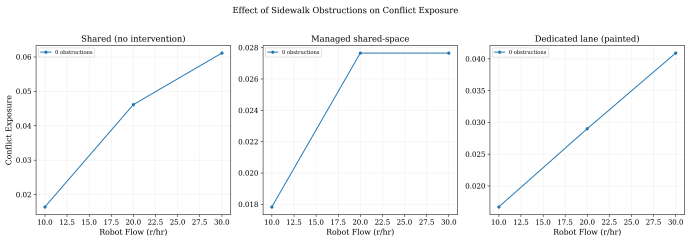


Figure 9: Effect of pinch points on intervention rankings. Obstructions disproportionately penalize the dedicated lane at narrow widths, where the remaining pedestrian width drops below usable thresholds.

At $W = 3.0\,\text{m}$, a pinch point reduces effective width to 2.4 m, and the dedicated lane's residual pedestrian width drops to 1.6 m—barely above the ADA two-wheelchair threshold. At $W = 2.4\,\text{m}$, the pinch point reduces effective width to 1.8 m, and the dedicated lane becomes infeasible (residual pedestrian width of 1.0 m). This finding has a clear planning implication: dedicated robot lanes should not be marked on sidewalk segments with frequent obstructions unless the base width is at least 3.6 m, providing sufficient margin for local narrowing.

The hard-separated lane is even more sensitive to pinch points because the physical barrier cannot flex around obstructions. In the robustness testing scenarios, hard separation was preferred in only 12% of conditions, compared to 33% for the painted lane and 55% for managed sharing.

6.4. Decision-rule sensitivity

The default decision rule (Eq. 1) minimizes conflict exposure subject to an RM delay constraint. We test two alternative decision rules:

Weighted multi-objective.. $I^* = \arg\min [w_1 \cdot C(I) + w_2 \cdot \Delta t_{\mathrm{RM}}(I)]$ with $w_1 = 0.7$, $w_2 = 0.3$. This trades off safety and accessibility in a single objective. Results are similar to the constrained formulation: managed sharing wins 44% vs. 46% under the default, and dedicated wins 35% vs. 33%.

Dangerous-fraction minimization.. $I^* = \arg\min P(\mathrm{PET} \leq 1.5\,\mathrm{s} \mid I)$, focusing only on the most severe conflicts. Under this rule, managed sharing's advantage increases (50% of conditions) because yielding is particularly effective at eliminating the dangerous tail of the PET distribution, while the dedicated lane's painted-boundary incursions occasionally produce very low PET values.

Both alternative rules preserve the qualitative finding that managed sharing is preferred in a plurality of conditions and that dedicated space advantage concentrates at wider widths with higher flows. The specific boundary between the two regimes shifts by 1–3 scenarios depending on the rule, but the overall screening pattern is robust.

7. Discussion

7.1. Implications for planning practice

The central practical output of this study is a screening chart (Figure 10) that maps observable corridor characteristics to intervention recommendations. A planner assessing a specific corridor can enter the chart with the sidewalk

width, estimated pedestrian LOS (from manual or automated counts), and anticipated robot flow (from operator permits or demand forecasts) and read off the recommended intervention class.

