Nolan Sherwood
Pete McNiff

Safe Haven or Hazardous Miscalculation: The Mid-Valencia Bike Lane Pilot

I. *Introduction*

The obesity epidemic has been plaguing the U.S. since the 1980s (Temple, N. J., 2022). It shows no sign of subsiding, with the prevalence of obesity in our adult population hovering around 40% over the past decade (Emmerich et al., 2024). According to the 2024 Heart Disease and Stroke Statistics: A Report of U.S. and Global Data From the American Heart Association, on average, someone in the U.S. dies from cardiovascular disease every 34 seconds (Martin et al., 2024). Several major cultural shifts are required if America is going to progress toward a healthier lifestyle, one of the most promising being active commuting. Active commuting, such as walking, running, and biking, has been proven to decrease obesity, improve blood pressure, and increase life expectancy (Pucher et al., 2010). The likelihood that individuals engage in active commuting primarily depends on the convenience and safety of traffic infrastructure, the latter of which is an area of concern in the United States. (Dill, 2009). For example, in 2009, on average, for every 10 million kilometers an individual cycles, they are roughly 20 times more likely to be injured in the U.S. compared to the Netherlands (Pucher & Buehler, 2010).

Our paper focuses on an innovative transportation infrastructure pilot program on San Francisco's Valencia Street intended to increase cyclists' safety while maintaining the corridor's economic vitality. In January of 2022, public outrage over the safety conditions on Valencia Street reached a fever pitch as protesters from the SafeStreetRebel group staged a Just A Minute (JAM) protest, effectively blocking traffic whenever a car was double-parked in the bike lane (Thorpe, 2023). The San Francisco Municipal Transportation Agency (SFMTA) responded by announcing the construction of a new center median bike lane, as shown in Figure 1. The 18-month pilot program has since been terminated, due to local businesses campaigning against the program for its supposed negative economic impact; however, we want to assess the impact this program had on public safety, with the hope that center median bike lanes can pave the way toward a healthier future in the United States.

II. Data

The data used for this project comes from the San Francisco Open Data database, which is a public database that makes administrative data available from many branches of local government. The specific dataset we use is "Traffic Crashes Resulting in Injury". This dataset has 61,900 observations from 2005 to 2025, showcasing in-depth records of the date, time, and the parties involved (motor vehicle, bicycle, pedestrian, etc.) in all traffic accidents resulting in injury in San Francisco. We then created Figure 2, a heat map of the dataset, to better understand the distribution of accidents throughout the city. The dataset is compiled from two sources for the particular years pertaining to our research, 2023 and 2024. The fatal crash data is sourced from the Office of the Chief Medical Examiner, but only includes cases that meet the Vision Zero Fatality Protocol. Then, the injury crash data is drawn from the SFPD's Interim Collision System. A key limitation of the dataset is that only crashes with valid geographic information are included, potentially excluding incidents without precise location data.

For our research, we filtered the data to include only observations from 2024 and only incidents on and near the Valencia Street corridor. Filtering our data significantly narrowed our observations to 83, which we then categorized into blocks per quarter. Our categorization method increased our total observations to 124, but poses a significant limitation by increasing our volume of zeros for blocks without any accidents. With the type of model we utilized, which we will describe more in-depth in the next section, it is vital to depict the exact location of each accident. Fortunately, the dataset not only shows which intersection a crash took place but also the direction and distance from that intersection. This level of precision allows us to accurately create distance variables for each accident, as shown in Figure 3.

III. Model

Our treatment of interest and precise location data sets the stage for an ideal opportunity to utilize a regression discontinuity model. Implementing a regression discontinuity model has many advantages, such as mitigating endogeneity concerns, eliminating arbitrary exclusion restrictions, reducing reliance on functional form assumptions, and avoiding distribution assumptions. The core method by which we're able to identify causality, and therefore mitigate endogeneity, is by exploiting the cutoffs where the bike lane begins and ends because locations on either side of the cutoff are likely to be similar in every way except for the presence of the bike lane, making them highly comparable. The quasi-random variation in treatment allows us to interpret any observed discontinuity in accident rates at the cutoff as a causal effect of the bike lane, rather than a reflection of confounding factors like weather patterns, road quality, or driver risk tolerance.

