Street Typology and Bicyclist Safety: A Systems Approach

By

Eric Vallabh Minikel

B.S. in Chinese Language and Literature
University of Wisconsin – Madison, 2004

Submitted to the Department of Urban Studies and Planning
and the Department of Civil and Environmental Engineering
in partial fulfillment of the requirements for the degrees of

Master in City Planning
and
Master of Science in Transportation

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2010

Author____
Department of Urban Studies and Planning
Department of Civil and Environmental Engineering
May 17, 2010

Certified by ____
Professor Joseph Ferreira
Department of Urban Studies and Planning
Thesis Supervisor

Accepted by____
Professor Joseph Ferreira
Chair, MCP Committee
Department of Urban Studies and Planning

Accepted by____
Professor Daniele Veneziano
Chair, Departmental Committee on Graduate Studies
Department of Civil and Environmental Engineering

Street Typology and Bicyclist Safety: A Systems Approach

By
Eric Vallabh Minikel

Submitted to the Department of Urban Studies and Planning and
the Department of Civil and Environmental Engineering on May 18, 2010
in partial fulfillment of the requirements for the degrees of
Master in City Planning and Master of Science in Transportation
at the Massachusetts Institute of Technology.

Abstract

Cycling is an attractive transportation mode but has not attained a large mode share in the United States, in part because it is correctly perceived as dangerous. Much literature on cyclist safety and the built environment has focused on bicycle facilities such as bike lanes, while fewer studies have addressed route choice, specifically side streets versus arterials. The cities of Berkeley, CA and Cambridge, MA have pursued opposite strategies, Berkeley creating "bicycle boulevards" out of residential side streets while Cambridge has added bike lanes to major arterial roads.

I hypothesize that side streets are safer than arterials regardless of which street has been treated for cyclists, both in terms of collision rate and fraction of collisions that are severe. For each city, I use cyclist count data with police-reported bicycle-motor vehicle collision data to compute relative collision rates for pairs of parallel streets that are alternate routes to reach the same destinations. In a detailed analysis of Berkeley I find that bicycle boulevards have categorically lower collision rates than arterials, with no difference in severity. I demonstrate, with very limited data, how to apply the same methodology to Cambridge and find results somewhat suggestive that side streets have lower collision rates than arterials. I highlight shortcomings in current data collection practices that make this research difficult.

Thesis Supervisor: Joseph Ferreira
Title: Professor of Urban Planning and Operations Research

Table of Contents

Acknowledgements

Dedicated to Mike Craig. Special thanks to Sonia Vallabh, Joseph Ferreira, Chris Zegras, Krista Nordback, Ezra Haber Glenn, Peter Furth, Anne Lusk, Jeff Parenti, Laura Kowalski, Kathleen Murphy, Cara Seiderman, Faye Nixon, Matt Beyers, Eric Anderson, Greg Raisman, Denver Igarta, Chris Monsere, Myeonwoo Lim, Oliver Smith, Norman Garrick, Chris McCahill, Wes Marshall, Meghan Winters, Daniel Daou and Aditi Mehta.

1 Introduction

1.1 Inspiration

Cycling holds great promise as a mode of transportation. It offers fresh air, physical exercise, and greater mobility than walking while using less street space than automobiles and producing no appreciable pollution. In cities, bicycles actually offer the freedom that automobile marketing has long taught us to associate with cars: you can go anywhere, anytime, skip through congestion, and find parking easily. Unlike public transit, your bicycle never comes 35 minutes late, you never just miss it as it pulls away, it never stops running at midnight, is never too crowded to get on, and you always get a seat.

But like a Shakespearean hero, cycling has a tragic flaw: it's just not very safe. In the United States, compared to motor vehicle occupants, cyclists are eleven times as likely to be killed per mile traveled, or three times as likely per trip made (Pucher and Dijkstra 2000). Surveys indicate that the public perception of this danger is one of the main things that keeps people from cycling more (Reynolds et al 2009). No wonder, then, that fewer than 1% of American commuters ride bikes to work (United States Census 2000). But cycling can be made safer—some of the world's most bicycle friendly countries, such as the Netherlands and Germany, have just a fourth the rate of cyclist fatalities per mile traveled as the United States (Pucher and Dijkstra 2000).

The question, then, is how. A number of cities in the United States have begun to install transportation facilities specifically for cyclists—bike lanes, bike paths, and more. These sorts of facilities are common in the countries where cycling is safest, so they seem a fair bet for improving safety. Oddly, though, the academic literature on the safety effects of these facilities is quite limited and delivers no clear verdict.

The debate over the safety of bicycle facilities often skips over a crucial point: on which streets should cyclists be encouraged to ride? Are they safer on major automotive thoroughfares, or on quiet residential streets? Of course, this question has not gone unnoticed by cyclists, and strong opinions fly in both directions. But academic literature seeking to answer this question is nearly nonexistent.

In a cyclist's dream city, every street is a bikeable street, but the truth is that cities start from nothing and make changes incrementally. The last two cities in which I have lived, Berkeley, California and Cambridge, Massachusetts, have spent the last fifteen years pursuing two very different approaches to beginning to accommodate cyclists. In Berkeley, a network of "bicycle boulevards"—traffic-calmed side streets improved and designated for bicyclist priority—now stretches over the whole city, but cyclists receive no special treatment on the busy arterial streets that motorists use. By contrast, Cambridge has already lined many of its major roads with bike lanes, and intends to reach most of the rest, while making few accommodations for cyclists on side streets.

Berkeley's bike plan also touches on convenience, attractiveness, minimizing delay, and even "ambiance" (Wilbur Smith Associates 1998). Cambridge's web page about bike lanes mentions cyclist comfort and stress level and, above all, direct routes to destinations (City of Cambridge. Bicycle Lanes, 2009). Both documents express a goal of getting more people cycling, and both emphasize cyclist safety.

One city is encouraging cyclists to use side streets, and the other is encouraging cyclists to use arterials. Which is safer? As a cyclist, should I choose to take side streets, or arterials, or simply whatever route the city has designated? As a city planner, should I promote the use of side streets or of arterials?

1.2 What, why and who cares

In this thesis, I seek to partially answer these questions through comparisons of parallel streets. In both Berkeley and Cambridge I compute relative collision rates for different streets by dividing the number of police-reported collisions by a measure of exposure gleaned from manual cyclist counts. I use these collision rates to compare Berkeley's bicycle boulevards to the untreated arterials to which they are alternatives. In Cambridge I compare treated or partially treated arterials with untreated side streets.

Such a comparison of parallel streets is valuable because it represents the real choice that cyclists face—on which street to ride—and the real choice that some cities face—where to encourage and enable cycling. My findings may not be easily generalizable to other cities, since I examine the *whole* street without isolating what factors (intersection and driveway frequency, traffic volume, roadway width, and so on) bring about the safety difference. On the other hand, the general idea of a residential, low-traffic street running parallel to an arterial is common to many cities, and if a significant safety difference exists for cyclists between side streets and arterials, officials in other cities should find it worth knowing about.

1.3 Road map

In the Literature Review I highlight the challenges facing cyclist safety research, characterize the current state of the field and point out that there is little research on street typology, particularly comparing parallel routes. Research does suggest, though, that motor vehicle volume and speed and the presence of heavy vehicles are all detriments to cyclist safety, so I formulate the hypothesis that side streets should be safer for cyclists than the arterials to which they run parallel.

In Methodology, I pledge to test this hypothesis in Berkeley, comparing bicycle boulevards to the arterials for which they are alternatives, and in Cambridge, comparing untreated side streets to arterials. In comparing safety, I compare both the rate of collisions per cyclist and the proportion of those collisions that result in severe injury. Key to my methodology is selecting identical segments of parallel streets so that the comparison represents a real choice—both for cyclists and for cities. I divide police-reported collisions by counts of cyclists to obtain collision rates for each street. I use Poisson regression to determine whether collision rates are significantly different, and examine collision and count data carefully to determine how well they represent full risk.

In Data, I provide extensive background on both Berkeley and Cambridge. In particular, I characterize their very different approaches to accommodating cyclists on city streets, where Berkeley has turned side streets into bicycle boulevards and Cambridge has applied bike lanes on arterials. For each city, I use police-reported collision data and manual cyclist counts. I describe the affordances and limitations of these datasets.

Analysis and Results are quite robust for Berkeley, demonstrating convincingly that bicycle boulevards are safer for cyclists than arterials, and that the difference is due to street typology and not safety in numbers or self-selection. Poisson regression indicates that the differences I observe are quite unlikely to have come about by chance without a true underlying safety difference, and my analysis of collision and count data indicate that those are reliable enough as not to call my findings into serious doubt. For Cambridge, data is limited and though I do find a lower collision rate for side streets than arterials, the difference is not statistically significant.

An obvious application of these findings is for city policy towards cycling but, as I explain in Conclusions, my results do not add up to a clear policy recommendation for cities, largely because choosing the safest route is not the only consideration for cities. I also suggest changes in data collection practices that could enable further research.

2 Literature Review

2.1 Definition of terms

Facilities built expressly for bicycle use generally take three forms: bike lanes, cycle tracks and bike paths. Bike lanes (Figure 1) are lanes designated for bike use only, separated from motor vehicle traffic by a stripe of paint. Cycle tracks (Figure 2) are lanes designated for bike use only and separated from motor vehicle traffic by a curb or other barrier. Bike paths (Figure 3) are linear facilities that carry only bikes; there are is no motor vehicle traffic to begin with.

Other types of infrastructure mix bicyclists with other road users. Multi-use paths (Figure 4) are like bike paths but they also allow joggers, rollerbladers, and other non-motorized users. Sharrows (Figure 5) exhort motorists to politely share the lane with cyclists, though in most cases cyclists are allowed by law to ride in the roadway regardless (Mionske 2007, 8-13).

All of these treatments are straightaway treatments which do not address how cyclists and motor vehicles should behave at intersections. These interventions may or may not be combined with intersection treatments. Examples of intersection treatments include colored bike lanes through the intersection, special traffic signals that allow cyclists to cross while other traffic is stopped, or “bike boxes” and advanced stop lines—pavement markings that allow cyclists to jump ahead of cars in the queue to make sure they are seen by motorists before the light turns green.

Bicycle boulevards (Figure 6) represent a more comprehensive attempt to give cyclists priority over motorists on certain side streets, including straightaway treatments and intersection treatments. These merit a more detailed discussion as they are a focus of this thesis.

Figure 1. Bike lane

Figure 2. Cycle track

Figure 3. Bike path

Figure 4. Multi-use path

Figure 5. Sharrow

Figure 6. Bicycle boulevard

Walker, Tressider and Birk define bicycle boulevards as “low-volume and low-speed streets that have been optimized for bicycle travel through treatments such as traffic calming and traffic reduction, signage and pavement markings, and intersection crossing treatments” (2009, 2). Motor vehicle access to each property is maintained, as is emergency vehicle access, but non-motorized modes receive priority. They assert that such streets may already be attractive to cyclists even before bicycle boulevard treatments are initiated, and that once a bicycle boulevard is complete, it is inviting to less experienced cyclists who are afraid of traffic. Bicycle boulevards are generally presented as alternatives to major automotive routes: “bicycle boulevards frequently parallel nearby arterial roadways on which many destinations are frequently located. The availability of a parallel arterial roadway also encourages motorists to use arterials rather than cutting through local streets" (2009, 8).

Cyclists can get hurt through a variety of types of unfortunate incidents. A fall is when a cyclist is involved with no other road user. The cyclist might hit a tree or a road sign, fly over the handlebars after getting a wheel stuck in a pothole, and so on. A collision is when a cyclist is involved with some other road user—a motor vehicle, another cyclist, a pedestrian. The term "crash" refers to either such event. I will not use "accident" because the term has fallen into disrepute as some road safety researchers feel it implies that such incidents are random, unpreventable, or inevitable.

2.2 Data challenges in cycling research

Researchers have studied the effects of bicycle facilities on either injury rates, crash rates or collision rates. These researchers always have to face a troubling shortage of data. Consider for a moment researching collision rate and ignoring severity. One way to express collision rate is shown in Equation 1.

Equation 1. An expression for collision rate

$$collision\ rate = \frac{collisions}{bicycle\ miles\ traveled}$$

To compute a collision rate, even a relative one, requires knowledge of both the numerator and denominator. The numerator of this fraction is available through either police reports, hospital records or surveys. Each of these sources brings its own limitations. Bicycle-motor vehicle collisions are tremendously underreported. One survey of cyclists found that of the crashes cyclists had been involved in, only about 40% were ever reported to police, and only about 5% involved hospital admission (Moritz 1997). This, combined with the fact that few people cycle to begin with, means it is difficult to amass enough crash data to establish statistical significance for any conclusions reached.

If the numerator is challenging, the denominator is far worse. Bicycle miles traveled is what is known as an "exposure" measure—an indicator of how much exposure to risk was required in order to generate the observed collisions. For motor vehicles, vehicle kilometers traveled can be gleaned from administrative data such as required safety checks or smog checks where the odometer mileage is recorded. Bicycles do not need smog checks, vehicle registration or fuel, so no such administrative data is available. If bicycle kilometers traveled cannot be calculated, a count of cyclists is the next best option, but this too is difficult. Census journey to work data offers an idea of how many people are cycling to work, but not of how many people cycle to other destinations, nor what routes are traveled, and anyway Census data only comes once every ten years, a long time in the cycling world—as I show later, cycling in Cambridge nearly doubled between 2003 and 2008 alone, so the 2000 Census data would be a poor representation of current levels of cycling. A number of means for electronically counting cyclists have been developed, using lasers, induction loops, and video cameras (Alta Planning and Design 2009). However, it appears that none of these technologies has yet been widely used in the

United States: the three cities for which I obtained count data—Berkeley, Cambridge, and Portland—all still conduct bicyclist counts by hand. Staff, contractors or volunteers stand at street corners for one or two hours in the AM or PM peak and record tally marks for each cyclist that passes, usually with direction of travel or turning indicated. The totals are then combined into a Microsoft Excel spreadsheet, with the direction and turning behavior of the cyclist lost at this stage in the case of Portland and Berkeley. If recorded over multiple years, the intersection counts allow for time series analysis, but are difficult to use for any evaluation of particular streets. Moreover, such counts are too costly to conduct with any frequency, and so are usually done just once or twice a year. If done at the same time of year, seasonal variations can be controlled for but not known.

Surveys offer hope of capturing unreported incidents, and of getting good exposure data by asking cyclists how much they travel on various facility types. On the surface, this seems perfect: lots of collisions and a direct measure of bicycle miles traveled. However, consider the many problems with this experimental design. Studies by Kaplan (1975) and by Moritz (1997, 1998) all obtain collision rates for different facility types by dividing the aggregate number of respondent-reported collisions on each street type by the aggregate number of respondent-reported miles on each street type. Rodgers (1997) similarly uses respondents' collisions and mileage along with dummy variables for "primary riding surface" among others in a logistic regression. Tinsworth, Cassidy and Polen (1994) use survey data as an exposure measure but hospital emergency room data as a crash measure. All five of these were national surveys, so consider one example of how unobserved variables can influence findings: bike lanes are found mostly in dense urban environments, which Marshall (2009) finds are safer for cyclists due to their street network, so the safety rating of bike lanes could reflect mostly these urban streets, while the safety rating of untreated major roads could reflect high-speed, high-volume arterials in suburbs. Another problem is that surveyed populations may not be representative of all cyclists. For instance, Kaplan's survey was of the League of American Wheelmen, who are presumably cycling enthusiasts, and Kaplan states that "League members were not chosen to represent the typical American bicyclist of today. This would be a gross misrepresentation of the facts" (1975, 12). Finally, Reynolds et al point out that, depending on the survey method, surveys may not capture serious collisions, including fatalities, that keep the cyclist from ever riding again (2009, 22).

Surveys, then, do not offer an easy way out of the data problem in studying cycling safety. The limitations of crash and exposure data are present in all of the studies of bicyclist safety that I have reviewed, and are certainly present in my own study as well. It is worth taking a moment to note how sparse the literature is, and how sparse the data within that literature. In their review, encompassing English language studies of transportation infrastructure and cyclist safety, Reynolds et al find just 23 studies, of which "Most... based their analyses and conclusions on at least 150 observations of injury or crash events, and seven studies based their analyses on more than one thousand observations." (2009, 14) Exposure data was even rarer. Take, for instance, the six studies of bike lanes that Reynolds et al reviewed. One adjusted for "city-wide" bicycle traffic volume based on counts at three intersections, just one of which was along one of the two streets studied (Smith and Walsh 1998). Another study controlled for exposure by assuming that certain types of collisions occurred independent of the presence of bike lanes and must therefore reflect cyclist volume. The remaining four were survey studies, subject to the unobserved variables and biases which I discuss above.

A recent development is the creation of interactive online mapping tools where users can report collisions or close calls. Joseph Broach's B-SMART tool for BikePortland.org has received more than 700 reports in its first year and a half (Broach 2008). This is promising in its ability to allow cyclists to share knowledge with each other and to allow researchers to see a large number of incidents or near-incidents that were never reported to police. It is still, in a way, a survey, subject to response bias—people who bother to report the incident to B-SMART may be more likely to be cycling enthusiasts than amateurs. However, it has an important advantage over the surveys used by Kaplan and by Moritz. It allows users to pinpoint the location of the incident, which means that researchers could later combine the user-submitted collision data with exposure data from other sources, if available, for each street. To my knowledge, no cycling safety studies have yet been done using such web user-submitted data.

In sum, the cycling safety literature to date has been limited in its strength by the low quality of data available. This is not to demean the researchers' efforts. To the contrary, they are making the best effort they can to study an important topic about which there is very little data.

2.3 Cyclist safety and the built environment

In a thorough literature review, Reynolds et al find that bicycling has significant benefits in terms of "physical and mental health, decreased obesity, and reduced risk of cardiovascular and other diseases" (2009, 4) and, compared to auto use, decreases the externalities of noise and pollution. They also note that only about 1% of trips in North America are made by bicycle, compared to 10 to 20% in some European countries (2009, 7).

The perceived danger of cycling may be one reason. Noland (1995) examines mail survey data from metro Philadelphia about perceived probability of injury, perceived expected severity of injury, and choice of auto, bicycle, walking or transit, and concludes that perceived risk is a significant determinant of mode choice. Moreover, he finds an elastic relationship, meaning that a 1% reduction in perceived risk would lead to more than a 1% increase in cycling. Decima Research (2000), in a Toronto telephone survey, finds that a plurality (35%) of utilitarian cyclists rank "careless drivers" as their number one concern about cycling, while among recreational cyclists, "unsafe traffic conditions" are the second most common reason for not cycling to work. Winters et al (submitted), surveying Vancouver, BC cyclists and would-be cyclists, list the top ten deterrents of cycling, nine of which relate to auto traffic, riding surfaces or lighting—all of which might indicate an underlying concern about safety.

Clearly, other factors matter as well—for instance, Winters et al (2007) find that weather is a significant determinant of cycling among Canadians, and Decima Research (2000) finds strong seasonality in Toronto mode choice, with fewer people cycling in the winter.