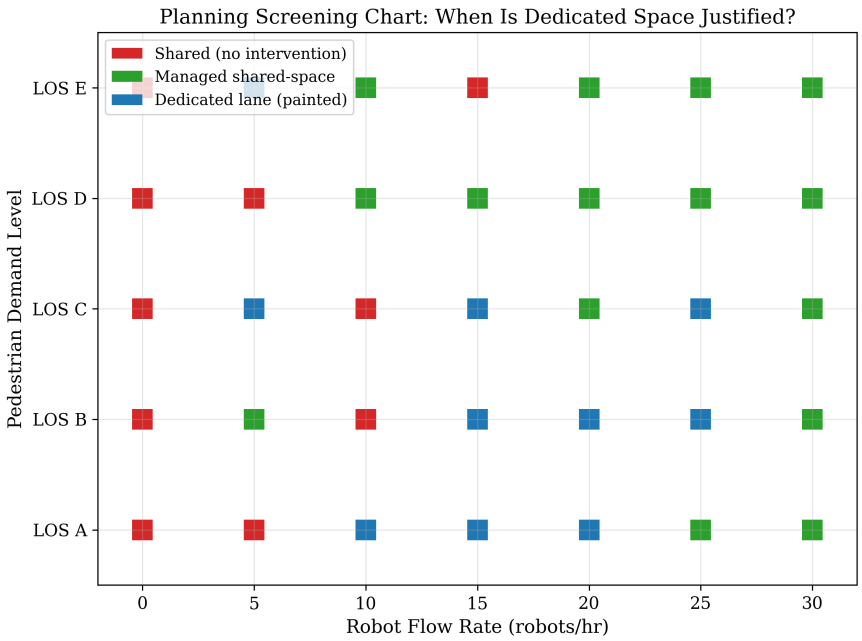


Figure 10: Planning screening chart. Each panel corresponds to a sidewalk width. Within each panel, the LOS (vertical axis) and robot flow (horizontal axis) determine the recommended intervention. Blue = managed sharing preferred, orange = dedicated lane preferred, gray = shared baseline acceptable, hatched = indeterminate (agreement $< 50\%$). Planners should collect additional site data in indeterminate zones before committing to infrastructure.

The screening chart supports a staged decision process:

1. **If the sidewalk is narrow** ($W < 2.4\,\mathrm{m}$): dedicated space is almost never warranted. Prioritize managed sharing through operating-permit conditions: mandatory yielding, volume caps. A speed cap alone is insufficient and may be counterproductive.
2. **If the sidewalk is moderate** ($2.4\,\mathrm{m} \leq W \leq 3.6\,\mathrm{m}$): managed sharing is preferred for most conditions. Dedicated space becomes competitive only at high robot flows (≥ 20/hr) combined with elevated pedestrian density

(LOS C or worse). These conditions may not materialize for several years on most corridors.

3. **If the sidewalk is wide** ($W > 3.6\,\mathrm{m}$): dedicated painted lanes are warranted when robot flow exceeds approximately 15/hr and pedestrian density reaches LOS B or higher. The space cost is manageable and the conflict reduction is substantial.
4. **In indeterminate zones**: the model cannot confidently distinguish between managed sharing and dedicated space. Planners should invest in local data collection—a short field study measuring actual robot–pedestrian interaction patterns on the corridor—before committing to infrastructure. A two-week observation campaign using existing traffic cameras may suffice to resolve the ambiguity.

A key finding for near-term policy is that behavioral regulation should precede infrastructure investment. Mandatory yielding rules, implementable through operating-permit conditions at negligible public cost, provide the largest single safety improvement. Volume caps are the second-most-effective tool. Physical lanes should be considered only when these management tools prove insufficient—a condition that our model places at relatively high robot penetration levels that few corridors have yet reached.

7.2. Limits of the bicycle lane analogy

The bicycle infrastructure evolution—shared road, painted lane, protected lane—has been invoked as a precedent for robot-lane planning. Our results suggest the analogy holds partially but breaks down in a policy-relevant way.

The similarity is structural: both cases involve allocating scarce right-of-way to a new mode that creates conflicts with existing users, and the preferred intervention depends on corridor conditions. The threshold-based logic (separation warranted above certain flow and density thresholds) transfers directly.

The difference is behavioral controllability. A cyclist's speed, path, and yielding behavior are autonomous; the only tools available to planners are infrastructure and signage. A delivery robot's behavior is programmable. The most

effective safety intervention we identify—mandatory yielding—has no bicycle equivalent, because a city cannot compel a cyclist to yield through firmware. This controllability means that the management-only solution space for robots is substantially larger than for bicycles, and the conditions under which physical separation is warranted are correspondingly narrower.

This does not mean dedicated robot lanes will never be needed. Our results show that at high robot penetration on wide sidewalks, dedicated space outperforms even well-managed sharing. As robot fleets scale, some corridors will enter this regime. But the transition will likely be slower and more geographically concentrated than the bicycle-lane transition, because software-based management provides a buffer that bicycles never had.

7.3. Transferability considerations

The model is calibrated to a single site (NAU campus) with a relatively homogeneous pedestrian population and controlled pathway geometry. Applying the screening framework to other contexts requires attention to three transferability issues.

First, pedestrian population heterogeneity. Campus pedestrians are predominantly young, able-bodied, and familiar with delivery robots. An urban commercial district would include older adults, children, wheelchair users, and pedestrians distracted by phones or shopping. These populations likely exhibit different PET distributions—potentially with a larger dangerous fraction—requiring recalibration of the mixture parameters. Until multi-site calibration data are available, planners should treat our PET parameters as optimistic and consider applying a safety factor (e.g., increasing the dangerous fraction by 20–30%) when assessing downtown corridors.