For our specific model, we rely on three assumptions to infer causality. First, in the absence of the bike lane, it is plausible to expect that the average number of accidents per block per quarter would vary smoothly, specifically just inside and outside the bike lane. Urban road conditions, traffic patterns, and built environment features such as intersections, stop signs, and lighting typically change gradually across adjacent blocks. Second, no other variables affecting accident rates, such as traffic volume, road width, or lighting, change discontinuously at the cutoff. As mentioned earlier, by limiting the distance between inside and outside the bike lane, the presence of confounding factors isn't likely. Third, cyclists cannot manipulate the location of their accident in proximity to the bike lane, which is highly likely since everyone on the road is incentivized to avoid accidents.

Our model, which is depicted below, estimates the effect of the Mid-Valencia bike lane on traffic accidents using a regression discontinuity approach. The outcome variable, $Accidents_i$, measures the number of accidents per block per quarter at location *i*. $BikeLane_i$ is a binary indicator equal to 1 if the location falls within the treated section and 0 otherwise. $CutoffDistance_i$ captures the distance from the cutoff where the bike lane begins. The intercept, β_0, reflects the baseline accident rate at the cutoff on the untreated side. β_1 captures the immediate change in accidents at the cutoff, showcasing the treatment effect. β_2 reflects how accidents trend with distance before the bike lane, while the interaction term β_3 allows that trend to differ within the treated area. Then the sum $\beta_2+\beta_3$ gives the slope of accident rates inside the bike lane segment. Figure 4 graphically captures the inner workings of our model if the bike lane succeeds in decreasing accidents.

$$Accidents_i = \beta_0 + \beta_1 \cdot BikeLane_i + \beta_2 \cdot CutoffDistance_i + \beta_3 \cdot (BikeLane_i \cdot CutoffDistance_i) + u_i$$

IV. Results

Table 1 shows our estimation of the model for the Traffic Accidents Resulting in Injury dataset. We find that the impact of the bike lane is small and positive (an increase of 0.148 accidents per block per quarter within the bike lane), but it is statistically insignificant. The estimation of our model is also depicted graphically in Figure 5. Table 1 also shows that as the distance from the bicycle lane cutoff increases, the number of accidents per block per quarter decreases; however, this remains statistically insignificant. These findings suggest that the beginning and end of the bike lane are creating potentially hazardous traffic conditions.

To dispel concerns about preexisting differences in accident rates between blocks, we also ran a placebo test using data from 2023 (pre-bike lane). Table 2 displays the result of this model, which is also shown graphically in Figure 6. The results confirmed no significant difference in accidents at the future cutoff point before the bike lane was installed. The results from our placebo test are reassuring because they mean that the bike lane was not selectively installed on blocks with a higher accident rate to begin with.

As a robustness check, we applied a difference-in-difference analysis strategy using the neighboring Mission Street as a counterfactual. We assume the adjacent streets had parallel traffic accident trends before implementing the bike lane. The findings from this model are displayed in Table 3. While this model suggests that Valencia fared slightly better than Mission Street (accidents per quarter increased at a lower rate between years), the high standard error means we cannot interpret this as meaningful. Additionally, we ran the difference in difference model a second time, observing only bike-related incidents. The results of this model, shown in Table 4, suggest that while overall traffic incidents per quarter on Valencia may have increased less than incidents on Mission Street, the proportion of accidents on Valencia involving bikes may have increased. Given that alternative methodologies find consistently insignificant impacts of the bike lane, our results appear robust.