Nevertheless, safety appears to be a significant concern both for cyclists and for would-be cyclists. Reynolds et al (2009) categorize cyclists as "vulnerable road users" because they have little or no physical protection in the event of a crash, and have far less mass than automobiles. This obvious point suggests that the perceived danger of cycling may be quite real. Pucher and Dijkstra (2000) use fatality statistics from the U.S. and Europe to show that this is in fact the case. In the U.S., compared to motor vehicle occupants, cyclists are 11 times as likely to be killed per mile traveled, or 3 times as likely per trip. Compared to cyclists in The Netherlands and Germany, American cyclists are 4 times as likely to be killed per mile traveled.

The comparison to Europe indicates that cycling can be made safer than it is in the United States today. Making cycling safer could be a public health boon: the direct toll of collisions would be reduced, and if the surveys reviewed here are correct that perceived danger deters cycling, then improved safety would also induce more people to cycle and capture the health benefits thereof.

Getting more people onto bicycles might in and of itself have positive effects on cyclist safety. One body of vulnerable road user safety addresses a phenomenon known as "safety in numbers" (for instance, Jacobsen 2003, Robinson 2005; additional studies reviewed in Reynolds et al 2009) Again and again, whether in time series or comparing across places, studies have found that when and where there are more cyclists (or pedestrians), there are also fewer collisions per cyclist (or pedestrians). Framed differently, cyclist collisions increase less-than-linearly with cyclist volume (the same for pedestrians). Though much has been written on the subject, the direction of causation is not well established. Jacobsen (2003) cites evidence that the presence of pedestrians changes motorist behavior, causing them to drive more slowly and carefully. The same can be imagined for cyclists, and while this is almost certainly part of the explanation, it also seems obvious that people are more likely to walk or bike where it appears safe to do so. Other factors—street network density and urban character, for instance—probably also play a role in making cycling and walking both safer and more common at the same time. It seems unlikely that causation flows exclusively in one direction, from numbers to safety.

Reynolds et al indicate that most studies of bicyclist safety in North America have focused on bicycle helmets, even though these only address injury severity and not the occurrence of collisions in the first place, and even then only address head injuries (2009, 5). For improving safety, they argue, the improvement of cycling infrastructure is a strategy superior to the encouragement or legislation of helmet use, because it reaches a population without requiring individuals to take initiative, and because it requires no enforcement (2009, 6). Meanwhile, some studies have suggested a positive correlation between bicycle commuting and the presence of bicycle facilities, though it is hard to establish cause and effect (Dill and Carr 2003, Nelson and Allen 1997).

All of this points to a need to understand the safety effects of various types of bicycle facilities. First, though, it is important to note that some cyclists are stridently opposed to any bicycle facilities. The most well-known of these advocates is John Forester. In his classic book *Effective Cycling* (1993), Forester offers instructions on how to safely cycle on the road alongside motor vehicles by behaving, and demanding to be treated, as a vehicle operator. He goes on to argue that such "vehicular cycling" is in fact far safer than cycling in bike lanes, on bike paths, or in other bicycle facilities.

Forester has made an important contribution to cycling by teaching cyclists how to claim their rightful space on the street and not be intimidated off the road by motorists. However, when he argues that special accommodations for cyclists are unsafe, he does not reconcile this view with certain observable facts: in a response to Forester, Pucher points out that cycling is both safer and more common in countries, and cities, with more bicycle facilities, which is hard to reconcile with the notion that bicycle facilities are so unsafe for cyclists (2001).

In a review of six North American studies of bike lanes, Reynolds et al (2009) find an average reduction in risk, variously measured in injuries, crashes or collisions per cyclist, of 50%, with just one of the six studies finding no effect. Only two of the studies compared the same streets before and after bike lanes were added; the other four compared streets with bike lanes and streets without bike lanes nationwide. This is problematic since the goal of such research is presumably to help policymakers decide whether to install bike lanes on a given street. Elvik and Vaa (2004) review 13 mostly European studies of bike lanes and find an average reduction in total injuries of just 4%.

Vehicular cyclists often criticize bike lanes for placing cyclists in the "door zone" of parked vehicles (John S. Allen quoted in Lombardi 2002) and for requiring that cyclists and motorists make turns across each others' paths (Allen, Cambridge Bike Lanes: Political Statement or Road Improvement?). While these are clearly valid concerns, it is a leap to argue, as some have (several individuals quoted in Lombardi 2002), that the bike lanes therefore make cyclists *less safe*. Take, for instance, the "door zone" issue: if one assumes that, in the absence of bike lanes, every cyclist will behave like a well-trained vehicular cyclist, then it would appear that the bike lanes make cyclists less safe. But this assumption is worth questioning: maybe most cyclists, in the absence of bike lanes, ride *even closer* to parked cars than cyclists do where there are bike lanes. In fact, there is evidence for this idea: the City of Cambridge commissioned a study on its own Hampshire St. which found that bike lanes not only caused fewer cyclists to travel in the door zone, but also caused motorists to park, on average, closer to the curb (Van Houten and Seiderman 2005).

Reynolds et al do not address cycle tracks specifically, but Elvik and Vaa (2004), reviewing 15 studies of cycle tracks, find a statistically insignificant 1% increase in bicyclist injuries. As for paths, Reynolds et al (2009) find evidence suggestive that, compared to on-road cycling, bike-only paths are safer but multi-use paths more dangerous.

Neither Reynolds et al nor Elvik and Vaa include in their meta-study Jensen's 2008 detailed before-after study of several hundred crashes in Copenhagen on streets where cycle tracks and bike lanes were installed. Per cyclist, Jensen finds that cycle tracks led to a 10% increase in crashes and injuries, and bike lanes led to a 5% increase in crashes and a 15% increase in injuries. The City of Copenhagen uses automatic counters to conduct frequent 12-hour counts of cyclists (Jensen, personal email, 2010), so Jensen had access to better exposure data than any of the U.S. studies discussed here.

He proposes that one reason why safety worsened after bicycle facility installation may have been the concurrent elimination of on-street parking, which may have induced more motorists to turn onto side streets in search of parking, thus increasing intersection conflicts. He also points out that the new bicycle facilities greatly influenced mode choice, increasing cyclist counts by 20% and decreasing motor vehicle counts by 10%, and speculates that the lower motor vehicle volumes may have made for higher motor vehicle speeds or more jaywalking by pedestrians.

It is interesting to pause and consider the source of such wide variation in conclusions about bike lanes. Bike lanes may have some consistent effect on safety which some of these studies were unable to measure correctly due to data limitations. On the other hand, the safety effects of bike lanes may be largely dependent on context—frequency of intersections and of driveways, speed of motorized traffic, and so on. The safety effects may also depend greatly on implementation. For instance, a narrow bike lane immediately next to parked cars might be less effective than a wider bike lane next to a curb. Intersections could be particularly important—bike lanes and cycle tracks are facilities for straightaway sections of road, where anything or nothing might be done to accommodate cyclists at the intersection.

Reynolds et al also address intersection treatments. At roundabouts, they find, cycle tracks perform better than bike lanes. For ordinary 4-way intersections, only two studies were available, with mixed or inconclusive results about raised crossings and colored crossings. Data on intersection treatments appear to be even sparser than data on straightaway treatments. Weigand reviews literature on several intersection treatments and found mostly studies that examined behavior modification (for instance, who yields to whom) and conflicts (generally, situations which the researchers believed *could* have led to a collision, though Weigand never provides a definition). Almost no studies were identified which addressed actual crashes, collisions or injuries (Weigand 2008).

Several studies have been conducted which address street typology rather than bicycle facilities in particular. Kaplan (1975) finds that minor roads saw slightly fewer collisions per bicycle mile traveled than major roads, with a relative danger index of .92 compared to 1.00 for major roads. Updating Kaplan's work decades later with a nearly identical survey, Moritz (1998) finds the opposite: a danger index of .66 for major roads and .94 for minor roads. However, in yet another survey Moritz (1997) finds a danger index of 1.26 for major roads and 1.04 for minor roads. Tinsworth, Cassidy and Polen (1994) find that major thoroughfares are 2.45 times as risky for cyclists as neighborhood streets. The evidence, then, is mixed, and in any case, these were all survey studies, subject to the problems I discuss earlier in this literature review.

Interestingly, both studies by Moritz find that sidewalks were by far the most dangerous places to ride. Suppose for a moment, though, that sidewalk riding is done mostly by novice or not particularly law-abiding cyclists, and then only on roads they (correctly) perceive as being particularly dangerous due to fast and voluminous traffic. If this were true, then Moritz's findings could not be taken as evidence that the sidewalk itself is dangerous.

With regards to the type of the street *network* rather than the street itself, Marshall (2009) finds that denser, more urban street grids are associated with more walking and cycling and fewer serious crashes. He selects 24 medium-sized California cities and characterizes their street networks by computing the number of intersections per square mile (an indicator of network scale) and the link-to-node ratio (a measure of the number of four-way intersections, thus an indicator of network connectivity). He finds that a high number of intersections per square mile—therefore a fine, dense street network—is strongly associated with lower fatality and severe injury rates, both per population and as a fraction of crashes. Both higher density and higher connectivity of street networks are associated with higher rates of walking and cycling. He speculates that lower motor vehicle speeds on dense street networks may be responsible for the difference.

Accident Prediction Models (APMs) are another area of research which can shed some light on the impact of street typology and design on bicyclist safety. Researchers collect collision data and count data for a large number of roads and then regress the collision rates with the road characteristics to determine the effect of those characteristics on collision rate. While many studies reviewed above deal with a bicycle facility, such as bike lanes, in isolation, APMs may address more different aspects of the street context, such as motor vehicle volumes, speed, and street width.

Turner, Roozenburg and Francis (2006) construct APMs for cyclists in New Zealand based on cyclist collisions, cyclist volume and motor vehicle volume. They also review a number of APMs for pedestrians addressing pedestrian volume and motor vehicle volume, all of which find safety in pedestrians' own numbers but danger from greater numbers of motor vehicles. The one cyclist APM that they review finds the same for cyclists in Sweden. Turner et al use collision and count data in three New Zealand cities to develop APMs for a variety of different intersection and link types and a variety of different collision types. They also develop "Total APMs" for signalized intersections, roundabouts and mid-block locations. The exponent for cyclist volume is always found to be less than one—total collisions increase less-than-linearly with cyclist volume, therefore there is correlative "safety in numbers" where there is high cyclist volume. In all cases the exponent found for motor vehicle volume is positive, meaning that more motor vehicles spell more cyclist collisions. So while there is safety in cyclist numbers, there is danger in motor vehicle numbers. For mid-block locations, the exponent for motor vehicle volume is greater than one, meaning that a doubling of motor vehicle volume would cause a more-than-doubling of cyclist collisions.

Allen-Munley, Daniel and Dhar (2004) develop an APM for Jersey City, New Jersey based on 314 collisions and several street characteristics. These researchers use no exposure data, instead looking only at collision severity. They find that more severe collisions are associated with, among other factors, trucks, wide lanes, steep grades, highways, one-way streets, and *lower* motor vehicle volume. Their explanation for this last counterintuitive finding is that perhaps speeds on heavily congested facilities are lower. It is worth noting that wide lanes, steep grades, highways and, according to the authors, even one-way streets, are all associated with higher motor vehicle operating speeds, a variable not included in the analysis.

Two studies, both in North Carolina, use only bicycle-motor vehicle collision data with no exposure measure and examine factors which influence the severity of crashes. Klop and Khattak (1999), looking at 1990-1993 data for two-lane state roadways, find that steep grades and high speed limits increase the likelihood of severity, while traffic volume decreases it. Kim et al (2006) use 1997-2002 data for all roads in the state, and find several factors which at least double the likelihood of a collision being severe, including high motor vehicle speeds and truck involvement.

2.4 Hypothesis formulation

The literature on bicycle lanes, tracks and paths is fairly mixed, with some studies suggesting a positive safety benefit and others a negative or no effect. Likewise, the literature on major vs. minor streets has been mixed and generally of poor quality; three of four studies reviewed above find that minor roads are safer (Kaplan 1975, Moritz 1997 and Tinsworth 1994 but not Moritz 1998).

There are a number of good reasons to expect that side streets ought to be safer for cyclists than the arterials which they directly parallel. Arterials tend to have high motor vehicle volumes, high speeds and function as truck routes and public transit bus routes. Turner, Roozenburg and Francis (2006) find, intuitively enough, that where there are more motor vehicles, there are more collisions per cyclist. Allen-Munley et al (2004) find that truck involvement and several factors associated with high motor vehicle speeds are correlated with more severe collisions. Kim et al (2006) also find that truck involvement and high speeds are correlated with severe injuries. Though Klop and Khattak (1999) and Allen-Munley et al (2004) both find that high traffic volume is correlated with a smaller proportion of collisions being severe, this does not mean that cyclists would be safer overall on busier streets, since Turner's model would predict that, compared to quiet streets, busy streets would have far more collisions to begin with. Further, those studies compared roads across an entire state (North Carolina) and city (Jersey City, NJ), meaning that unobserved street design and street network variables can creep in—the high-volume streets may also be more urban in character.

Though the literature is sparse, it appears that for cyclists to use minor side streets with low motor vehicle volumes and speeds, which are not truck or bus routes, should reduce their collision rates and severities substantially compared with traveling on major streets. In other words, side streets ought to be safer alternatives to arterials. It also follows that encouraging the use of side streets through treating them as bicycle boulevards ought to be a good strategy for improving city-wide cyclist safety.

To my knowledge, no study has yet examined the safety of bicycle boulevards relative to arterials. Walker, Tressider and Birk state that "Although the safety benefits specifically attributed to bicycle boulevards has [sic] yet to be studied, the safety benefits of traffic calming are well documented to reduce both the frequency and severity of collisions" (2009, 58).

Based on the findings of the above literature, I formulate the following hypothesis:

Side streets are safer for cyclists than adjacent parallel arterial routes, with both a lower collision rate and a lower proportion of collisions that lead to severe injury.

Whereas much of the literature I review above asks the question, "All else being equal, what is the safety effect of X intervention," my hypothesis starts from the point that all else is not equal in real life: cyclists—and cities—often face a choice between two streets that are different in a great many respects: traffic volume and characteristics, land uses, roadway width, driveway frequency and so on. Instead of controlling for these variables, I package them and treat the package as a study variable. I ask, "Of two streets A and B, which has the safer package of attributes?"

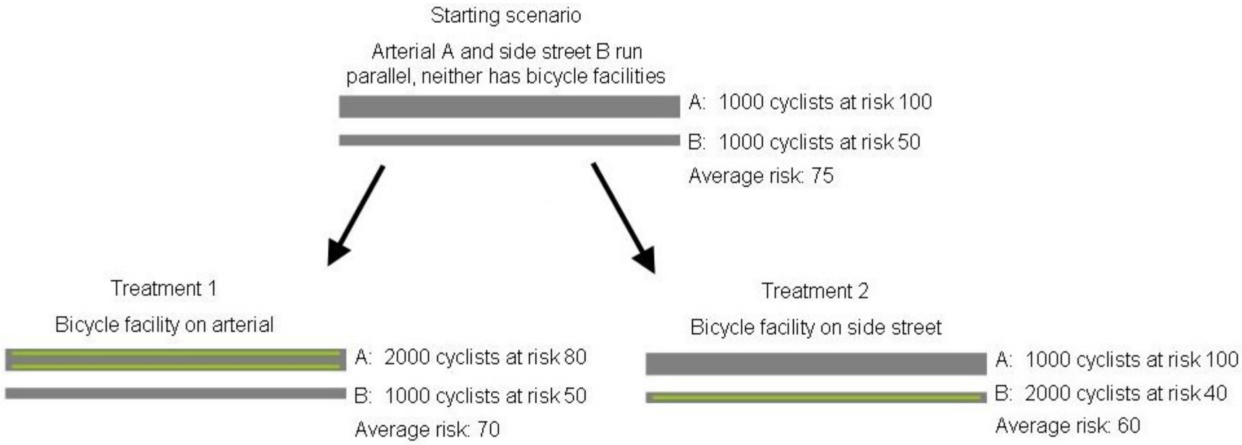
An analogy to dietetics may be useful. One strategy for studying the effects of diet on heart disease, obesity and so on is to isolate the effects of a particular nutrient or ingredient—for instance, fiber, saturated fat or trans fat. Another strategy is to compare the health effects of two complete dietary choices: for instance, a fast food diet versus the Mediterranean diet. Critics will point out that such an approach has a low resolution or specificity, since it does not seek to determine *what about* the Mediterranean diet is healthier. On the other hand, for some patients this comparison may represent a realistic choice of what to eat, and precisely since it touches on so many different underlying factors, a doctor's advice to "eat Mediterranean" may be more effective in producing positive health outcomes than advice to "cut out saturated fat."

I believe that research which attempts to isolate the safety effects of a bicycle facility (like trying to isolate the health effects of trans fat) is useful. I have chosen here to take the approach of looking at the whole street (like looking at the whole diet) because it is missing from the literature to date and I believe the answer to my question is useful to both cyclists and cities.

It is valuable to cyclists because it represents a real choice that they face. For instance, when I bike home from M.I.T. at the end of the day, I can ride on Massachusetts Avenue, a busy arterial with buses and trucks but also bike lanes, or on Green Street, a one-lane, one-way street more residential in character but with no bicycle facilities. The two streets are one block apart and either will allow me to reach home. My decision might hinge on which street I believe is safer.

The answer is valuable to cities because, when they use a design intervention to encourage or enable cycling, they may wish to know that their intervention will be encouraging cyclists to use the safer of two routes. Hypothetically, imagine a city with no bicycle facilities but with known bicycle-motor vehicle collision rates on two streets, A and B, each of which are currently traversed by 1000 cyclists per day. Cyclists on A experience a relative risk of 100, and those on B a relative risk of 50. This could be measured, for instance, in collisions per million bicycle miles traveled, but the units are unimportant here. It is believed that, for the street where they are applied, bike facilities reduce the collision rate by 20%, and also stimulate cycling by 100%. The city has funds and political will to treat only one of the two streets. Consider the two treatments depicted in Figure 7.

Figure 7. Two approaches to accommodating cyclists. In this hypothetical scenario, the citywide risk reduction achieved depends on which street is safer to begin with.



In either case, the city can boast that it has improved cyclist safety, because the street where bicycle facility was applied is now safer than it was. This way of thinking is supported by most of the literature on bicycle facilities. Yet consider the safety implications of the two approaches: In the first case, averaged over all cyclists on both routes, the relative risk faced is 70. In the second case, the average relative risk faced by a cyclist is just 60.

Of course, this hypothetical scenario is terribly simplistic. In real life, the relative risk could never be known so precisely; the ability a bicycle facility to stimulate demand would be unequal on two different streets, depending, for instance, on how well the route aligns with cyclist desire lines; different types of cyclists would use the two different routes, and so the relative collision rates might represent a difference in cyclist behavior and not in underlying safety characteristics of the street; and so on.

The point of this exercise, though, is to show that when a city installs bicycle facilities, it is probably anticipating an increased volume of cyclists, and so it might do well to consider whether it is encouraging those cyclists to ride on the safest route. Naturally, this is not the only consideration—other factors may include accommodating cyclist desire lines, working within political constraints of where street capacity can be turned over to cyclists, and promoting the visibility of cycling as a lifestyle in the city. Furthermore, route choice need not be "either-or"—in the above example, the city would best serve its goals of promoting cycling and improving cyclist safety if it installed bike lanes on both streets. While many cyclists would prefer to live in a city where every single street was safe and comfortable for cycling, the reality is that no city upgrades its entire street network overnight: cycling improvements are made piecemeal, and an increase in cycling after one intervention may be needed in order to justify the next intervention.