Second, sidewalk context. Campus pathways are relatively free of obstructions, grade changes, and intersections. Urban sidewalks feature curb cuts, street furniture, transit stops, and driveway crossings that interrupt robot flow and create additional conflict points not captured by our corridor-segment model. The pinch-point analysis (Section 6.3) provides partial guidance, but a full ur-

ban application would benefit from segment-level GIS analysis of obstruction frequency.

Third, robot behavior variation. Our model assumes Starship-type robots (six-wheeled, 0.55 m wide, 1.79 m/s operating speed). Larger robots (e.g., Nuro R3, which operates in roadways but has been proposed for wide pedestrian areas) or faster robots would produce different interaction patterns. The model accommodates speed variation through the v_r parameter, but width and behavioral differences would require respecification of the encounter model.

Despite these limitations, the screening framework is designed for transferability in its structure if not its specific parameter values. The inputs (width, pedestrian flow, robot flow) are universal corridor descriptors, and the decision logic (compare conflict exposure across interventions subject to accessibility constraints) applies regardless of the local calibration. We encourage cities considering robot-lane allocation to use the framework with locally recalibrated parameters, treating our NAU-based parameters as initial defaults for the first screening pass.

7.4. Limitations

Several limitations bound the interpretation of our results.

Reduced-form abstraction.. The model does not track individual agent trajectories and therefore cannot capture emergent phenomena such as platoon formation, pedestrian herding, or strategic robot routing. It assumes steady-state conditions and uniform corridor geometry. These abstractions are appropriate for screening-level analysis but insufficient for detailed design of lane markings or signal timing.

No network effects.. The model evaluates individual corridor segments independently. In practice, dedicating a lane on one segment affects robot routing on the network, potentially concentrating flows on dedicated corridors and dispersing them elsewhere. A network-level analysis would be needed to capture these redistribution effects.

No intersection modeling.. The 50 m segment model does not include intersections, crosswalks, or driveways, where a large fraction of real-world pedestrian–robot conflicts may occur. Intersection analysis would require a different model structure (e.g., gap-acceptance models) and additional calibration data.

Exogenous demand.. Pedestrian and robot flows are treated as fixed inputs. In reality, sidewalk conditions affect demand: unpleasant pedestrian environments may reduce walking, and congested sidewalks may lead operators to reroute robots. Endogenous demand effects could amplify or dampen the intervention differences we report.

Single-site calibration.. As discussed in Section 7.3, the empirical foundation is a single campus site. Multi-site calibration is the highest-priority extension of this work.

No cost analysis.. We compare interventions on safety and accessibility metrics without incorporating implementation costs (painting, bollards, enforcement, monitoring). A benefit–cost analysis would likely strengthen the case for managed sharing at low robot flows (zero infrastructure cost) and may alter the boundary conditions where dedication becomes cost-effective.

8. Conclusion

This paper addresses a planning question that will confront an increasing number of cities: under what conditions should municipalities dedicate sidewalk space to delivery robots rather than managing shared operation? We developed a reduced-form stochastic screening model, calibrated to 201 observed robot–pedestrian interactions [1], and tested over 5,250 core scenarios spanning five sidewalk widths, five pedestrian demand levels, and seven robot flow rates.

Three principal findings emerge.

Finding 1: Managed sharing is preferred more often than dedicated space, but neither dominates universally.. Across all tested conditions, managed sharing (speed cap + mandatory yielding + volume limit) wins in 46% of scenarios; the dedicated painted lane wins in 33%; and the shared baseline is acceptable in the remaining 21%. The preferred intervention depends on corridor conditions, reinforcing the need for a screening approach rather than a one-size-fits-all policy.

Finding 2: Sidewalk width is the primary determinant of the allocation decision.. The dedicated lane's win share increases from 22.9% at $W = 1.8\,\text{m}$ to 42.9% at $W = 4.2\,\text{m}$, with a threshold at approximately 3.6 m above which dedicated space becomes competitive for a plurality of conditions. On narrow sidewalks, the space cost of dedication is prohibitive; on wide sidewalks, the conflict-reduction benefit justifies the allocation.

Finding 3: Mandatory yielding is the single most effective management tool; speed caps alone can backfire.. Policy ablation reveals that mandatory yielding reduces conflict exposure by 28.9% and volume caps by 31.6%, while a speed cap alone *increases* conflict exposure by 21.1% due to prolonged robot dwell time. Cities currently relying only on speed limits in their PDD statutes should consider adding yielding mandates as a higher-priority intervention.

These findings support several policy recommendations. First, cities should prioritize behavioral regulation through operating-permit conditions before investing in dedicated robot infrastructure. Second, dedicated robot lanes should be considered primarily on sidewalks wider than 3.6 m where robot flows exceed approximately 15 units per hour. Third, planners should use the screening chart (Figure 10) as a first-pass assessment tool, collecting additional site-specific data for corridors that fall in the indeterminate zone (agreement below 50%).