V. Conclusion

Utilizing publicly available San Francisco government data, we find the Mid-Valencia center running pilot program had a statistically insignificant impact on public traffic safety. This finding is in line with Goodno et al (2013) did not find a positive effect on safety from a similar bike lane in Washington D.C. A major limitation we experienced was that our dataset only includes accidents that resulted in an injury. More minor incidents that didn’t report injuries are not captured. Future investigation would require finding information on how the number of riders per quarter was impacted. The policy implication of this study is that San Francisco made the correct decision to remove the bike lane. An expensive center running bike lane project is not worth the cost because it has little effect on safety and is unpopular among city residents.

Nolan Sherwood
Pete McNiff

VI. References

Dill, J. (2009). Bicycling for Transportation and Health: The Role of Infrastructure. *Journal of Public Health Policy*, *30*, S95–S110. http://www.jstor.org/stable/4020725

Emmerich, S., Fryar, C., Stierman, B., & Ogden, C. (2024). Obesity and Severe Obesity Prevalence in Adults: United States, August 2021–August 2023. *Obesity and Severe Obesity Prevalence in Adults: United States, August 2021–August 2023*, *508*. https://doi.org/10.15620/cdc/159281

Martin, S. S., Aday, A. W., Almarzooq, Z. I., Cheryl A.M. Anderson, Arora, P., Avery, C. L., Baker-Smith, C. M., Bethany Barone Gibbs, Beaton, A. Z., Boehme, A. K., Commodore-Mensah, Y., Currie, M. E., Mitchell S.V. Elkind, Evenson, K. R., Generoso, G., Heard, D. G., Hiremath, S., Johansen, M. C., Kalani, R., & Kazi, D. S. (2024). 2024 Heart Disease and Stroke Statistics: A Report of US and Global Data From the American Heart Association. *Circulation*, *149*(8). https://doi.org/10.1161/cir.0000000000001209

Pucher, J., & Buehler, R. (2010). Walking and Cycling for Healthy Cities. *Built Environment*, *36*(4), 391–414. https://doi.org/10.2148/benv.36.4.391

Pucher, J., Buehler, R., Bassett, D. R., & Dannenberg, A. L. (2010). Walking and Cycling to Health: A Comparative Analysis of City, State, and International Data. American Journal of Public Health, 100(10), 1986–1992. https://doi.org/10.2105/ajph.2009.189324

Temple, N. J. (2022). The Origins of the Obesity Epidemic in the USA–Lessons for Today. *Nutrients*, *14*(20), 4253. National Library of Medicine. https://doi.org/10.3390/nu14204253

Thorpe, A. (2023). Disrupting Infrastructure: Space, Speed and Street Governance. Ssrn.com. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4661456