Therefore this research is not intended to address all aspects of a city's decision of how to accommodate cyclists, nor is it intended to portray what type of street network would be ideal for cyclists. Besides being useful to cyclists deciding where to ride, it will probably be most useful to cities that currently have few or no bicycle facilities and are deciding where to begin.

3 Methodology

3.1 How to measure safety

There are many different ways to measure safety. For a moment, leave cyclists aside and consider road safety in general. The total number of traffic fatalities in the U.S. is one measure of road safety in this country, but a lousy one: suppose that the number of traffic fatalities has increased since last year, but by a proportion less than the increase in population. That means that the number of fatalities per person—or per 100,000 people—has dropped, which in fact seems to have made everyone safer. However, suppose that the aggregate number of vehicle miles traveled (VMT) dropped in this same year, despite the population increase. That would mean that the rate of fatalities per VMT has risen even though the rate of fatalities per population has dropped. In this case, are we safer than last year? It depends on your perspective and goals. If you are a citizen who for whatever reason must drive the same amount this year as last, you should be concerned that each mile is more dangerous than last year, so your risk is greater. If you are a public policy maker concerned with helping your citizens, on average, live longer, then fewer deaths per population is an improvement.

Death, meanwhile, is not the only risk that road users face. Suppose that emergency medical services improve, meaning that despite the same number and severity of collisions, more people survive them. This is an improvement in public health, but not in road safety, and one must ask all sorts of related questions—are we restoring people to full health or merely to paralysis? In any event, people on the road are (or should be) scared not only of dying, but also of paralysis and other serious injuries, of expensive hospitalization, and even of property damage, though it seems trivial in comparison.

For the purposes of this study, I am concerned with severe injuries per bicycle mile traveled. For the numerator, I choose severe injuries because fatalities are thankfully too rare to study at a street-by-street level, yet I believe it is still important to somehow differentiate severe incidents from minor incidents. The severity of injury is of great importance to the cyclist injured, of course, and literature I review above suggests that the likelihood of severity depends on traffic characteristics. For the denominator I choose bicycle miles traveled, because this reflects the extent to which cyclists expose themselves to the risks of the road. It is also possible to imagine making cyclists safer by densifying cities and bringing destinations closer, so that each bicycle trip is fewer miles, and *cyclists* are safer even if *bicycle miles traveled* are not safer. However, this study seeks to inform cyclists' daily choice of route and cities' medium-term choice of which routes to encourage cyclists to use. Long-term changes in overall urban density are quite beyond the scope of this study.

I focus my study on bicycle-motor vehicle collisions. I formulated my hypothesis based on the motor vehicle traffic characteristics of side streets, so it is only bicycle-motor vehicle collisions that I expect to be less common and less severe there. In my literature review I did not uncover any theory or evidence as to what causes two-cyclist collisions, cyclist-pedestrian collisions or one-party cyclist falls. I have no particular reason to believe that such crashes would be reduced on side streets compared to arterials, and so I exclude them from the main course of my study.

Bicycle-motor vehicle collisions form the bulk of serious risk for cyclists: in the United States in 2006, 688 cyclists died in "motor vehicle traffic" while 238 died due to "other" causes (Heron et al 2009, Table 18). This means that motor vehicle traffic was responsible for 74% of cyclist deaths in that year.

Still, other types of crashes do form part of the full profile of risk for cyclists. If I found, for instance, that one street type has a significantly lower rate of bicycle-motor vehicle collisions but a higher rate of bicycle-pedestrian collisions, then it might not be reasonable to conclude that street type is "safer." For this reason, I also include an analysis of all crashes in the available datasets in Appendix 3 (Berkeley) and Appendix 5 (Cambridge). It is worth cautioning, though, that data on other crashes is likely very incomplete. For instance, the collision data for Cambridge appears to include only three single-party falls, though it seems likely that more than three cyclists had a fall in six years. Such crashes may be even more underreported than bicycle-motor vehicle collisions.

The risk of severe injury from bicycle-motor vehicle collisions may be expressed as shown in Equation 2.

Equation 2. An expression for cyclist risk

$$risk = \frac{severe\ injuries}{bicycle\ miles\ traveled}$$

To study this risk, it is useful to decompose that risk into two terms as shown in Equation 3.

Equation 3. Decomposition of severe injury risk

$$\frac{severe\ injuries}{bicycle\ miles\ traveled} = \overbrace{\frac{severe\ injuries}{bicycle\text{-}motor\ vehicle\ collisions}}^{\text{Term 1}} * \overbrace{\frac{bicycle\text{-}motor\ vehicle\ collisions}{bicycle\ miles\ traveled}}^{\text{Term 2}}$$

This decomposition is useful because in some cases it is only possible to study Term 1, for instance due to lack of exposure data, while in other cases it is only possible to study Term 2, for instance if too few collisions are severe to draw meaningful conclusions. Of the studies I review above, most focus on Term 2, collision rate, without regard to the severity of injury. A handful of them look only at Term 1. In my study of Berkeley and Cambridge I compare Term 2 street by street, but since severe injuries are thankfully rather rare, I compare Term 1 across the entire city.

As Hauer (2001) points out, the specification of Term 2, bicycle-motor vehicle collisions / bicycle miles traveled, is not quite right. Since bicycle-motor vehicle collisions can involve more than one bicycle, (bicycles involved in collisions with motor vehicles) / (bicycle miles traveled) would give an unbiased estimate of the risk to each cyclist. However, at least in the datasets that I use here, multi-cyclist collisions are exceedingly rare; almost all collisions with motor vehicles involve just one bicycle.

Since it is simpler and makes no meaningful difference in the results, I use collisions, not cyclists, as my numerator.

3.2 Application to Berkeley and Cambridge

As applied to Berkeley, my hypothesis is that each bicycle boulevard is safer than the arterial(s) to which it is an alternative. Most of these arterials are untreated for cyclist use. In Cambridge, my hypothesis is that, where side streets run parallel to arterials, they are safer than the arterials. The arterials I test are treated for cyclist use, to varying degrees.

Therefore, as applied to both cities, the hypothesis becomes this: side streets are safer than parallel arterial routes regardless of whether either street is treated for cyclist use.

This hypothesis requires a rigorous definition of "safer." Side streets generally have lower motor vehicle volume than arterials so I expect them to have a lower collision rate. They also generally have fewer heavy vehicles and lower speeds, so I expect a smaller proportion of collisions there to be severe. This means that both terms of Equation 3 will be smaller, so the overall risk of severe injuries per exposure unit will also be smaller.

Comparing parallel streets requires having comparable street segments. For each pair to be compared, I select segments of identical length, starting at ending at the same cross streets. I then select collisions that occurred along the relevant road segment for each street. Since my interest is in which of the two routes cyclists choose, I am interested in collisions that happened to cyclists traveling along, or at least turning onto or off of, the street of interest. I do not include collisions that happened on cross streets, or at intersections but to cyclists traveling perpendicular to the street of interest.

Next I need to divide the number of collisions for each street by an appropriate exposure measure. The ideal measure of aggregate cyclist exposure to a particular street's risk would be bicycle miles traveled on that street. A collision rate defined as collisions divided by bicycle miles traveled would show the risk that each cyclist faces for every mile traveled on a given street. In practice this is not possible. The nearest proxy available for bicycle miles traveled is a count of cyclists. While this cannot give an absolute measure of risk such as collisions per bicycle mile traveled, it can give the relative risk of two streets. If 500 cyclists are observed on street A and 1000 on street B and identical segments of the two streets have an equal number of collisions, then a best estimate is that street A has twice the collision rate of street B. In order for cyclist counts to be a good measure of exposure, several things must be true:

- The counts are taken at the same point along A and B—for instance, at the same cross street.

- The bicycle traffic at that particular point is proportional to bicycle miles traveled along the whole street segment under study.

- The counts on A and B must be taken at the same time, for the same length of time, and bicycle traffic during that time segment must be representative of the entire day and the entire year.

The advantage of looking at parallel routes, besides the fact that it represents a real-life choice for cyclists and cities, is that it automatically controls for some of these variables. It would be implausible to suggest that a two hour afternoon peak count in the fall is in some way representative of the amount of cycling that goes on in an average two hour period of the entire year. However, it is not nearly as hard to believe that the *ratio* of bicycle traffic on two parallel routes that are close substitutes for one another is relatively constant over time.

Similarly, even if the cyclist count at one point along a street is not representative of bicycle traffic everywhere along the street, the ratio of the parallel streets' counts might still be relatively constant. Not necessarily, though. Imagine a simple city with just one street, along which all residents live at a uniform density and ride bicycles everywhere they go. Everyone rides to work and other destinations, all of which are uniformly distributed along the street, in locations uncorrelated with employees' and customers' home locations. In this city, one would expect traffic to be heaviest at the center of the street segment and taper off to zero at both ends. Since Berkeley's bicycle boulevards are of finite length, ending at city limits or sooner, one might expect low bicycle traffic at their ends. The same is true of Cambridge's unconnective side streets. Both cities' arterials, however, continue into nearby towns, and thus might have higher bicycle traffic at the point where the parallel side street ends. So the assumption that the ratio is relatively constant anywhere along the two streets is worth questioning and, to the degree that data allows, I do examine it later in this thesis.

In practice, the count data available are intersection counts—counts of cyclists at an intersection grouped by approach and direction of travel. So for a standard four-way intersection there are 12 counts: straight, right and left from each approach (cyclists almost never make U-turns at intersections). This raises the important question of how to count turns. I reason that I could determine north-south flow by counting at a mid-block location just to the north or south of my intersection. Cyclists traveling straight through the intersection on the north-south street would be counted in either case, but a count just to the north would capture cyclists making four of the eight possible turns at the intersection, and a count just to the south would capture cyclists making any of the other four possible turns. The equivalent is true for east-west travel. Therefore every cyclist that turned, I counted as half a cyclist on the north-south street and half a cyclist on the east-west street, whereas cyclists that proceeded straight were counted as whole cyclists for their respective streets. This method uses all the data available at the intersection, and has a smoothing effect—it is the same as if I conducted four mid-block counts and determined north-south and east-west flow each as averages of two counts on opposite sides of the intersection.

Dividing the number of collisions by a count-based measure of exposure estimates Term 2 of Equation 3, the collision rate. I also test Term 1, the proportion of collisions that are severe. Since just a few percent of collisions are labeled as severe in either Berkeley or Cambridge, it would be difficult to compare the proportion of collisions that are severe on a street-by-street basis. Instead I conduct a city-wide comparison of the proportion severe on different street types.

3.3 Establishing statistical significance and causation

To examine the statistical significance of my findings requires a number of different approaches. Comparing the proportion of collisions that are severe on side streets versus arterials is easy enough: a two-tailed difference of proportions test will do the trick. But comparing collision rates is complicated. Statistical tests are generally designed for random samples, which is not the nature of the underlying data here. The collisions are not a sample at all, but in fact the complete set of police-reported collisions for the period of study. This, in turn, is some subset of all collisions that occurred, but by no means a random one—for instance, more severe collisions are probably more likely to be reported. The counts are not random samples either. They are samples designed to maximize the number of cyclists counted: they were obtained at peak hours at major intersections. This raises important questions about their representativeness of total volume.

In this section I address separately the statistical and uncertainty issues about three figures: collision rate or collisions/counts, collisions alone, and counts alone. Then I put this all together to reiterate the importance of comparing only comparable parallel street segments, and finally discuss how to derive causation from correlation and show that street typology is at least one reason why collision rates differ.

3.3.1 Collisions/counts

Although the collisions are not a sample but a complete set of police-reported incidents, it is still appropriate to treat them as a sample. This is because one year is, at some level of simplification, a sample of what happens in a typical year. In other words, if we could magically run 2003 over again, we might see different results the second time around. A simple model, ignoring the reality that different cyclists probably have different risk levels, is that collisions are produced on each street as a result of a Poisson process with some constant collision rate per bicycle mile traveled, or a Bernoulli process with some constant collision rate per cyclist trip. The period of study is a sample of each street's process, and one would like, based on this sample, to estimate the underlying parameter of each street's process. To the extent that street design is the same today as during the period of study, this allows us to predict which street is safer to ride on today.

This view of the collisions necessarily assumes that one knows perfectly the number of trials on each street that produced the collision results. Assuming this for now, and pledging to come back to the issue of count reliability later, it is possible to compare the collision rates of each bicycle boulevard-arterial pair. This asks the question, "Assuming we know the number of collisions and number of cyclist trips perfectly, then for each pair of streets, what is the chance we would see these results if in fact the two streets were equally safe?" I answer this question using Poisson regression, which I describe in more detail in Appendix 2.

3.3.2 Collisions

The collision numbers themselves represent not a sample, but the complete set of police-reported collisions for the period of study. However, it is well-known that bicycle collisions are underreported, so the validity of the collision numbers for this analysis could still be compromised if the reporting rate differed for different streets. Since bicycle boulevards and arterials are so different in character, they may attract different types of cyclists with different levels of inclination to report collisions. Imagine, for instance, that side streets are populated primarily by risk-averse cyclists who get very scared in even a minor motor vehicle collision and choose to report it to police, whereas arterials are used by fearless cyclists who do not think twice about a minor scrape. In this case, the true difference in collision rates between the two street types would be even larger than my numbers show. Conversely, it is also possible to imagine that side streets attract moderate cyclists who do not care to report a minor scrape, while arterials draw highly political cyclists who will report any collision just for the sake of reporting it. This would mean that at least some of the safety difference I have shown is an illusion created by reporting rates.

I am not aware of any surveys or studies which have addressed which type of cyclist is more likely to report a collision with a motor vehicle. There is no statistical test that I can apply to address the issue of inclination to report, so I limit discussion of this to a quick sensitivity test: if A has twice as many reported collisions as B, it would need to have twice the reporting rate in order for the true number of collisions to be identical.

Differences in collision severity rates could also cause a difference in reporting rates. For instance, suppose street A has twice the rate of reported collisions as street B and the two streets have an equal proportion of their collisions designated as severe. This could be because collisions truly are more likely to occur (per cyclist) on street A, but are equally likely to be severe on either street. Or it could be that collisions are equally likely to occur (per cyclist) on either street, but twice as likely to be severe on street A, and so twice as many are reported. With only reported collisions to study, it is unfortunately not possible to determine which of these scenarios is the case. Thankfully, the *product* of collision rate and proportion severe—risk of severe injury per cyclist, the ultimate dependent variable of this study—is the same either way.

3.3.3 Counts

There are three questions it is important to ask about the count data:

1. How consistent are counts on a given street over different points along that street?
2. How consistent are counts at a given point on a given street over different years?
3. How well does weekday PM peak traffic represent total weekly traffic?

The answers to these questions are probably something like “not very.” But remember, I am not computing absolute collision rates, but only relative rates between two parallel and adjacent or nearly adjacent streets which function as alternate routes to reach the same destinations. This makes these questions somewhat tamer:

1. How consistent is the ratio of counts on A to B at different points along the two streets?
2. How consistent is the ratio of counts on A to B over different years?
3. How well does the ratio of A's weekday PM peak traffic to B's represent the ratio over the whole week?

Where data are available, these questions can be asked on a case-by-case basis. Question 1 could be addressed simply by comparing the ratio of counts on A to B at several different points, Question 2 by comparing the ratio of counts on A to B over several years at the same point. Question 3 could be answered by comparing the ratio of counts on A to B at AM peak, PM peak, mid-day, evening, weekends, and so on. For each question, the more consistent the ratio is found to be, the better an estimate of true volume we may believe the counts to be. Unfortunately count data is limited and it is not possible to adequately answer all of these questions for Berkeley or Cambridge. I address these questions to the extent possible in Analysis and Results.

3.3.4 Synthesis

The complications I describe above underscore the importance of limiting my comparisons to parallel adjacent streets. Comparing the collision rate on one street in Berkeley to one street in Cambridge would be exceedingly difficult—differences in weather and cycling seasonality would render count data non-comparable, different standards for reporting between different police departments might bias collision data, and so on. Moreover, the two cities have different demography, topography, and so on—so differences in collision rate might not be due to street design or typology. Instead I compare parallel routes within each city, and in the Conclusion, I also compare the patterns and differences observed—or not observed—between cities.

3.3.5 Causation

Where I do find that there is a difference in collision rate between side streets and arterials, it is still important to ask what drives the difference. My hypothesis is based on motor vehicle traffic characteristics, and so I expect that this is what is causing any difference in safety observed. However, there are other imaginable explanations for any observed difference in safety.

For instance, perhaps self-selection is at work: arterials street attract "rogue cyclists" who by nature are risk-seeking: they ride fast, they run red lights, and so on. Perhaps side streets attract decent, law-abiding, risk-averse cyclists. The difference in observed safety between the two streets could be due to a difference in the innate risk behavior of the two populations of cyclists rather than due to motor vehicle characteristics or to any difference in risk behavior brought about by street design. This theory is dramatized in Figure 8.

Figure 8. The self-selection theory: arterials and side streets simply attract different types of cyclists who carry with them different collision rates

This may be part of the story, but in order for this theory to account for all of the observed difference in safety, it would have to be true that the cyclists carry all the risk with them—force a rogue cyclist to ride on a side street instead of his preferred arterial, and he will still manage to get into just as many scrapes with motor vehicles, even if there are far fewer motor vehicles to get into scrapes with.

I address this proposition by looking at which party police found at fault in collisions in each city. Findings of fault may not tell the full story of who or what brings about a collision, but they can provide some insight. If cyclists are found at fault in the vast majority of cases, this would be consistent with the self-selection theory, since cyclist behavior is what brings about risk, while if drivers are more often found at fault, that would suggest that cyclist behavior cannot account for most of the difference in risk.

Another point is that the self-selection theory can only explain away streets' collision rate differences to the extent that any given cyclist behaves the same way no matter which street she rides on—in other words, street design has no effect on behavior. It's not hard to imagine, though, that the same cyclist who rides in the street on a bicycle boulevard might ride on the sidewalk on an arterial, and so on. Unfortunately, I do not have data on individual cyclists' behavior on different streets for either city.

Another theory is that of causative "safety in numbers"—perhaps, for any given level of exposure, a side street and an arterial would have an equal collision rate, but one street type carries more bike traffic, achieving safety in numbers, and so has lower collision rates. In Berkeley I address this issue by plotting collision rate versus exposure for each street and showing that the two street types differ in collision rate at any level of exposure. In Cambridge, I simply observe that the street with more exposure is not found to have a lower collision rate.

4 Data

4.1 Berkeley

4.1.1 Background

Berkeley is a city in the Bay Area in Northern California. It is located in Alameda County on the eastern shore of the San Francisco Bay, an area known as the East Bay. Berkeley is best known for the University of California flagship campus there, but also features a lively downtown and transit access to the Bay Area's larger employment hubs in San Francisco and Oakland.