The ranking uncertainty we document near the management–dedication boundary (40–46% agreement) is itself informative. It signals that for many corridors, the choice between managing shared space and dedicating robot space

is genuinely ambiguous given current evidence. Rather than treating this ambiguity as a limitation, we suggest that cities use it as a trigger for adaptive management: begin with managed sharing, monitor interaction patterns, and transition to dedicated space only when local data confirm that the corridor has entered the dedication-preferred regime.

Several extensions would strengthen the framework. The highest priority is multi-site calibration of the PET distribution, incorporating observations from urban commercial districts, residential neighborhoods, and transit corridors to test whether the NAU campus parameters generalize. A microscopic simulation benchmark—comparing screening-model recommendations against agent-based model predictions for a subset of scenarios—would validate the reduced-form approach against higher-fidelity models. Network-level analysis would capture redistribution effects when dedicated lanes alter robot routing. Finally, an equity assessment examining how intervention choices affect different pedestrian subpopulations (older adults, wheelchair users, visually impaired pedestrians) would ensure that space-allocation decisions do not disproportionately burden vulnerable users.

Sidewalk delivery robots are transitioning from novelty to infrastructure challenge. The planning frameworks that guided bicycle-lane allocation over four decades provide a structural precedent, but the software-controllable nature of robot behavior creates a wider and potentially more efficient intervention menu. Our screening model offers cities a first quantitative tool for navigating this new allocation problem—grounded in empirical interaction data, transparent in its assumptions, and calibrated to the observable corridor characteristics that planners work with daily.

CRediT authorship contribution statement

[Author 1]: Conceptualization, Methodology, Software, Formal analysis, Writing – Original Draft, Visualization. **[Author 2]**: Supervision, Writing – Review & Editing, Funding acquisition. *[To be completed upon acceptance.]*

Data availability

The simulation code and scenario configuration files used in this study will be made available on GitHub upon acceptance. The empirical calibration data are derived from **(author?)** [1], which is publicly available.

Acknowledgments

[Blinded for review.]

References

[1] S. R. Gehrke, A. Akhavan, T. F. Welch, Observed interactions between sidewalk autonomous delivery robots and other pathway users, Transportation Research Interdisciplinary Perspectives 18 (2023) 100789. doi:10.1016/j.trip.2023.100789.

[2] I. Bakach, A. M. Campbell, J. F. Ehmke, A two-tier urban delivery network with robot-based deliveries, Frontiers in Future Transportation 2 (2022) 773240. doi:10.3389/ffutr.2021.773240.

[3] D. Jennings, M. Figliozzi, Study of sidewalk autonomous delivery robots and their potential impacts on freight efficiency and travel, Transportation Research Record 2673 (6) (2019) 317–326. doi:10.1177/0361198119849398.

[4] N. Boysen, S. Schwerdfeger, F. Weidinger, Scheduling last-mile deliveries with truck-based autonomous robots, European Journal of Operational Research 271 (3) (2018) 1085–1099. doi:10.1016/j.ejor.2018.05.058.

[5] M. Kovacic, D. Milakis, Regulating sidewalk robots: insights from a comparative analysis of urban governance approaches, Urban Geography 45 (3) (2024). doi:10.1080/02723638.2023.2298925.

[6] Z. Weinberg, D. A. Tedjopurnomo, Z. Zheng, J. Gao, T. Nelson, K. Sila-Nowicka, Sharing the sidewalk: observing delivery robot and pedestrian interactions, Mineta Transportation Institute Publications 7 (5) (2023) 53.

[7] R. Buehler, J. Dill, Bikeway networks: a review of effects on cycling, Transport Reviews 36 (1) (2016) 9–27. doi:10.1080/01441647.2015.1069908.

[8] J. E. Schoner, D. M. Levinson, The missing link: bicycle infrastructure networks and ridership in 74 US cities, Transportation 41 (6) (2014) 1187–1204. doi:10.1007/s11116-014-9538-1.

[9] G. Franchi, N. Fazeli, C. Berger, Robotability score: an infrastructure-aware measure of sidewalk traversability for delivery robots, in: Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems, ACM, 2025.

[10] X. Tong, M. D. Simoni, Robust route planning for sidewalk delivery robots under pedestrian uncertainty, arXiv preprint arXiv:2507.12067 (2025).