Nolan Sherwood
Pete McNiff

Figure 1: The Valencia Street Center Median Bike Lane

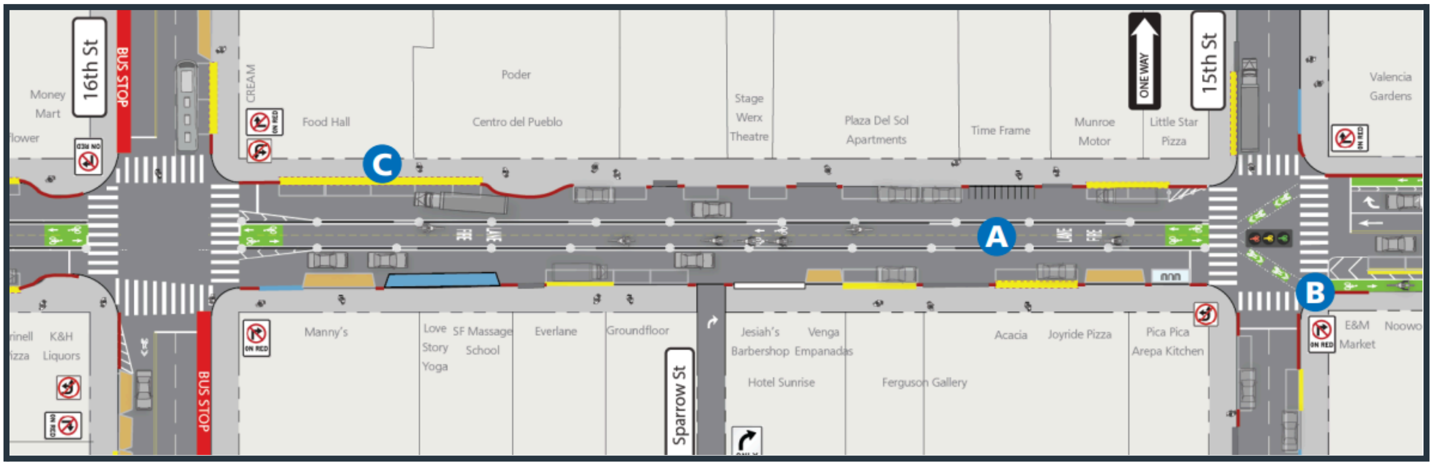


Figure 2: Heatmap of Traffic Crashes Resulting in Injury Dataset

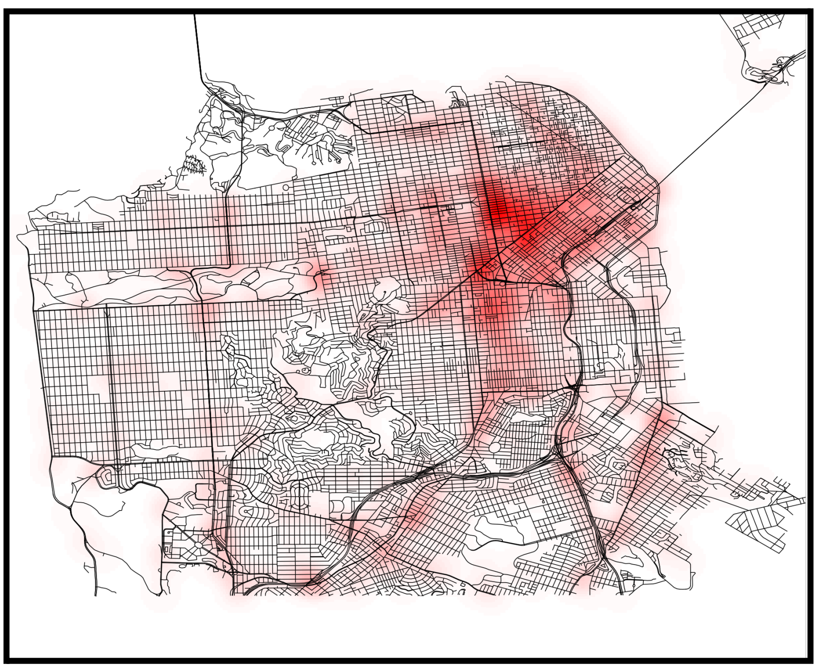


Figure 3: Primary Data Depiction

Primary Rd	Secondary Rd	Distance	Direction	Intersection	Cutoff Distance
VALENCIA ST	15TH ST	22	North	Midblock > 20ft	1.0
VALENCIA ST	16TH ST	162	South	Midblock > 20ft	2.0
VALENCIA ST	16TH ST	0	Not Stated	Intersection ≤ 20ft	1.0
VALENCIA ST	16TH ST	5	South	Intersection ≤ 20ft	1.0
VALENCIA ST	17TH ST	0	Not Stated	Intersection ≤ 20ft	2.0
VALENCIA ST	17TH ST	0	Not Stated	Intersection ≤ 20ft	2.0
VALENCIA ST	17TH ST	0	Not Stated	Intersection ≤ 20ft	2.0
VALENCIA ST	17TH ST	125	South	Midblock > 20ft	2.0
VALENCIA ST	17TH ST	24	North	Midblock > 20ft	2.0
VALENCIA ST	17TH ST	140	South	Midblock > 20ft	3.0