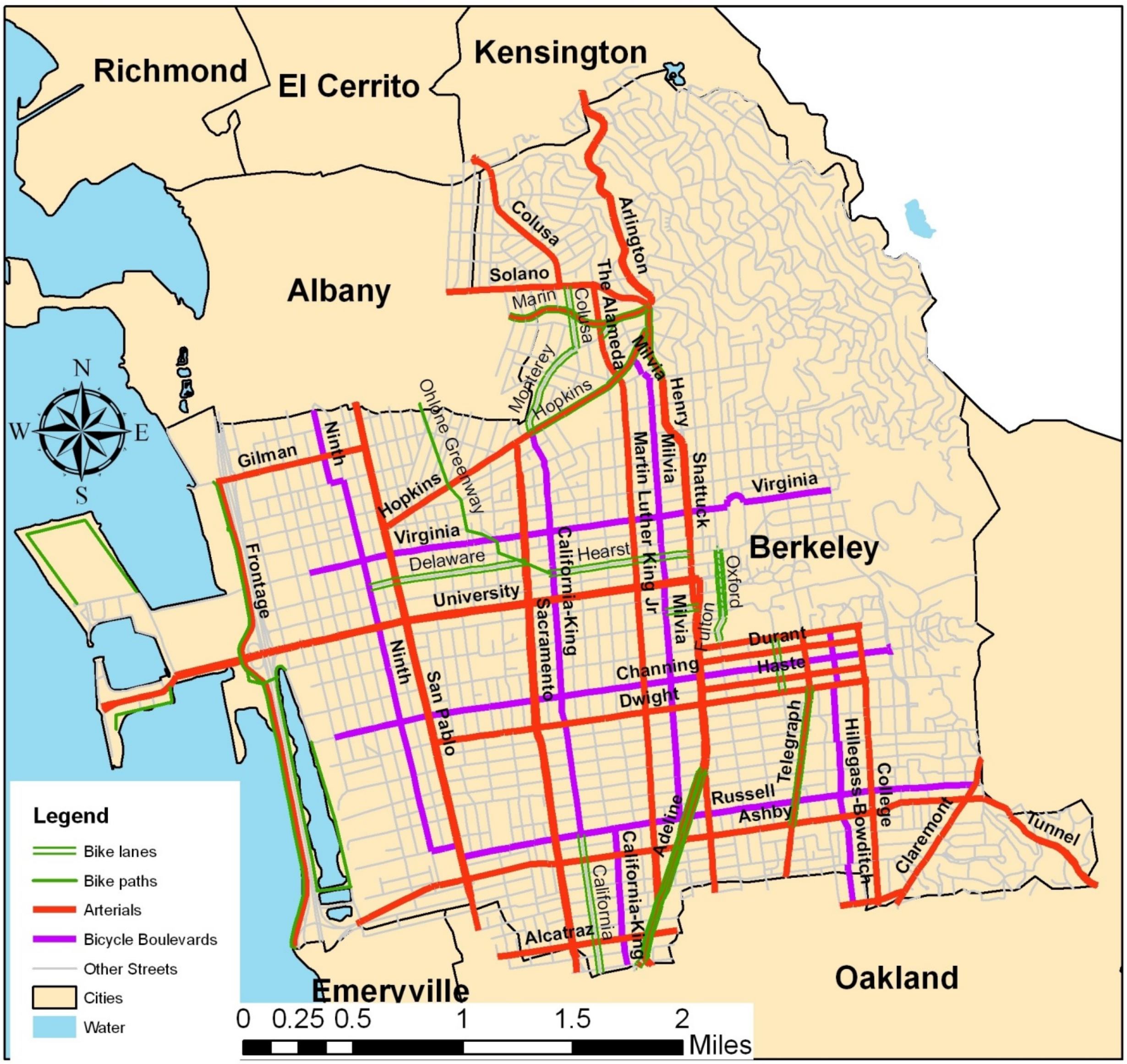
Figure 9. Berkeley viewed from the eastern hills. UC campus front and center, downtown slightly to the left, the "flatlands" beyond and the marina jutting into the San Francisco Bay.

With a population of around 109,000 people (2006-8 American Community Survey data, accessed through Bay Area Census) spread over 10.5 square miles, Berkeley's population density is about 10,000 persons per square mile—somewhat less than the 16,000 or so of Cambridge, MA. About 17% of Berkeley's 45,000 households do not own cars, compared to 11% in Alameda County and 9% nationwide. 5.6% of workers who live in Berkeley bike to work, compared to 1.2% in Alameda County and 0.4% nationwide (United States Census 2000).

Around the time those census data were collected, the city embarked on an ambitious program to create a bicycle boulevard network throughout the city. According to the city's Bicycle Boulevard History page, city council adopted a bike plan in 1999 which called for, among other things, a network of seven bicycle boulevards. Following a public engagement process, the city began implementing the bicycle boulevards in 2001, with bulk of the work completed in winter 2002-3 (City of Berkeley, Bicycle Boulevard History). The regular street grid that covers most of the city is especially well suited to a bicycle boulevard strategy, since parallel and adjacent to each arterial are quieter neighborhood streets that are equally direct and connective.

The bicycle boulevard network, shown below, forms the backbone of Berkeley's bicycling network. Berkeley also has a few streets with bike lanes, and a bike path fronting the Bay, reachable by a ped-bike bridge over I-80, and the Ohlone Greenway, a short segment of bike path running on top of the BART right-of-way near North Berkeley Station. However, as can be seen in Figure 10, the bicycle boulevard network is quite extensive, basically covering the whole city, and each bicycle boulevard runs parallel to at least one arterial road, providing a quieter alternative for cyclists to reach key destinations on the arterials. For these reasons it seems fair to say that bicycle boulevards have been, to date, Berkeley's major strategy for accommodating cyclists on its roads.

Figure 10. Seven bicycle boulevards (purple) run parallel to key arterials (red).



Streets: ESRI data & maps 9.3 [electronic resource]. Redlands, CA : Environmental Systems Research Institute, Inc., c2008.
City and water: ESRI StreetMap premium [electronic resource] : North America. [Redlands, Calif.] : ESRI, c2008-2009.
Arterials, bicycle boulevards, bike lanes and bike paths selected from streets or drawn manually based on City of Berkeley Bikeway Map available at http://www.ci.berkeley.ca.us/uploadedFiles/Public_Works/Level_3_-_General/Bikeway_Network.pdf

The city's Bicycle Boulevard Network web page provides a list of major destinations reachable by bicycle boulevards and lists the arterials for which each is a good alternative. This information is summarized in Table 1.

Table 1. Berkeley's seven bicycle boulevards and the arterials to which they are alternatives

Bicycle Boulevard	Arterial
Ninth	San Pablo Ave.
California-King	Sacramento St.
Milvia	Shattuck Ave. Martin Luther King, Jr. Way (MLK)
Hillegass-Bowditch	College Ave. Telegraph Ave.
Virginia	University Ave.
Channing	Dwight Way
Russell	Ashby Ave.

Source: City of Berkeley, Bicycle Boulevard Network. I have modified the city's table by eliminating Sixth St. and Cedar St., which are somewhat less arterial in character than others, and by considering only Virginia, and not Channing, to be an alternative to University Ave.

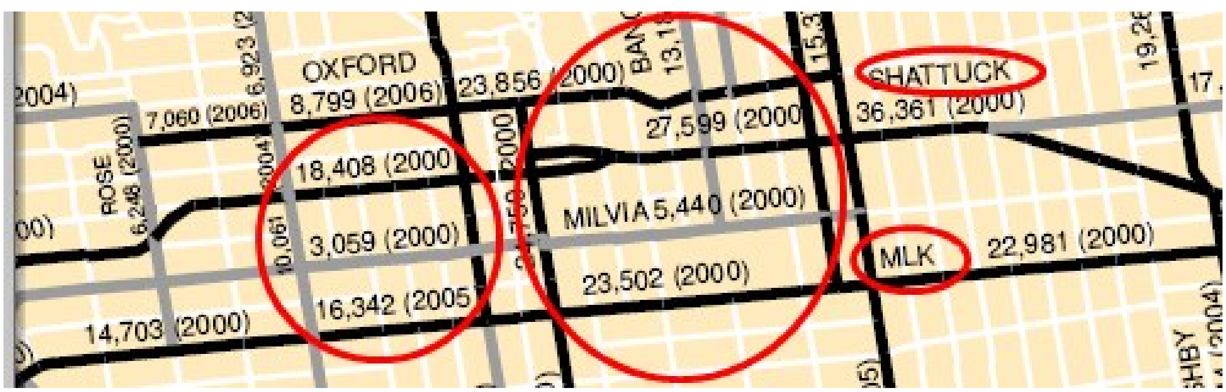
Each bicycle boulevard is a two-lane street with two lanes of parallel parking, street trees and sidewalks. The land uses along these streets are almost exclusively residential—though a notable exception is the central section of Milvia St. which runs through downtown past Berkeley High School, City Hall and a number of other institutional buildings. The arterials, by contrast, are highly connective through streets, most of which are at least partly lined with commercial or industrial uses. Most are four lane divided roads, though not necessarily for their entire length, and serve as truck routes and public transit bus corridors.

Therefore, the three risk factors on which I base my hypothesis—heavy vehicles, speed and volume—all differ between the two street types. Since arterials are often truck and bus routes, it seems clear that heavy vehicles are more common there than on bicycle boulevards, though no counts are available. As for speeds, though official speed limits are 25 mph on all streets in Berkeley save a few arterials such as Sacramento St. and San Pablo Ave. where 30 mph is the limit (City of Berkeley, Speed Limits), speed limits are rarely posted. The design and land use characteristics of the bicycle boulevards, described above and depicted in Figure 11, may suggest to the driver a lower acceptable speed than on arterials.

Figure 11. Ninth St. (left), a bicycle boulevard, looks and feels very different from parallel San Pablo Ave. (right), an arterial.

One would expect motor vehicle volumes to be much higher on the arterials than on the bicycle boulevards. The motor vehicle traffic counts that the city has makes available are mostly from 2000, and for the most part only arterials were counted. The only bicycle boulevard included is Milvia St., which carried around a fifth as many motor vehicles as did Martin Luther King Jr. Way (hereafter MLK) and Shattuck Avenue, the two arterials which flank it—see Figure 12. Milvia St., as noted above, is the one bicycle boulevard that traverses a significant section of dense, non-residential development. For other bicycle boulevards, one would expect an even steeper ratio. When I observed California St., a bicycle boulevard, and Sacramento St., an arterial, during morning rush hour, I saw about 10 or 20 cars on California St. in the entire two hours I spent there, while at least 20 cars passed every minute on Sacramento St. This suggests a difference in motor vehicle traffic volumes of about two orders of magnitude.

Figure 12. City of Berkeley motor vehicle counts. Milvia carries about a fifth as many cars as its neighbors.



Source: City of Berkeley, Average Total Daily Traffic Volume.

This huge differential in traffic volumes is probably not due solely to the streets' inherent characteristics. A splendid variety of different traffic calming techniques have been applied to the bicycle boulevards, along with signage and pavement markings intended to make the streets recognizable as part of the bicycle boulevard network. It is worth noting that some of the traffic calming devices and bike lanes were present before the streets were designated as bicycle boulevards (City of Berkeley, What is a Bicycle Boulevard?). A photographic survey of the bicycle boulevards follows.

Branding and wayfinding

Berkeley has branded its bicycle boulevards in purple with white lettering. Periodic signs alert the user that she is on a bicycle boulevard; some signs also provide information about destinations reachable by bicycle.

Figure 13. Branding and signage of Berkeley's bicycle boulevards

Reduction of motor vehicle volumes

The bicycle boulevards use a number of different barriers to prevent automotive through-traffic while allowing bicycles through. Vehicular access is preserved for every property—someone driving to an address on a bicycle boulevard must simply travel most of the way on an arterial, and then turn onto the bicycle boulevard for the final few blocks. Types of barriers include curbs, flowerpots and even dead ends at public parks. Most of these are permeable to emergency vehicles.

Figure 14. Curbs and flowerpots as traffic barriers on Milvia St.

Figure 15. This park on California St. is a dead end for motor vehicles, but bicyclists can pass through.

Similar to barriers, the boulevards also use diverters which force traffic to turn at a particular cross street. These include flowerpots and guard rails, as shown in Figure 16.

Figure 16. Guard rail diverters on Virginia St. (left) and flowerpot diverters on Ninth St. (right)

Reduction of motor vehicle speeds

Other traffic calming devices serve to reduce motor vehicle speeds. Speed bumps are used, as well as chicanes which narrow the street and make it curve back and forth. Traffic circles are common, sometimes along with stop signs but often without, which allows cyclists to pass through without slowing significantly.

Figure 17. Speed bump on Milvia St. (left) and traffic circle on Channing St. (right)

Figure 18. Chicanes on Milvia St.

Establishing of cyclist priority

Several other devices remind, or force, drivers to give cyclists priority on bicycle boulevards. Pavement markings serve as a near-constant reminder that cyclists can use the full lane. In a few places, such as the busy section of Milvia St. through downtown, there are bike lanes on the right instead of markings for cyclists to use the full lane. These bike lanes pre-dated the bicycle boulevard system; bike lanes were not planned as part of bicycle boulevard implementation (City of Berkeley, What is a Bicycle Boulevard?). Bicycle-activated signals help cyclists to cross busy streets, and sometimes protected areas are provided for cyclists to wait at the intersection.

Figure 19. Pavement markings on Virginia St. allow cyclists the whole lane (left); a bike lane on Milvia St. near downtown defines separate spaces for cars and bikes (right).

Figure 20. A bicycle-activated signal on Russell St., a bicycle boulevard, helps cyclists cross Telegraph Ave., an arterial.

Figure 21. Cars on Channing St., a bicycle boulevard, are forced to turn right on Martin Luther King Jr. Way, an arterial, but cyclists can proceed straight and have a curb-protected waiting area with a bicycle-activated signal.

It should be noted that not all of these traffic calming devices are unique to bicycle boulevards within Berkeley. Flowerpot diverters, for instance, are not uncommon in the city's residential neighborhoods, and some of those on the bicycle boulevards probably pre-date bicycle boulevard implementation. In any case, the combination of so many measures for traffic calming and cyclist priority creates an exceptionally calm and pleasant cycling environment.

The arterials by and large do not have any bicycle facilities (see Figure 10). One exception is Telegraph Ave., where the most minimal accommodations have been made for cyclists by painting sharrows on the street and offering brief segments of bike lane near intersections, fading to an edge line at mid-block locations. There are also wide painted bike lanes on part of Adeline St. near the Oakland border. Adeline St., which runs parallel to the King St. portion of the California-King bicycle boulevard, was ultimately not included in this study due to lack of data about King St.

Figure 22. This northbound bike lane on Telegraph Ave. runs for about 100 feet after the intersection with Ashby Ave.

Figure 23. Sharrows on Telegraph Ave.

Figure 24. Bike lane on Adeline St.

4.1.2 Collisions

Collision data for Berkeley comes from the Statewide Integrated Traffic Records System (SWITRS) maintained by the California Highway Patrol (CHP). This dataset includes all collisions reported to local or state traffic enforcement agencies (Metropolitan Transportation Commission, Safety Toolbox: Problem Identification). For bicycle collisions, the vast majority occurred on local streets and so data comes ultimately from local police reports. The Alameda County Public Health Department keeps a Microsoft Access database of SWITRS collisions in the county from January 1996 through October 2008, and a contact there generously provided this database for my research.

The SWITRS data is a relational database with three tables: collision, party and victim. The Alameda County database additionally contains a table of geocoded collisions which overlap with, but are not a perfect subset or superset of, the table of all collisions. It is not clear why this is so, and my contact at Alameda County was unable to explain further. I provide more details on the database schema and how I process the data in Appendix 1, while key decisions that I make as part of my methodology are outlined below.

The database contains an attribute for injury severity levels, the metadata for which are shown in Table 2. "PDO" stands for "Property Damage Only"—no injury. "Severe," as I use the term here, encompasses levels 1 and 2. Unfortunately, the metadata do not indicate that collisions are assigned severity levels by any objective criterion such as whether the cyclist was transported to the hospital in an ambulance. It may be that assigning severity levels is at the discretion of the police officer at the scene.

Table 2. Severity levels in the SWITRS database

1 - Fatal
2 - Injury (Severe)
3 - Injury (Other Visible)
4 - Injury (Complaint of Pain)
0 - PDO

The table of geocoded collisions contained 136,912 collisions. Of these, 7707 involved bicycles and of those, 1989 were in Berkeley. I run a handful of tests on these 7707 bicycle collisions to see how well they match up with general findings from the literature review, and to determine whether the severity levels are meaningful despite their lack of objective criteria. Table 3 shows that the data match quite well with the findings from, for instance, Kim et al (2006), as it appears that alcohol and heavy vehicles do make collisions more severe, and Marshall (2009), since Berkeley's dense urban street grid has produced fewer-than-expected severe and fatal collisions compared to the rest of Alameda County, much of which is suburban in character. I conducted a two-tailed difference of proportions test on each factor versus its reference case at the top of its section of the table. This asks, for instance, what is the chance that 72 of 487 alcohol-related collisions (14.8%) would "happen" to be severe if in fact alcohol did not increase the likelihood of severity above and beyond the 6.4% given by 491 of all 7707 collisions being severe? The answer: less than 0.1%. All of these findings are statistically significant at the .05

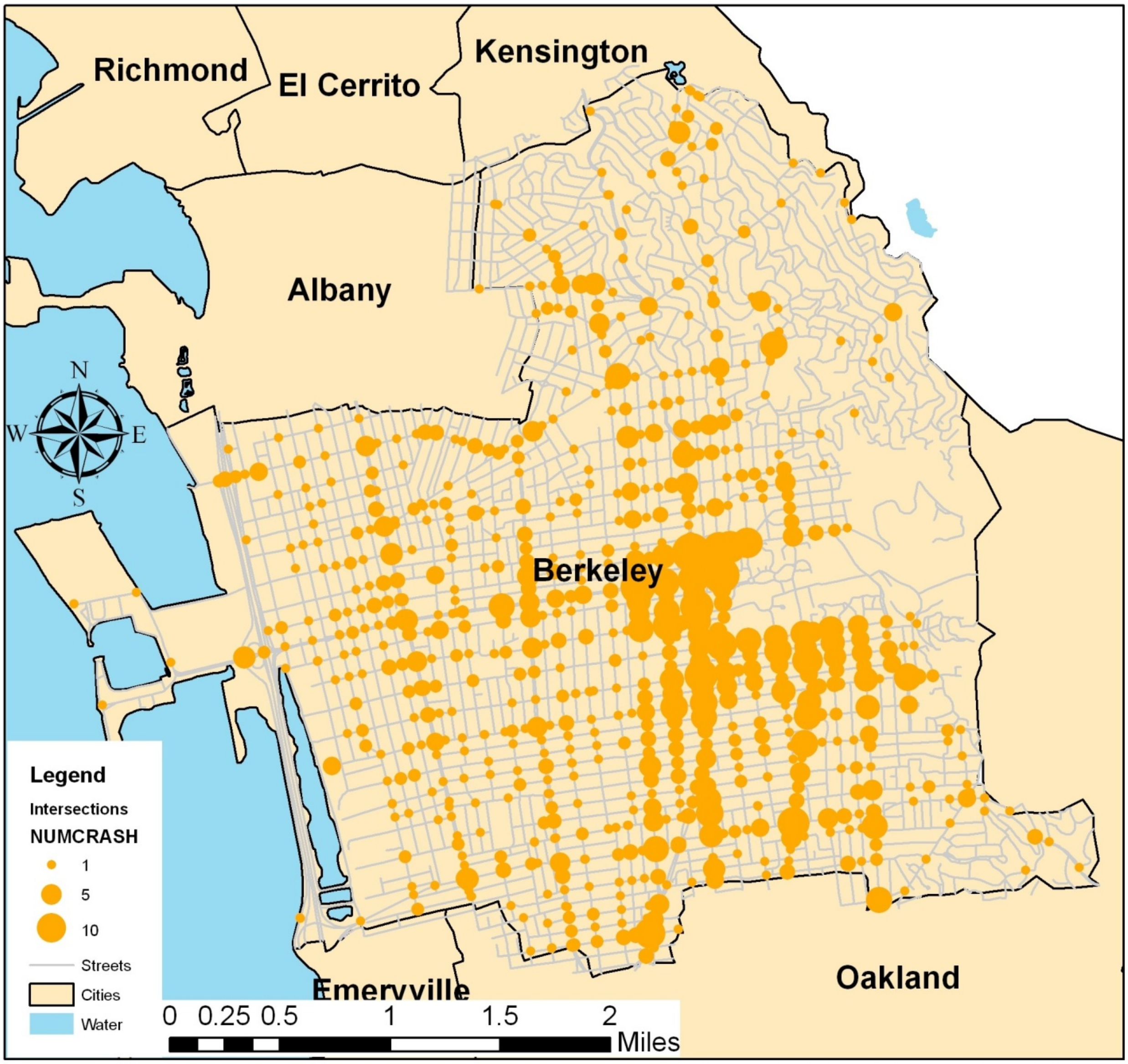
threshold or better. The results in Table 3 give confidence that the severity attribute in the SWITRS database is indeed meaningful.