[11] H. Yang, X. Chen, Z. Li, An analytical model for designing sidewalk autonomous delivery robot fleet systems, Transportation Research Part C: Emerging Technologies 160 (2025) 104978. doi:10.1016/j.trc.2024.104978.

[12] U. Weidmann, Transporttechnik der Fussgänger: Transporttechnische Eigenschaften des Fussgängerverkehrs (Literaturauswertung), Tech. Rep. 90, ETH Zürich, Institut für Verkehrsplanung, Transporttechnik, Strassen- und Eisenbahnbau (1993).

[13] R. Bennett, R. Vijaygopal, R. Kottasz, Autonomous delivery robots and vulnerable road users: understanding pedestrian safety concerns, Journal of Product & Brand Management 33 (2) (2024) 181–194. doi:10.1108/JPBM-01-2023-4309.

[14] United States Access Board, ADA standards for accessible design, Tech. rep., Department of Justice, Washington, DC (2010).

[15] M. M. Rahaman, J. Strauss, G. H. d. A. Correia, Sidewalk accessibility for people with disabilities: a review of barriers and solutions, Journal of Urban Planning and Development 148 (3) (2022) 03122002. doi:10.1061/(ASCE)UP.1943-5444.0000854.

[16] S. Zangenehpour, J. Strauss, L. F. Miranda-Moreno, N. Saunier, Are intersections with bike lanes safer? A study based on automated surrogate safety analysis using video data, Accident Analysis & Prevention 86 (2016) 69–78. doi:10.1016/j.aap.2015.10.015.

[17] S. Agrawal, Z. Zheng, S. Peeta, J. He, Evaluating robot influence on pedestrian behavior models using immersive virtual reality, arXiv preprint arXiv:2409.14844 (2024).

[18] S. Agrawal, Z. Zheng, S. Peeta, J. He, PeRoI: a pedestrian-robot interaction dataset for intention-aware and safety-aware autonomous driving, arXiv preprint arXiv:2503.16481 (2025).

[19] J. Song, S. Lee, J. Lee, Effects of delivery robot speed on pedestrian comfort and safety perception in shared sidewalk environments, Journal of Transport Geography 114 (2025) 104048.

[20] National Association of City Transportation Officials, Urban bikeway design guide, Tech. rep., NACTO, New York, NY (2014).

[21] R. L. Sanders, M. Branion-Calles, T. A. Nelson, M. Winters, L. D. Frank, K. Laberee, E. Setton, Competitive versus supportive: understanding cyclist needs using the Bicycle Level of Traffic Stress, Journal of Transport Geography 53 (2016) 1–11. doi:10.1016/j.jtrangeo.2016.04.003.

[22] J. Alonso-Mora, S. Samaranayake, A. Wallar, E. Frazzoli, D. Rus, On-demand high-capacity ride-sharing via dynamic trip-vehicle assignment, Proceedings of the National Academy of Sciences 114 (3) (2017) 462–467. doi:10.1073/pnas.1611675114.

[23] A. Karndacharuk, D. J. Wilson, R. Dunn, A review of the evolution of shared (street) space concepts in urban environments, Transport Reviews 34 (2) (2014) 190–220. doi:10.1080/01441647.2014.893038.

[24] Transportation Research Board, Highway Capacity Manual, 5th Edition, National Academies of Sciences, Engineering, and Medicine, Washington, DC, 2010.

[25] D. Helbing, P. Molnár, Social force model for pedestrian dynamics, Physical Review E 51 (5) (1995) 4282–4286. doi:10.1103/PhysRevE.51.4282.

[26] M. Jafari, X. Liu, A heterogeneous social force model for personal mobility vehicles sharing pedestrian spaces, Simulation Modelling Practice and Theory 131 (2024). doi:10.1016/j.simpat.2023.102871.

[27] Y. Li, J. Chen, Z. Wang, C. Xu, Social force model parameter calibration for pedestrian crosswalk simulation via YOLOv5 detection, Sensors 24 (15) (2024) 5011. doi:10.3390/s24155011.

[28] C. F. Daganzo, Fundamentals of transportation and traffic operations, Emerald Group Publishing (1997).

[29] M. Batty, Big data, smart cities and city planning, Dialogues in Human Geography 3 (3) (2013) 274–279. doi:10.1177/2043820613513390.

[30] A. P. Vánky, S. Verma, T. Courtney, P. Santi, C. Ratti, Effect of weather on pedestrian trip count and duration: city-scale evaluations using mobile phone application data, Environment and Planning B: Urban Analytics and City Science 47 (7) (2020) 1309–1324. doi:10.1177/2399808319830870.