Figure 4: Graphical Representation of Our Model

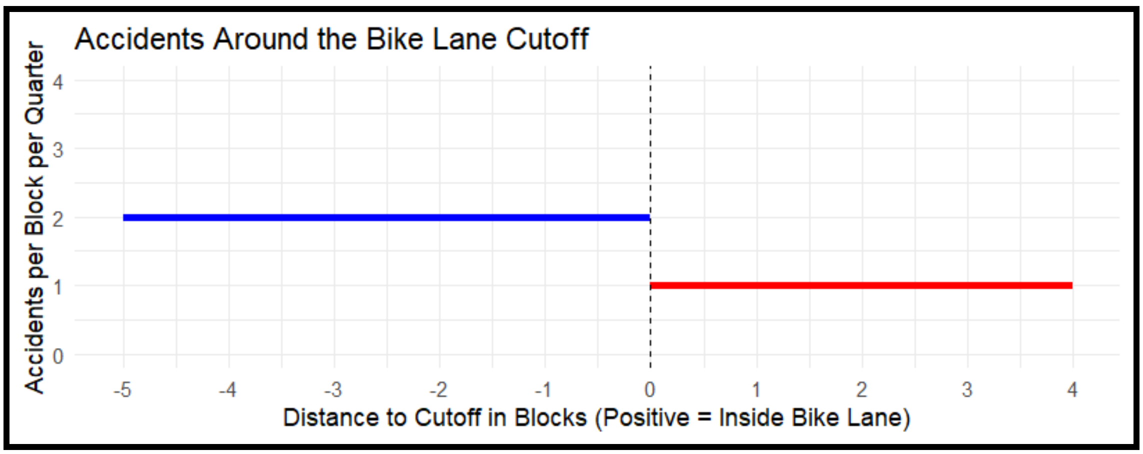


Figure 5: 2024 Regression Discontinuity Results

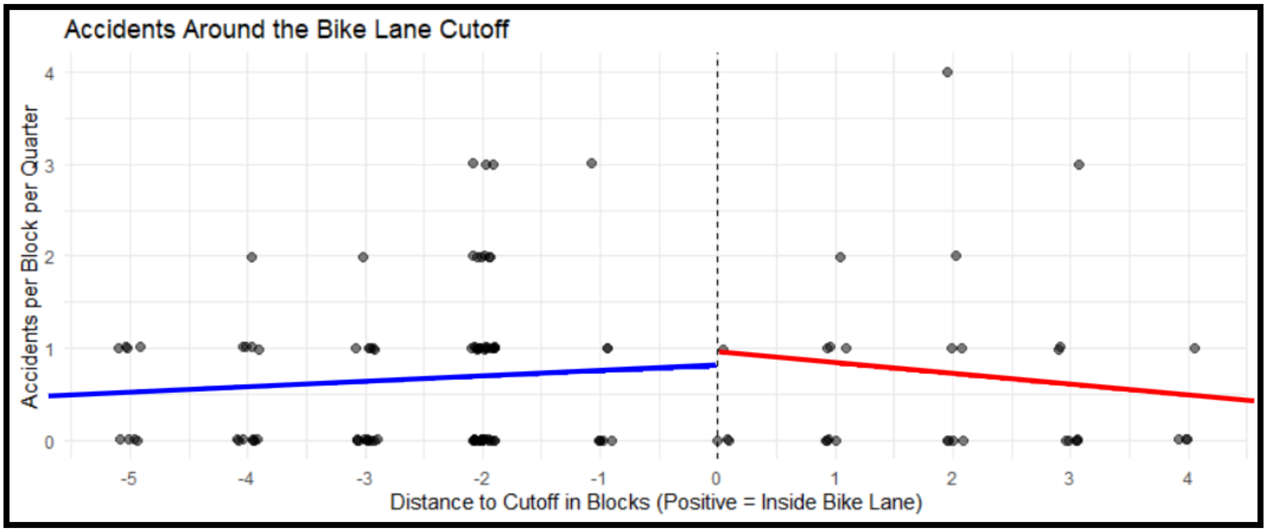
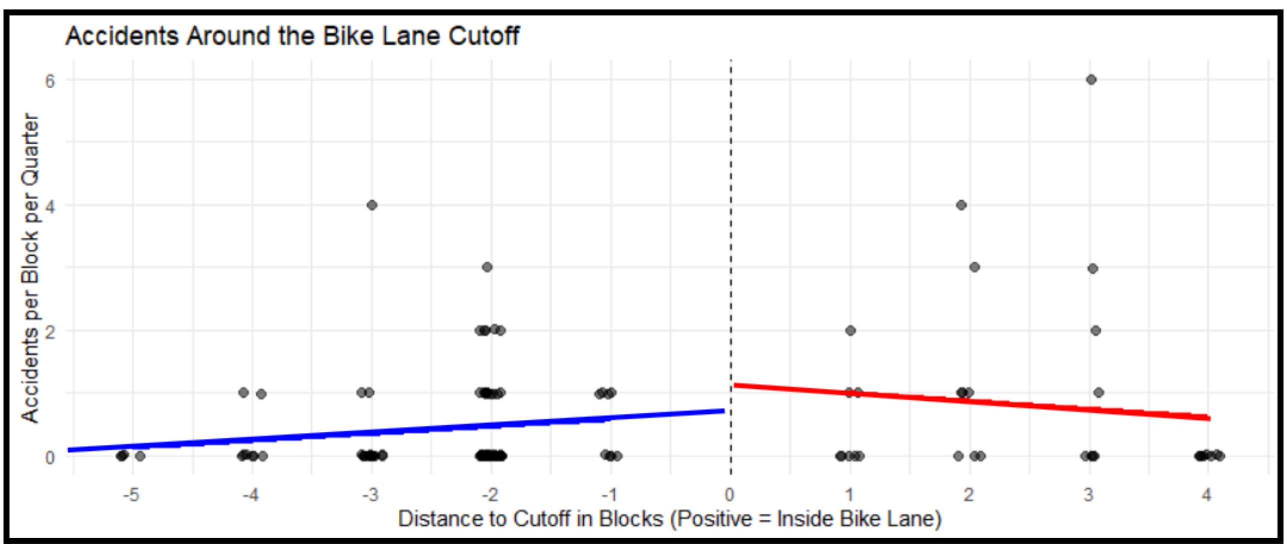


Figure 6: 2023 Regression Discontinuity Results



Nolan Sherwood
Pete McNiff

Table 1: Impact of Bike Lane on Accidents

	Accidents Per Block Per Quarter
Bike Lane	0.148
	(0.445)
Cutoff Distance	0.056
	(0.078)
Bike Lane× Cutoff Distance	−0.171
	(0.174)
Constant	0.794***
	(0.205)
Observations	124
R^2	0.009
Adjusted R^2	−0.016
Residual Std. Error	0.861 (df = 120)
F Statistic	0.368 (df = 3; 120), $p = 0.776$

Note: $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table 2: Impact of Bike Lane on Accidents

	Accidents Per Block Per Quarter
Bike Lane	0.455
	(0.511)
Cutoff Distance	0.113
	(0.111)
Bike Lane × Cutoff Distance	−0.238
	(0.190)
Constant	0.670*
	(0.288)
Observations	120
R^2	0.049
Adjusted R^2	0.024
Residual Std. Error	0.975 (df = 116)
F Statistic	1.977 (df = 3; 116), $p = 0.121$

Note: $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Nolan Sherwood
Pete McNiff

Table 3: Impact of Bike Lane on Accidents

	Accidents
Bike Lane	-24.500^{***}
	(4.108)
Post	3.000
	(4.108)
Bike Lane × Post	-1.000
	(5.809)
Constant	33.250^{***}
	(2.905)
Observations	16
R^2	0.862
Adjusted R^2	0.827
Residual Std. Error	5.809 (df = 12)
F Statistic	24.95 (df = 3; 12), $p = 1.92e{-}5$

Note: $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$