Table 3. A few factors affecting severity in Alameda County bicycle collisions

	Collisions	Severe	Proportion severe	Significance
All	7707	491	6.4%	reference case
Involving alcohol	487	72	14.8%	$p < 0.001$
Involving a heavy vehicle	165	20	12.1%	$p = 0.005$
	Collisions	Severe	Proportion severe	Significance
All Alameda County	7707	491	6.4%	reference case
Berkeley Only	1989	102	5.1%	$p = 0.045$
	Collisions	Fatal	Proportion Fatal	Significance
All Alameda County	7707	37	0.48%	reference case
Berkeley Only	1989	1	0.05%	$p = 0.011$

The collisions table (as opposed to the *geocoded collisions* table) contained an additional 131 bicycle collisions in Berkeley which had not been geocoded. I used ArcGIS to geocode these collisions and found matches for 125 of them, for a total of 2114 bicycle crashes in Berkeley. The geographic distribution of these 2114 crashes is shown in Figure 25. Since they are geocoded snapping to the nearest intersection, the points appear on top of one another and had to be aggregated into bubbles sized for the number of collisions per intersection in order to display meaningfully.

Figure 25. Geographic distribution of 2114 police-reported bicycle crashes in Berkeley, CA 1996-2008



Streets: ESRI data & maps 9.3 [electronic resource]. Redlands, CA : Environmental Systems Research Institute, Inc., c2008.
City and water: ESRI StreetMap premium [electronic resource] : North America. [Redlands, Calif.] : ESRI, c2008-2009.
Crash data: Alameda County SWITRS Dataset, aggregated by intersection

Of these 2114 crashes, 1715 (81%) were bicycle-motor vehicle collisions. 332 involved only one party—the cyclist—and should actually be called "falls" rather than "collisions." 51 involved a bicycle and a pedestrian but no motor vehicle. 16 involved two cyclists and no motor vehicle. As described in Methodology above, I formulate my hypothesis based on motor vehicle traffic characteristics on different streets, so my hypothesis is that bicycle-motor vehicle collisions should be less frequent on side streets than arterials. For this reason I conduct my tests on the 1715 bicycle-motor vehicle collisions. However, since the other collision types are available in this dataset and represent part of the full profile of risk to cyclists (and, in the case of bicycle-pedestrian collisions, risk that cyclists pose to others), I also discuss them separately, and re-run my analysis with all 2114 collisions included, in Appendix 3.

The 1715 bicycle-motor vehicle collisions represent incidents from as early as 1996 and from all streets in Berkeley, including those that are neither bicycle boulevards nor arterials, so not all collisions feature prominently in my study.

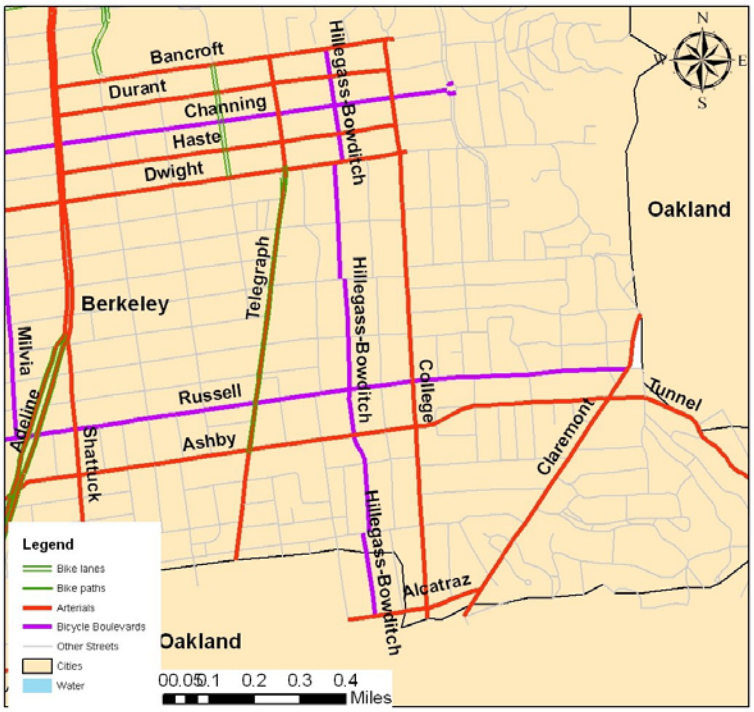
Of these 1715 collisions, just five involved two cyclists, and none involved more than two. One could raise the issue of whether collisions, or affected cyclists, ought to be the variable of interest in assessing risk. If the counts being used as a denominator are of individual cyclists, then one could argue that cyclists involved, not collisions, is the proper numerator for risk. On the other hand, many cyclists were transporting children, who might or might not be counted as separate "cyclists", whether in the police report of a collision or in the counts. The issue is complicated, yet unimportant. Unless all of the five two-cyclist collisions occurred on the same street, which they did not, it would make no difference in the results. For simplicity's sake, I treat collisions, not cyclists, as the variable of interest.

As a general remark about the SWITRS collision data, it is worth reiterating a point made in the literature review: bicycle collisions are chronically underreported to police, so these data represent just a fraction of actual collisions. In the case of the Alameda County SWITRS data, I can offer first hand evidence of the extent of underreporting. My own sister was involved in a hit-and-run collision in Oakland in 2007 in which she, a cyclist proceeding straight through an intersection, was struck by an oncoming left-turning motor vehicle. She called the police to report the incident but they never arrived, and no accident report was filed. Therefore, even though she suffered a broken clavicle, was briefly unconscious and had to be transported to the hospital in an ambulance, this collision does not appear in the SWITRS dataset. This collision would surely have qualified as "severe," so its absence shows that it would not be safe to assume that all severe collisions are reported. Moreover, if even a collision as serious as this can go unreported, just imagine the hundreds of more minor collisions in Berkeley which go unreported each year.

As I describe in Methodology above, I take great care to select *comparable* street segments so that my comparison represents a real choice for cyclists and cities. This comparability means paired street segments should meet two criteria: they should be of identical length, starting and ending at the same cross streets.

There are two pairings where these two criteria could not be simultaneously met. One involves the Hillegass-Bowditch bicycle boulevard, an alternative to Telegraph Ave. and College Ave.—see Figure 26. For much of its length, Telegraph runs at an angle, not parallel, to Hillegass-Bowditch, so I had the choice of either selecting the same *length* of street or selecting the same *segment*, but not both. I chose to select the same segment, again reasoning that this represents the actual choice that cyclists face—if riding on Telegraph involves more mileage, then that affects the risk associated with that choice. Of course, depending on one's destination, Telegraph could be a trip-shortening hypotenuse whereas Hillegass-Bowditch would require a brief trip on a perpendicular street of unknown risk. In any case, it is not important—in fact, both street lengths round off to 1.1 miles. Another complication is that Telegraph hits the jagged Berkeley-Oakland border at Woolsey St. whereas College Ave., the other arterial adjacent to Hillegass-Bowditch, hits Oakland at Alcatraz Ave., one block further along. As luck would have it, there were no collisions on Hillegass-Bowditch in the one block between Woolsey St. and Alcatraz Ave., so the entire length could be compared to Telegraph or to College without any differential treatment.

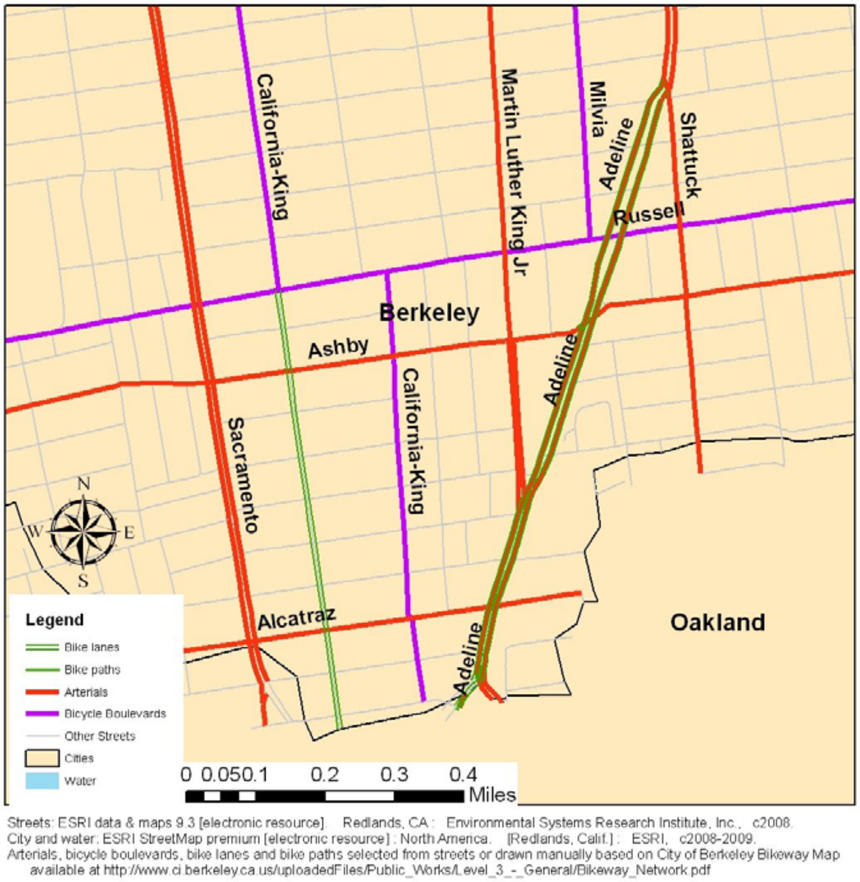
Figure 26. The alignment of Hillegass-Bowditch, Telegraph and College



Streets: ESRI data & maps 9.3 [electronic resource]. Redlands, CA : Environmental Systems Research Institute, Inc., c2008.
City and water: ESRI StreetMap premium [electronic resource] : North America. [Redlands, Calif.] : ESRI, c2008-2009.
Arterials, bicycle boulevards, bike lanes and bike paths selected from streets or drawn manually based on City of Berkeley Bikeway Map available at http://www.ci.berkeley.ca.us/uploadedFiles/Public_Works/Level_3_-_General/Bikeway_Network.pdf

The other difficult pairing involved the California-King bicycle boulevard, which consists mostly of California St. but makes a jog along Russell St. to reach its southernmost segment along King St. I chose to treat California St. and King St. as separate bicycle boulevards, comparing equal segments of California St. to Sacramento St., and separately comparing King St. to equivalent segments of Sacramento St. and MLK / Adeline St. In the end, however, no count data were available for King St., so it was eliminated from the study.

Figure 27. The alignment of the California-King bicycle boulevard, Sacramento St., and other nearby streets



Assigning collisions to streets of interest posed an additional challenge. Recall from my Methodology section that I include in each street-by-street comparison only those collisions which happened to cyclists traveling *on* the street of interest. For instance, in counting collisions on Milvia St., which runs north-south, I would not include incidents where a cyclist travling east-west on a cross street was struck by a motor vehicle proceeding north on Milvia St.

Collisions in the SWITRS database are located only by Primary Rd, Secondary Rd, and a Jurisdiction number which reveals the city; geocoded points are snapped to the intersection of the two roads, even for mid-block collisions. The metadata for the database gives no indication of how Primary Rd or Secondary Rd are defined, and it seems likely to vary according to the police officer recording the incident. One can imagine Primary Rd reserved for the road that one or more parties were traveling on prior to the collision, or for the more major of the two roads. Therefore, the question of which collisions involved cyclists *on* the street of interest is nontrivial.

To manage this issue, I used several variables from the database. I did not consider which road was listed as Primary Rd or Secondary Rd. For each street of interest, I selected the appropriate segment in ArcMap, applied a buffer large enough to capture all of the collisions along it, and labeled those collisions as belonging to that street. For instance, collisions along Milvia St had “Milvia” in the NS_St column, and collisions along Virginia St had “Virginia” in the EW_St column.

At this stage, I had only selected for geocoded location based on intersection; my selections did not indicate whether the collision had actually *occurred* on a street of interest. For instance, the collisions labeled “Milvia” would surely include some which were at midblock locations on cross-streets and thus had nothing to do with Milvia St. To select collisions relevant to the north-south street of interest, I developed the following two sets of criteria:

1. Generous: All intersection collisions, all collisions with multiple cyclists (since more than one direction might be involved), and all collisions where cyclist direction of travel was north or south. This only excludes midblock collisions on the east-west street. This should certainly include all relevant collisions, though it also includes quite a lot of irrelevant ones.

2. Conservative: Collisions with cyclist’s direction of travel as north or south, OR at an intersection and cyclist’s movement preceding collision as right, left or U-turn. This includes all collisions where the cyclist was traveling north or south along the street of interest as well as those where the cyclist turned onto the street of interest at an intersection. This seems to reliably exclude any irrelevant collisions, though it may also exclude a handful of relevant ones—for instance, if the cyclist was intending to turn onto the North-South street but had not yet begun turning and so was recorded as “Proceeding Straight.”

SQL code used to implement these criteria is included in Appendix 1.

Of these two criteria, I believe the conservative one is a more accurate reflection of collisions along a given street, and so my discussion of relative collision rates later on focuses on the conservative measure. I include the numbers obtained by both criteria in Appendix 3, however, to show that it makes little difference in the results.

It is worth noting that, even according to the conservative criteria, a collision can be assigned to more than one street. This makes sense: if a cyclist turns from one street onto another and is hit by a car at that intersection, that is a strike to the safety record of both streets.

Applying the above criteria I obtain measures of how many bicycle-motor vehicle collisions occurred on each street of interest. The final challenge is in selecting the appropriate time period over which to compare collisions. As I explain above, my hypothesis for Berkeley is that bicycle boulevards are safer than the arterials they parallel, so it would be ideal to choose a time period after the bicycle boulevards had already received most of the treatments depicted in Figure 13 through Figure 21.

According to City of Berkeley’s Bicycle Boulevard History web page, several steps were undertaken to implement bicycle boulevards during 2001-2002, with the “first phase of bicycle boulevard implementation” ending when signage and pavement markings were completed in “Winter

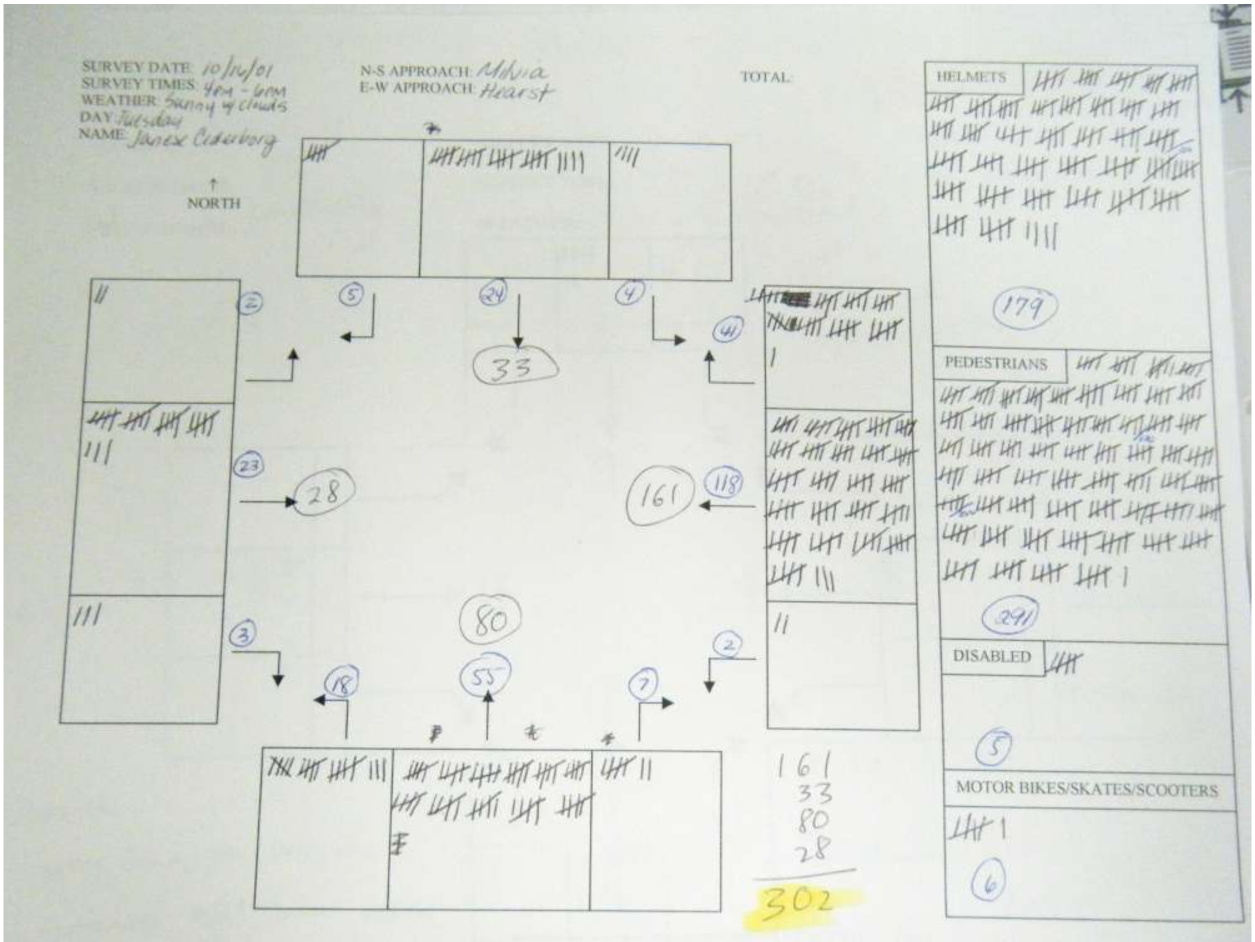
2002/3". I could not find any document with much more detail on the subsequent phases. The Bicycle Boulevards: Current Projects page lists at least one project (the Hillegass/Ashby intersection improvement) which I believe has not been completed to date (City of Berkeley, Bicycle Boulevards: Current Projects). The bicycle boulevards, it seems, are a decade-long work in progress—with some pre-existing traffic calming even before they became bicycle boulevards (City of Berkeley, What is a Bicycle Boulevard?).

This timeline offers no single point in time where one can declare that the bicycle boulevards were complete. This presents a tradeoff: setting a later start date from which to include data means less data are available to study, but an earlier start date means capturing more time when the bicycle boulevards were still relatively untreated side streets. I choose to begin at the start of 2003, around the time when the "first phase" was completed. I include all collisions from then until October 2008, the last month for which data is available.

4.1.3. Exposure

Since 2000, the City of Berkeley has conducted sporadic counts of cyclists. These counts are collected manually, by staff or volunteers standing on street corners making tally marks on pages. One of the more legible such pages is shown in Figure 28. All counts are conducted at intersections save those at the ped-bike bridge over I-80, and most of the intersections chosen are major ones for cyclists, where the two intersecting roads are bicycle boulevards, arterials, or one of each. The city was quite ambitious in its 2000 counts, conducting 47 separate counts at 36 different intersections over the summer and fall, mostly at weekday PM peak (4:00p – 6:00p). In subsequent years, the counts became gradually less thorough, and in 2007-8 none were conducted.

Figure 28. A sample cyclist count sheet from a count conducted by the City of Berkeley



Depending on the year, volunteers were often asked to record helmet use, cyclist sex and/or other factors alongside direction of travel or turn. However, when the city entered its data into Excel, only total count was consistently preserved; all data about direction of travel was thrown out. While visiting Berkeley in March 2010, I was able to obtain copies of most of the original count sheets with direction of travel indicated. Several were illegible—for instance, volunteers had not indicated which street was which on their diagrams of the intersection, or had made hundreds of individual marks rather than grouping their tallies into fives as is customary. I also did not bother to digitize counts at the ped-bike bridge, as my study is of city streets. Therefore the data in Table 4 reflects the usable counts that I was able to obtain after examining the hard copies.

As of this writing, the city's most recent counts were in 2009. To update counts and examine cycling conditions for myself, I and another volunteer conducted nine additional counts in Berkeley in March 2010. The final row in Table 4 reflects these counts.

Table 4. Number of usable cyclist counts in Berkeley

Year	Counts	Intersections
2000	47	36
2001	10	10
2002	12	11
2003	24	18
2004	10	9
2005	7	7
2006	0	0
2007	0	0
2008	0	0
2009	11	11
2010	9	9

It is unclear whether these counts are best described as "cyclist counts" or "bicycle counts." In my own counts I did not include people walking bikes, and I made only one tally mark, not two, when an adult cyclist passed with a child in tow. While a handful of the city's volunteers actually indicated how they were handling such complications, most did not.

To obtain flow for each street based on my own counts and the city's counts, I use the method described in Methodology, of counting straight-through cyclists as a full cyclist and turning cyclists as a half for each street. However, for 2000, most of the original count sheets were not available. Instead, someone had already typed up total north-south and east-west counts, assigning flow based on direction of *approach* to the intersection—thus missing some of the nuance of my method.

In creating an exposure measure from counts, I use only counts taken in 2003 or later—after my start point for including collision data. Though collision data is not available after October 2008, I use count data up to the present. Count data for relevant streets is in short supply, and from my experience living in Berkeley I believe that the design of the bicycle boulevards has been relatively static for the last few years, so counts taken today should still be fairly representative of cyclist traffic patterns over that time period. While it would be ideal to use only count data from the exact same time period as the collision data, it is also important to have count data for more streets of interest, and to have enough repeated counts to be able to average out—and examine—fluctuations in counted volumes. In Analysis and Results I address in detail the question of how well these count data can approximate true exposure.

Using many different counts to create a single exposure measure for each street requires some way of combining or averaging the count data. A given street—again, take Milvia as an example—may have several counts, conducted over several years and at a few different cross streets, with some repetition. I chose to average first over time, so that the traffic on Milvia at Channing is considered to be the average of the traffic observed at that point in 2003, 2004, 2005, 2009 and 2010. Next I average over space, so that the traffic on Milvia is considered to be the average of the traffic observed at Channing and at Hearst, the two points where Milvia has been counted. For consistency, I use only weekday PM peak (4:00p-6:00p) counts in all cases. Exposure measures obtained in this manner are shown in Table 5.

Table 5. Exposure measures obtained for each street in Berkeley

Street	Type	Average 2-hour PM peak volume	Number of counts	Number of cross streets
Hillegass-Bowditch	B	107.6	8	3
Telegraph	A	114.5	3	1
College	A		0	0
Milvia	B	188.1	10	2
Shattuck	A	185.0	1	1
MLK	A	44.8	5	3
California	B	39.8	9	3
Sacramento	A		0	0
9th	B	36.4	7	3
San Pablo	A	43.2	3	3
Virginia	B	61.3	4	1
University	A	28.9	6	2
Channing	B	87.2	13	6
Dwight	A		0	0
Russell	B	63.6	9	4
Ashby	A	29.6	6	4

Unfortunately, no count data were available for three streets—College, Sacramento and Dwight. Also, only one count was available for Shattuck, so I specifically address the reliability of Shattuck's exposure measure when I discuss count reliability in Analysis and Results.

4.2 Cambridge

4.2.1 Background

Cambridge is a city in the Boston Metropolitan Area, located in Middlesex County immediately across the Charles River from Boston proper. It is famous as home to Harvard University and the Massachusetts Institute of Technology. It has also boasts several bustling retail districts, biotech and high-tech industries and easy transit commute access to downtown Boston. It has been ranked America's Best Walking City (Americawalks.org 2008) and, of American cities with average commute times below the national average, it has the highest percentage of non-car commuters, 58% (Bandyk 2009).

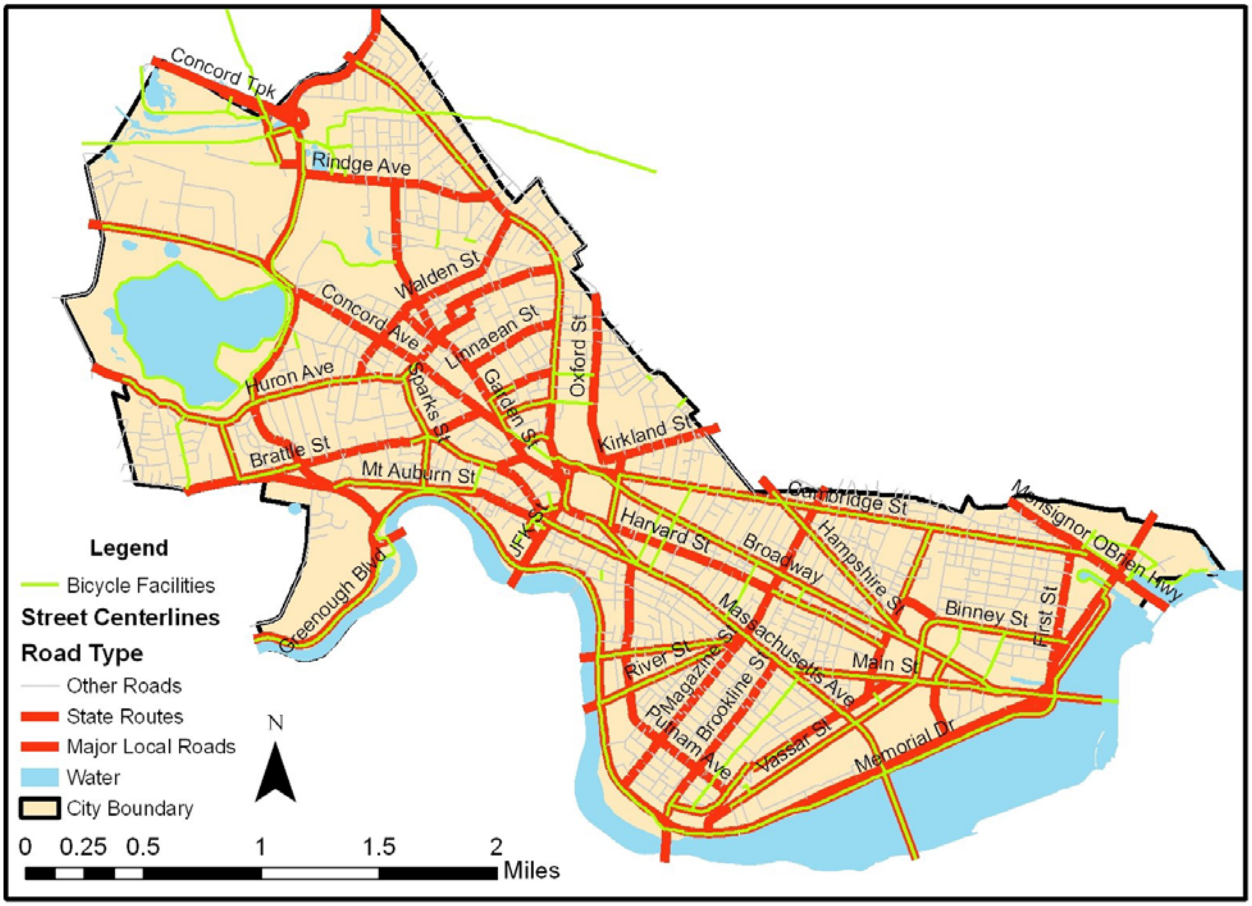
Figure 29. Cambridge, MA viewed from M.I.T. Campus in the foreground, Central Square and Harvard Square straight ahead.

Cambridge has a population of about 101,000 (United States Census 2000). Spread over 6.4 square miles of land, this represents a population density of about 16,000 people per square mile (about half again that of Berkeley), making it among the denser urban areas in the U.S. About 32% of Cambridge's 41,000 households do not own cars, compared to just 10% in Middlesex County and 9% nationwide. 3.9% of workers who lived in Cambridge in 2000 biked to work, but this number will probably change substantially in the 2010 Census, as the City of Cambridge's bicycle counts indicate that the number of cyclists on the city's roads approximately doubled between 2003 and 2008 alone—see Exposure below.

While the bike path along the Charles River dates back further, most of the city's bicycle infrastructure exists thanks to a Bicycle Program created in 1992 by the city's Vehicle Trip Reduction Ordinance (Seiderman, personal email, 2010). No document exists which spells out the history of all the city's bicycle facilities, but it seems that the program really took off in the late 1990s and gained steam in the 2000s.

Cambridge's strategy for accommodating cyclists is quite the opposite of that taken in Berkeley. The City of Cambridge has added or plans to add bike lanes on most of the city's major arterials, while making relatively few bicycle-specific accommodations on side streets. Figure 30 overlays bicycle facilities built as of 2009 on the major roads. Note that the green lines in Figure 30 represent anything that the city considers to be a bicycle facility, including "edge lines"—stripes of paint which demarcate a space between traffic and parked vehicles but are not very wide nor specifically marked for bicycle use. The pictures following Figure 30 provide a survey of bicycle facilities in Cambridge.

Figure 30. Bicycle facilities and major roads in Cambridge, MA



Streets: City of Cambridge GIS, Street Centerlines 2009
Bicycle Facilities: City of Cambridge GIS, Bike Facilities 2009
City Boundary: City of Cambridge GIS, City Boundary 2004
Water: City of Cambridge GIS, Hydrography 2004

Figure 31. Bike lane. Massachusetts Avenue, Central Square

Figure 32. Cyclists on edge lines.
Massachusetts Avenue north of Central Square

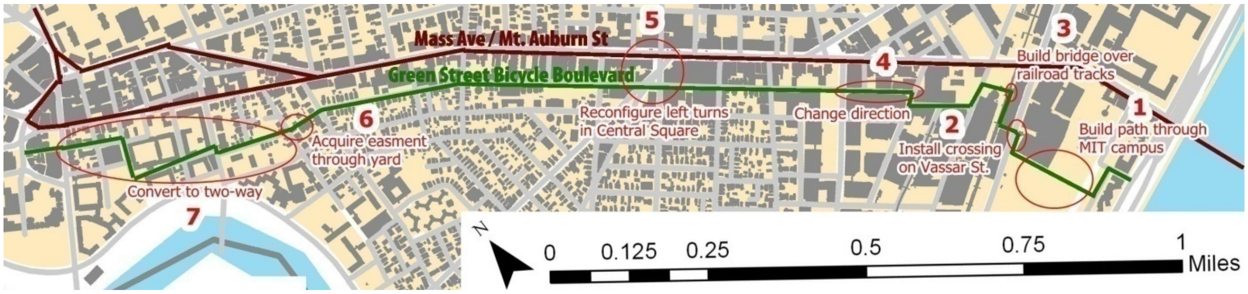
Figure 33. Cycle tracks. Vassar St. on MIT campus.

Figure 34. Dr. Paul Dudley White Bicycle Path along the Charles River

This difference in strategy probably stems from Cambridge's highly irregular street network. Arterials run at odd angles to one another and are the only really connective routes in the city. Side streets are short, unconnective and often one-way.   To make a high-quality bicycle boulevard network in Cambridge is narrowly imaginable, but not realistic.

As a thought experiment, consider what it would take to link all the way from the Charles River to Harvard Square on a bicycle boulevard alternative to the Massachusetts Avenue (hereafter "Mass Ave," as it is known locally) / Mt. Auburn Street corridor. Figure 35 illustrates just a few of the adjustments this would require. M.I.T. would have to agree to cut a route or path through its west campus (1), perhaps between the athletic fields and the gym, create a crossing at Vassar St. (2) cut a route through the parking lots and industrial buildings lining the railroad tracks and build a bridge over the railroad tracks (3) to link with Cross St. Next the city would have to reverse the one block of Green St. (from Landsdowne St. to Sidney St.) that runs southbound, so the whole street ran northbound (4), and do the converse with parallel Franklin St. to provide continuous southbound travel, or else make Green St. two-way, which would require removing at least one side of curb parking. The MBTA #70 bus route, which runs on Green St. because it cannot turn left onto Western Ave. from Mass Ave., would have to be relocated, and the traffic which turns left on Green St. from River St. because left turns aren't allowed at Mass Ave. would likewise have to be diverted elsewhere (5). Extensive traffic calming and diversion measures would be needed to prevent cut-through traffic. At the dead end of Green St., the city would have to purchase an easement through the yard of the house on the other side of the dead end, knock down the fence and put a path connector through the yard (6). Next Harvard University would have to agree to make all of Grant St., Mill St., Holyoke St. and Winthrop St. two-way through its Riverside campus extension (7), allowing cyclists to link all the way to the non-motorized block of Winthrop St. at the heart of Harvard Square.

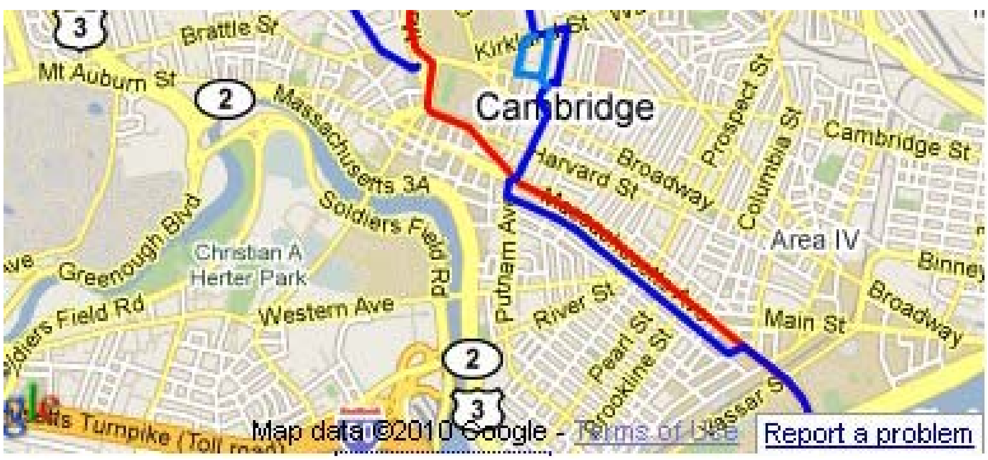
Figure 35. A few of the challenges in implementing a true bicycle boulevard alternative in the Mass Ave / Mt. Auburn corridor



It would be a very pleasant, and fairly direct, route. If planning and community meetings began today, the project might be completed by 2150.

Considering all this, Cambridge's decision to accommodate cyclist desire lines along its arterials seems inevitable. Still, in Cambridge as elsewhere, some cyclists have sought out lower-traffic routes to ride on. Cambridgebikes.org, a group of the City of Cambridge Bicycle Committee recently proposed a plan of side street bikeways as a supplement to the city's arterial bike network. Unlike the scenario shown in Figure 35, the group's proposed Green St. alternative to Mass Ave., shown in detail in Figure 36, does not reach the M.I.T. campus or Charles River to its south, simply terminating at Landsdowne St., and does not reach Harvard Square, instead weaving east of Harvard University's main campus.

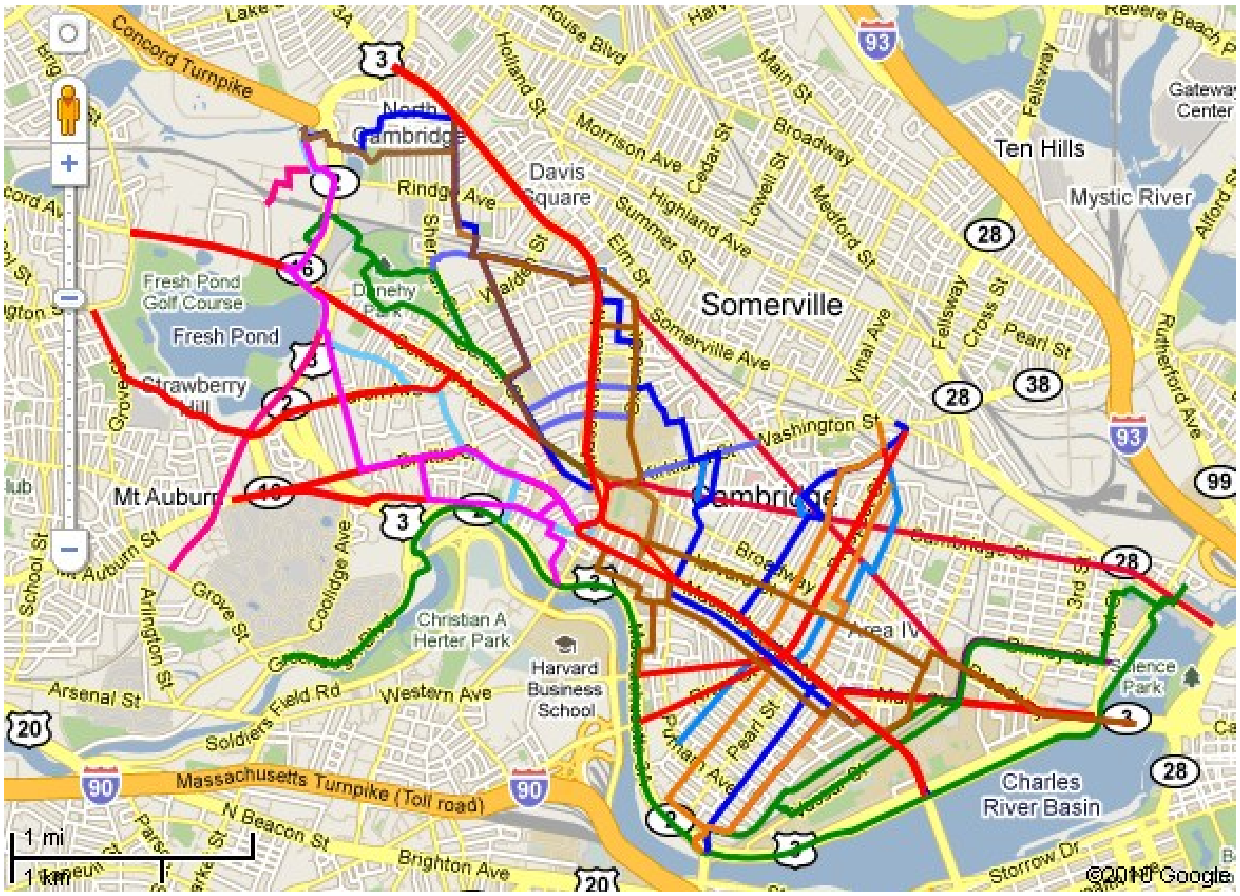
Figure 36. Cambridgebikes.org proposed alternative (blue) to Mass Ave. (red)



Source: cambridgebikes.org/Network

The full Cambridgebikes.org proposal, depicted in Figure 37, shows that a side street network is possible in Cambridge, yet also shows how inferior such a network would be to that in Berkeley. First, the side streets do not quite form a complete network, and there are several segments where cyclists would still have to travel on arterials. Another drawback is that, whereas Berkeley's bicycle boulevards run parallel to their arterials and are straight and direct, many of the routes shown in Figure 37 require cyclists to travel far out of their way. Finally, some of the "side streets" used, while quieter than arterials, are still fairly busy. For instance, Brookline St. has a daily motor vehicle count of about 13,000, nearly as much as the 16,000 or 19,000 measured at two points on Western Ave. (City of Cambridge, 24 Hour Average Daily Traffic Counts) . Recall that in Berkeley, even the downtown section of the Milvia bicycle boulevard carries just a fifth the motor vehicle traffic of its neighboring arterials.

Figure 37. Cambridge Bicycle Committee recommended bicycle network. Direct routes are in red, side street alternatives in other colors.



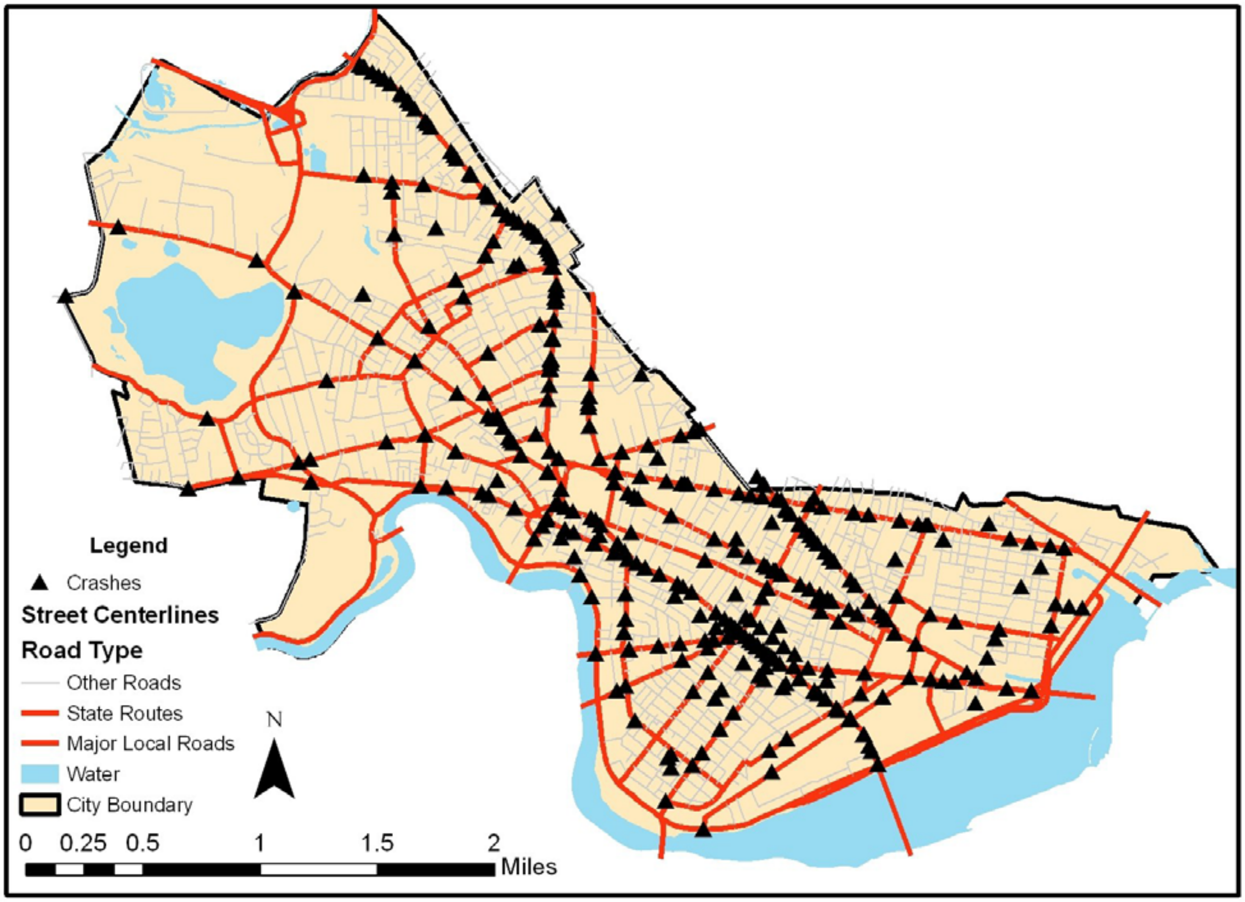
Source: cambridgebikes.org/Network

4.2.2 Collisions

Cambridge's Traffic, Parking and Transportation Department, with the help of the Cambridge Police Department, maintains an Excel file summarizing police-reported bike and moped crashes on city streets. As of this writing, the database contains data from January 2004 through September 2009, inclusive, for a total of 524 crashes. Included are nine bicycle-pedestrian collisions, three moped-motor vehicle collisions, three cyclist falls, and a handful of miscellaneous incidents somehow involving a cyclist. About 500 are bicycle-motor vehicle collisions.

Figure 38 shows the locations of the bicycle crashes. Note the concentration along arterials. Mass Ave. alone accounts for 183 of the dataset's 524 crashes—more than one third.

Figure 38. Geographic distribution of police-reported bicycle crashes 2004-9 in Cambridge, MA



Streets: City of Cambridge GIS, Street Centerlines 2009
City Boundary: City of Cambridge GIS, City Boundary 2004
Water: City of Cambridge GIS, Hydrography 2004
Crashes: City of Cambridge Bicycle Crash Data 2004-9

The database's major limitation is a lack of data on collision details, at least in a format easy to operationalize. 480 of the records contain at least some text from the police narrative describing the incident, in some cases with quite a bit of detail. However there are no separate columns for direction of travel or street of travel which would make it possible to use SQL to quickly extract all the collisions related to a particular street. Severity levels for the "injury status (box 32)" column are shown in Table 6, though the contact at the Cambridge Police Department who provided me with this information acknowledged that the assignment of severity level is subjective, at the reporting officer's discretion.

Table 6. Severity levels (injury status box 32) in the Cambridge collision dataset

1	fatal
2	incapacitating injury
3	non-incapacitating
4	possible injury
5	no injury
99	unknown

Source: Murphy, personal interview, 2010.

Another drawback Cambridge's crash dataset is that some of the data are questionable or missing. Two records contain '1' in the "injury status (box 32)" column, indicating fatalities, even though no bicycle fatality has occurred in Cambridge since 2002. A number of entries contain no textual information on location and can be located only by their geocoded XY point.

An interesting point in the dataset is that it includes a column for the type of collision, such as "door," "left hook," and so on. Though I do not make extensive use of it here, this variable could allow for rich analysis. For instance, it is notable that 105 of the 524 crashes (20%) are doorings.

The fact that the crash dataset does not contain columns that make it possible to operationalize the sorting of collisions by street of interest means that I must examine each collision's description separately and decide to which street it belongs. Due to this limitation, I do not conduct a city-wide analysis of Cambridge's arterials and side streets. Instead, I just make two comparisons to illustrate that my methodology can be applied in Cambridge, and to see if any surprising results emerge.

The first comparison is of Green St. and Mass Ave. Green St. is a one-way, one-lane street, mostly residential in character. It certainly has lower motor vehicle traffic volume than Mass Ave., though unfortunately no official counts are available. However, Green St. does have some heavy vehicle traffic—the MBTA #70 bus route, postal service trucks entering the loading docks of the Cambridge Post Office, and an assortment of utility trucks. It also sees a sizeable volume of cut-through traffic, in part because motorists turning left from River St. are required to use Green St. instead of Mass Ave. From my own observations, operating speeds may actually be higher on certain parts of Green St. than on corresponding parts of Mass Ave. To the extent that this is true, it limits my ability to hypothesize a lower severity of collisions on Green St.

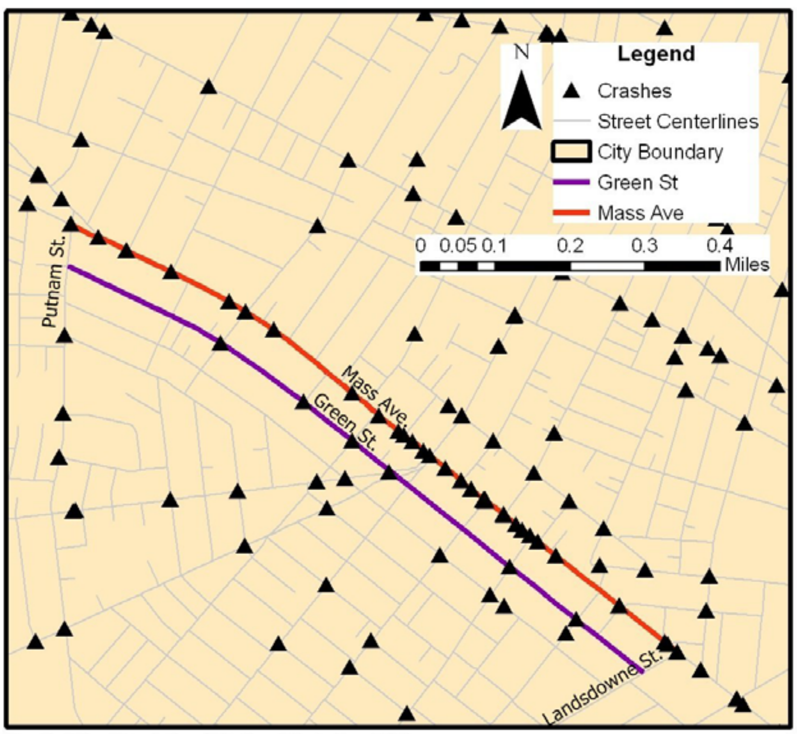
The equivalent segment of Mass Ave. has one travel lane and one parking lane in each direction for most of its length, with two travel lanes for the southernmost few blocks. It currently has bike lanes for nearly its entire length—from Landsdowne St. all the way to Cambridge City Hall between Pleasant St. and Sellers St., where the bike lane becomes an "edge line." It is worth noting, though, that the period for which collision data are available, January 2004 – September 2009, represents a few different states of Mass Ave. The section through Central Square has had bike lanes since 1997 (Seiderman, personal interview, 2010), but the southernmost section only had its bike lanes completed a few years ago, and was under construction for much of the time period of study, which might have increased risk for cyclists.

Figure 39. Green St.

Figure 40. Mass Ave. parallel to Green St.

I selected in GIS the identical 0.9 mile segments of Green St. and Mass Ave.—from Landsdowne St. to Putnam Ave. in each case. I then buffered the street segments to select by location the geocoded collisions that occurred on the street.

Figure 41. The 0.9 mile segments of Mass Ave and Green St that I compare



Streets: City of Cambridge GIS, Street Centerlines 2009
City Boundary: City of Cambridge GIS, City Boundary 2004
Crashes: City of Cambridge Bicycle Crash Data 2004-9

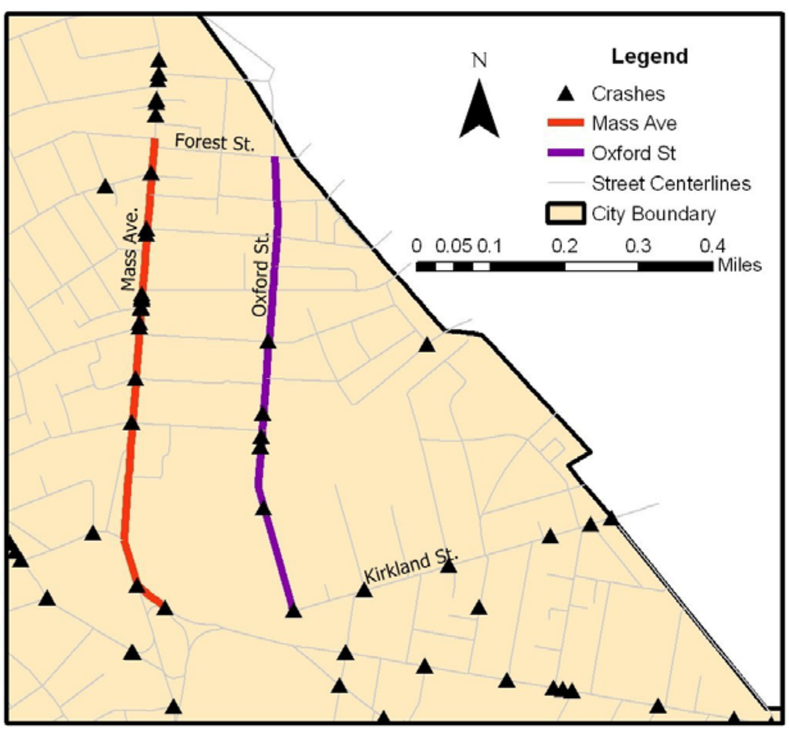
The second comparison is of Oxford St. and Mass Ave. between Harvard University and Porter Square. Oxford St. is a quiet two-way street with one travel lane and one parking lane in each direction, traffic lights and a few speed bumps, traversing residential neighborhoods and the back side of Harvard University's campus. Daily motor vehicle traffic counts along Oxford St. are around 7,000 (City of Cambridge, 24 Hour Average Daily Traffic Counts). The equivalent segment of Mass Ave. has two travel lanes and one parking lane in each direction, an unadorned concrete median, and sporadic sharrows for cyclists. The nearest motor vehicle count on Mass Ave., near Porter Square to the north, is about 54,000 (City of Cambridge, 24 Hour Average Daily Traffic Counts).

Figure 42. Oxford St.

Figure 43. Mass Ave. parallel to Oxford St.

Just as before, I selected the same segment of each street—0.6 miles in this case, from Forest St. in the north to Kirkland St. in the south, though the end point for Mass Ave is approximated since it does not intersect Kirkland St. These street segments are shown in Figure 44. It would have been ideal to extend the northern edge up to Roseland St., but half of the last block of Oxford St. is in Somerville and it is not clear that the Cambridge Police Department would have any record of cyclist collisions there.

Figure 44.The 0.6 mile segments of Mass Ave and Oxford St that I compare



Streets: City of Cambridge GIS, Street Centerlines 2009
City Boundary: City of Cambridge GIS, City Boundary 2004
Crashes: City of Cambridge Bicycle Crash Data 2004-9

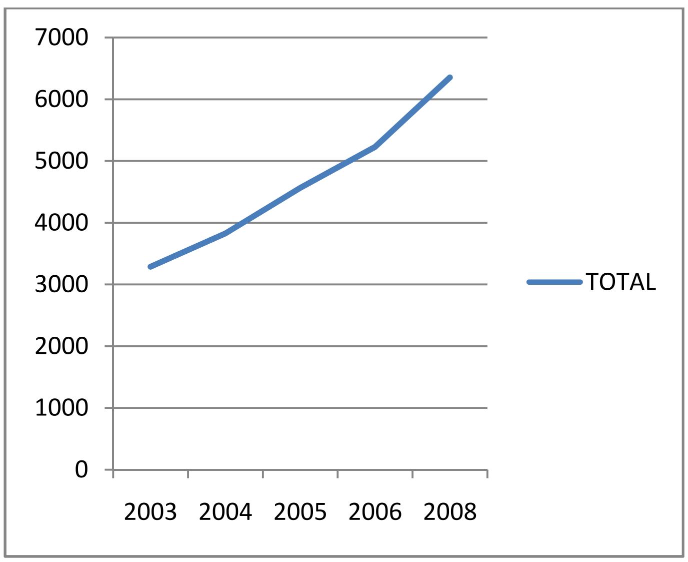
For both street pairings, after buffering the street segments to obtain lists of collisions that occurred on each street, I read through each collision's description to determine if it belonged in the study. I threw out collisions where the cyclist was traveling on a cross street rather than the street of interest. I also removed those where no motor vehicle was involved, though I analyze these separately and all crash types together in Appendix 5.

4.2.3 Exposure

The City of Cambridge has conducted highly organized cyclist counts since 2003 and has made detailed data available electronically. The city counted the same 16 intersections in September of 2003, 2004, 2005, 2006 and 2008, for one hour each in AM and PM peak with direction of travel information included, allowing for counts to be summed by street.

The counts reveal an impressive increase in cycling over the five years from 2003 to 2008. The total number of cyclists counted at all intersections has approximately doubled, as shown in Figure 45. In fact, every intersection has seen at least a 30% increase in bike traffic over the five years, and some intersections have tripled their counts.

Figure 45. Total cyclists counted at 16 intersections in Cambridge in one hour each of AM and PM peak, 2003-8



The major limitation of these counts, with regards to this research, is that the streets counted are mostly through streets which carry a substantial volume of automobile traffic. A couple of side streets with bicycle facilities which act as connectors, such as Sparks St., are included as cross streets, but many of the quieter side streets that Cambridgebikes.org has proposed as alternate routes have never been counted.

For this reason I was unable to use the city's count data to compute collision rates. Instead, I and another volunteer conducted simultaneous counts of Mass Ave. and the two side street alternatives.

Since Green St. is one-way, it is probable that cyclist traffic there is substantially different in AM and PM peak. Cyclists who prefer a side street alternative to Mass Ave. might use Green St. northbound and the neighboring Franklin St. southbound. Mass Ave. and Green St. are only about 250 feet apart (centerline to hypothetical centerline) so it is possible, standing at just the right point, to count them at the same time. From 8:00a to 9:00a and 4:30p to 5:30p on two consecutive days, I counted both streets. The first day, I counted at Pleasant St., in the heart of Central Square, where I would anticipate highest traffic on both streets. The second day, I counted at Sidney St., essentially the beginning of Green St., where I would anticipate lowest traffic, at least for Green St.

To obtain count data for Oxford St. and Mass Ave., I and another volunteer stood at the intersection of each street with Wendell St., around the midpoint of the street segments of study, again from 8:00-9:00a and 4:30-5:30p.

I computed an average two-hour exposure measure by adding the AM and PM counts and, in the case of Green St. and Mass Ave., averaging over the two cross streets where I had counted. Results are shown in Table 7.

Table 7. Average exposure measures computed for streets in Cambridge

Street	Exposure
Green St.	24.3
Mass Ave. parallel to Green St.	311.5
Oxford St.	159.0
Mass Ave. parallel to Oxford St.	269.5

5 Analysis and results

5.1 Berkeley

Table 8 compares the collision rates on all the streets of study in Berkeley. Bicycle boulevards are shown in purple, followed by the arterial(s) they parallel, shown in red. The collisions column indicates the total number of bicycle-motor vehicle collisions, by conservative inclusion criteria, that occurred on the street between January 2003 and October 2008, inclusive. This figure is available for all streets. The exposure column indicates the average two-hour PM peak count—averaged first over time and then over space, as I describe in Methodology. Some streets had no count data available. For those with count data, the relative collision rate listed in Table 8 is the quotient of the preceding two columns. Note that the units of the relative collision rate column are not meaningful—collisions in about six years, divided by cyclists in two hours. Also, each bicycle boulevard is comparable only to the arterial(s) immediately following it in the table; streets from different groups are not comparable because the length of street segments may be different.

Table 8. Collisions, exposure and relative collision rate for Berkeley streets based on January 2003 – October 2008 bicycle-motor vehicle collision data

Street Name	Street Type	COLLISIONS	EXPOSURE	RELATIVE COLLISION RATE
Hillegass-Bowditch	B	5	107.6	**0.05**
Telegraph	A	21	114.5	**0.18**
College	A	30		
Milvia	B	26	188.1	**0.14**
Shattuck	A	47	185	**0.25**
MLK	A	18	44.8	**0.40**
California	B	9	39.8	**0.23**
Sacramento	A	15		
Ninth	B	3	36.4	**0.08**
San Pablo	A	26	43.2	**0.60**
Virginia	B	6	61.3	**0.10**
University	A	17	28.9	**0.59**
Channing	B	19	87.2	**0.22**
Dwight	A	11		
Russell	B	7	63.6	**0.11**
Ashby	A	26	29.6	**0.88**

For each and every pair of bicycle boulevard and arterial where count data are available, the bicycle boulevard has the lower collision rate. The risk ratio—collision rate on the arterial divided by collision rate on the bicycle boulevard—ranges from 1.8 for Shattuck v. Milvia to 8.0 for Ashby v. Russell. Appendix 3 shows that including all crashes rather than only bicycle-motor vehicle collisions, or including bicycle-motor vehicle collisions by generous rather than conservative criteria, give very similar results.

To analyze the proportion of collisions that are severe, I use all collisions since 2003 across Berkeley's entire street network. Collisions which occurred on one arterial, on an arterial and a non-bicycle boulevard cross street, or at the intersection of two arterials are categorized as "A". Similarly, collisions on a bicycle boulevard, a bicycle boulevard and a non-arterial cross street, or two bicycle boulevards are "B". Those at the intersection of an arterial and a bicycle boulevard are "AB", and those that involved neither are "O." Only comparable segments of arterials and bicycle boulevards are included in this analysis, just as for collision rates, so a collision which occurred on an arterial for which no bicycle boulevard alternative existed would be classified as "O." The results in Table 9 show that the bicycle boulevards actually have a higher proportion of collisions that are severe than do arterials, though with a p-value of about .6 for a two-tailed difference of proportions test, the difference is far from being statistically significant.

Table 9. Proportion of bicycle-motor vehicle collisions that are severe on different street types in Berkeley, January 2003-October 2008.

Street type	Number of Collisions	Number Severe	Proportion Severe
A	575	21	3.7%
B	185	9	4.9%
AB	3	0	0.0%
O	952	41	4.3%

Results in Table 8 and Table 9 indicate that bicycle boulevards have lower collision rates than their corresponding arterials, with about the same proportion of those collisions being severe.

To test the statistical significance of the findings in Table 8, I first conduct a Poisson regression on each street pair. Details of this method are provided in Appendix 2. Table 10 summarizes the results.

Table 10. Risk ratio and statistical significance of Poisson regressions on street pairs in Berkeley

Arterial	Bicycle Boulevard	Risk ratio	p value
Telegraph	Hillegass-Bowditch	3.9	0.0058
Shattuck	Milvia	1.8	0.0128
MLK	Milvia	2.9	0.0005
San Pablo	Ninth	7.3	0.0011
University	Virginia	6.0	0.0002
Ashby	Russell	8.0	< 0.0001

Recall from Methodology that the Poisson regression assumes that collision and count data are representative of total collisions and total exposure. A separate discussion of the reliability of these measures follows.

As I describe in Methodology, reported collisions could be an unreliable measure of total collisions if the streets being compared have different populations with different levels of inclination to report collisions. I have no data with which to formally address this concern. In my time observing streets in Berkeley, I did not notice a categorical difference in cyclists on the two street types. Bicycle boulevards and arterials alike had some fast riders in spandex and some slow casual riders; some commuters, some students, and so on. I did see more people carrying small children on the bicycle boulevards than on arterials.

A quick sensitivity analysis is worthwhile. Take Milvia vs. Shattuck, the bicycle boulevard-arterial pair with the smallest difference in estimated collision rate. Even for this pair, in order for my result of the bicycle boulevard being safer to be invalid, it would have to be that reporting rates on the arterial are higher, and by almost a factor of two. So if Milvia had a reporting rate of 40%, as one study has found for bicycle crashes (Moritz 1997), Shattuck would need a reporting rate of almost 80%. Given that the profile of cyclists on the two street types was not so obviously different, this is somewhat hard to imagine.

It is also possible that a difference in collision severity drives a difference in reporting rates, but as I explain in Methodology, this does not alter the product of collision rate and proportion severe, in other words the risk of severe injury per unit exposure.

As for counts, I pose three questions about how well they represent the ratio of volume between two streets.

1. How consistent is the ratio of counts on A to B at different points along the two streets?
2. How consistent is the ratio of counts on A to B over different years?
3. How well does the ratio of A's weekday PM peak traffic to B's represent the ratio over the whole week?

The City of Berkeley's counts provide some, albeit limited, ability to answer the first two questions. The third is more difficult to answer with available data, and will be left largely unresolved.

In 2003, there were two bicycle boulevards – Ninth and Russell—for which the city conducted counts of the boulevard and the neighboring arterial at multiple cross streets, making it possible to see how consistent the ratio of bike traffic was. For Ninth v. San Pablo, all three counts were simultaneous—4:00p to 6:00p on the exact same day.

Table 11. Ratio of bicycle boulevard to arterial bike volumes at different cross streets according to Berkeley cyclist counts.

Season	Cross Street	BBlvd Date	Arterial Date	BBlvd Vol	Arterial Vol	Ratio
Russell v. Ashby						
Spring	Mabel	2003/6/4	2003/6/4	27	26	**1.04**
Spring	MLK	2003/4/9	2003/5/28	75.5	46.5	**1.62**
Spring	Hillegass	2003/3/18	2003/3/19	50.5	22	**2.30**
Spring	California	2003/5/20	2003/6/11	48	26.5	**1.81**
Ninth v. San Pablo						
Season	**Cross Street**	**BBlvd Date**	**Arterial Date**	**BBlvd Vol**	**Arterial Vol**	**Ratio**
Spring	Parker	2003/5/27	2003/5/27	41.5	43.5	**0.95**
Spring	Channing	2003/6/3	2003/6/3	32.5	38	**0.86**
Spring	University	2003/5/29	2003/5/29	34.5	48	**0.72**

Russell has consistently a bit more bike traffic than Ashby, and Ninth has consistently a bit less bike traffic than San Pablo. The ratio varies a bit, but never to the point that it would call into doubt that there is a difference in collision rate. Since Ninth has only 3 collisions and San Pablo has 26, Ninth would have to drop to .11 of San Pablo's bike traffic in order for there to be no difference in collision rate, whereas the lowest ratio seen here is .72. Likewise, Russell has just 7 collisions to Ashby's 26, and so the ratio of bike traffic would need to drop to .27, whereas the lowest ratio seen here is 1.04. So while the ratios vary a bit depending on where they are measured, they do not vary nearly enough to call the difference in collision rate into doubt. It is regrettable that such comparisons are not possible for other streets in this study. For those streets with smaller risk ratios—say, Milvia v. Shattuck with 1.8, variation in count ratios at different cross streets *could* be enough to erode the difference.

The second question asked above is how consistent is the ratio of volume on parallel streets A to B over different years. Here, unfortunately, we have less to go on, because after 2003 the city stopped doing side-by-side counts such as those in Table 11. In my 2010 counts, however, I conducted some side-by-side counts. On consecutive weekdays I conducted one-hour PM peak counts at Ninth & Channing and at San Pablo & Channing, and found 43 cyclists on Ninth and 32.5 on San Pablo, for a ratio of 1.32, which is even more favorable to the Ninth's relative safety record than the city's 2003 numbers are. Variation is revealed, but in a direction that supports my test hypothesis—that bicycle boulevards have lower collision rates than arterials.

Another way of getting at this question is to compare perpendicular streets' counts over time, since the city counted many of the same intersections year after year. If it is true that overall cycling volumes exhibit yearly trends but that route choice is consistent, then an intersection should have about the same ratio of N-S traffic to E-W traffic in repeat counts over different years. Table 12 examines this proposition, showing the mean, maximum and minimum of the ratio of N-S to E-W traffic at six of the most counted intersections in Berkeley. Each row in Table 12 represents an intersection that was counted N times over 2003-2010.

Table 12. Range of ratios in NS to EW bicycle traffic observed at intersections over repeat counts

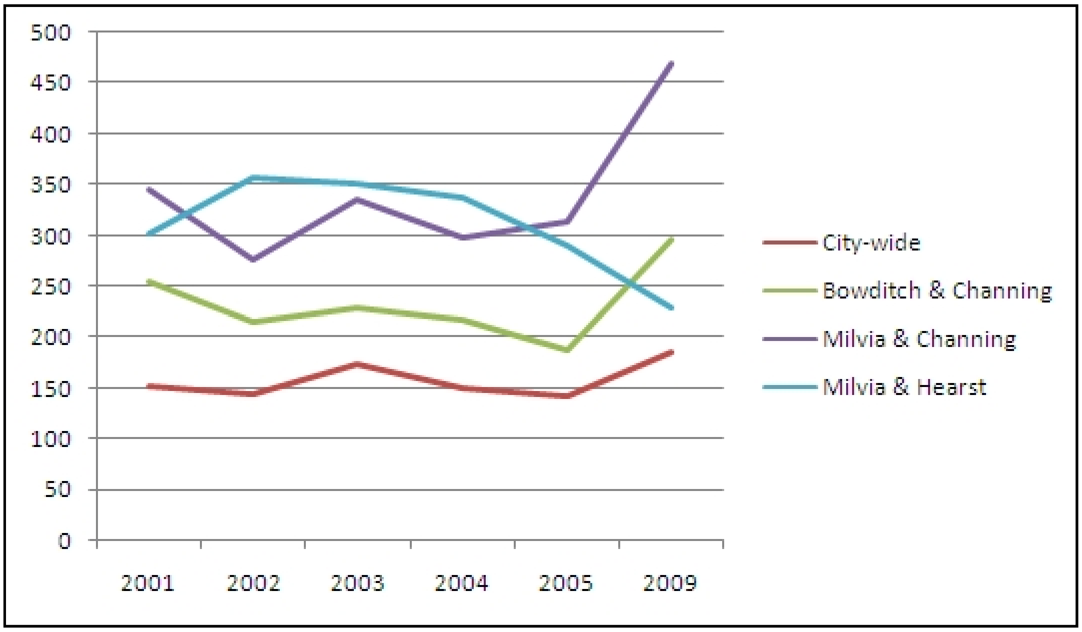
N-S Street	E-W Street	N	Ratio NS/EW		
			Mean	Max	Min
Ninth	University	5	2.6	6.4	0.6
California	Russell	4	1.1	1.2	1.0
California	Virginia	4	0.7	1.0	0.5
Hillegass-Bowditch	Channing	4	2.5	2.8	2.0
Milvia	Channing	5	1.9	2.3	1.6
Milvia	Hearst	5	1.0	1.2	0.8

Ninth & University has a rather large range, but in fact, this was entirely caused by one count. In four of the five counts, Ninth had multiple times the volume of University, and in just one count, it was almost exactly reversed. This raises suspicion that perhaps the volunteer conducting the count simply mislabeled the two streets.

Other than that, the other intersections show a fair amount of consistency: some have about equal traffic on both streets (Milvia & Hearst or California & Russell), while others have consistently more traffic on one street than the other. This supports the notion that route choice is consistent, and so that the ratio of cyclist volume on two *parallel* streets, A and B, is probably consistent.

Relatedly, one might ask if there is a discernible trend in cycling overall during the period of study, 2003-2008. Some arterials have seldom or never been counted since 2003, while bicycle boulevards have been counted more consistently. If bicycle traffic were on the rise during this time, then exposure measures based on these counts would be biased, showing more bicycle traffic on bicycle boulevards, relative to arterials, than is really the case. The city's own analysis of its count data provides a summary table of *total* (north-south and east-west lumped together) fall PM peak counts at its three most often counted intersections, as well as a "city-wide" average of several of the most counted intersections. Figure 46 graphs these counts.

Figure 46. Change in Berkeley cyclist counts at selected intersections, 2001-2009



Source: Berkeley 2000-2009 Bicycle Counts – Chart Update II – Trend Line Data

A linear regression on the counts in Figure 46 finds wildly varying slopes, with Milvia & Hearst even suggesting a decrease in cycling. None of the slopes are statistically significant at the .05 threshold. Note also that two of the intersections singled out (Milvia & Channing and Bowditch & Channing) are intersections of two bicycle boulevards: if an increase in cycling would be expected anywhere in the city, it would be expected here. It seems, then, that there is no trend in overall cycling or cycling on bicycle boulevards since 2001. The lack of a trend gives confidence that the year in which different streets were counted has not biased the relative collision rates computed.

Of course, the fact that cycling has not increased—especially given that Figure 46 begins in 2001, when the bicycle boulevards were just beginning to be implemented—is worrisome from the standpoint of advocating bicycle boulevards as a public policy instrument to promote cycling and promote the use of safer streets. As I demonstrate in this paper, bicycle boulevards have lower collision rates than arterials, so shifting cyclist traffic from arterials onto bicycle boulevards might improve safety even if more people did not begin cycling. Yet such a shift, unless overall cycling were actually on the decline, would come with an increase in cycling on bicycle boulevards, which is not apparent from Figure 46. In Appendix 4 I try a few different ways of examining whether cyclist traffic has shifted from arterials to bicycle boulevards, but data are not numerous enough to conclude anything. I also discuss this issue further under Policy Implications in the Conclusion.

A separate point to make about Figure 46 is that, whether due to yearly variation or daily variation, there is some variation in the number of cyclists counted at the same intersection in repeat counts. This might cause one to question the reliability of the relative collision rate that I computed for Shattuck Ave., since it was based on a single count as exposure data. In fact, that single count was one that I myself conducted at Shattuck and Channing in 2010, paired with a simultaneous count on Milvia and Channing. Those counts yielded 249 cyclists on Milvia and 185 on Shattuck. Since the count of Milvia was averaged with several other city-conducted counts, the average volume computed for Milvia dropped to 188. But this paired count actually suggests an even larger difference in collision rates for Milvia and Shattuck, 26/249 = .10 versus 47/185 = .25. Paired counts are more valuable because they inherently control for daily and yearly variation.

The city also conducted paired counts of Milvia and Shattuck in 2000 and non-paired counts in 2001 and 2002. These counts, summarized in Table 13, show that even before the first phase of bicycle boulevard implementation was completed in 2003, Milvia was often carrying more bicycle traffic than Shattuck. The average exposure measures used in this paper, 188.1 for Milvia and 185.0 for Shattuck, show Milvia with 1.02 times the volume on Shattuck. Three of the four counts in Table 13 indicate an even higher ratio, and one just slightly lower. So although the exposure measure I have used for Shattuck was based on just one count, there is plenty of reason to believe that the difference in risk between Milvia and Shattuck is quite real, and not just a by-product of one unusual count.

Table 13. Cyclist counts on Milvia and Shattuck before 2003. (PM peak 4:00p – 6:00p counts)

Milvia Count Date	Shattuck Count Date	Cross Street	Milvia Volume	Shattuck Volume	Ratio Milvia/Shattuck
2000/7/12	2000/7/12	Channing	110	124	0.89
2000/7/13	2000/7/13	Hearst	139	56	2.48
2001/10/9	2001/10/16	Channing	212.5	190.5	1.11
2002/10/9	2002/10/18	Channing	152	96	1.58

Other than Shattuck, the other streets' exposure measures were all based on at least three counts, as shown in Table 5.

The third question posed about counts is whether the ratio of A to B at weekday PM peak is representative of the ratio more generally. AM peak, at least, looks to be about the same as PM peak—the three intersections that have been counted both at AM peak and PM peak show quite similar numbers for both. Table 14 shows that both the absolute number of cyclists, and the ratio of north-south cyclists to east-west cyclists, are fairly consistent at given intersections whether counted at AM peak or PM peak. In fact, since all the streets I study in Berkeley are two-way, there is no particular reason to expect morning vs. evening variations. Consistency in the ratio of north-south to east-west traffic at an intersection, as shown in Table 14, does not exactly imply that the ratio of traffic on two parallel streets will necessarily also be consistent. However, it does support the more general theory that cyclists' route choice does not vary much between AM and PM peak, and so this may be true of parallel streets as well.

Table 14. AM vs. PM peak cyclist counts at three intersections in Berkeley, 2003-2010

Date	Time	Intersection	NS St	EW St	NS Volume	EW Volume	NS/EW
2010/3/23	8:25a-10:25a	Hillegass & Ashby	Hillegass	Ashby	73	18	4.06
2003/3/19	4p-6p	Hillegass & Ashby	Hillegass	Ashby	67	22	3.05
2004/10/20	4p-6p	Hillegass & Ashby	Hillegass	Ashby	57	19	3.00
2009/11/3	4p-6p	Hillegass & Ashby	Hillegass	Ashby	96.5	17.5	5.51
2003/10/23	7a-9a	Milvia & Channing	Milvia	Channing	149.5	95.5	1.57
2003/10/21	4p-6p	Milvia & Channing	Milvia	Channing	235.5	100.5	2.34
2004/10/21	4p-6p	Milvia & Channing	Milvia	Channing	193.5	100.5	1.93
2005/10/27	4p-6p	Milvia & Channing	Milvia	Channing	200	112	1.79
2009/10/29	4p-6p	Milvia & Channing	Milvia	Channing	286	183	1.56
2010/3/22	4p-6p	Milvia & Channing	Milvia	Channing	249	135	1.84
2003/10/22	7a-9a	MLK & Russell	MLK	Russell	18.5	68.5	0.27
2010/3/22	8a-10a	MLK & Russell	MLK	Russell	30	110	0.27
2003/4/9	4p-6p	MLK & Russell	MLK	Russell	31.5	75.5	0.42
2003/10/21	4p-6p	MLK & Russell	MLK	Russell	16.5	83.5	0.20
2009/11/3	4p-6p	MLK & Russell	MLK	Russell	67.5	221.5	0.30

key:

morning counts
evening counts

Supposing we assume that PM peak counts are, in fact, representative of a two-hour PM peak and a two-hour AM peak. This still represents just 20 hours a week, out of 168. Particularly for streets near campus, where it's likely that many cyclists are students, a majority of bicycle miles traveled probably occur outside of peak times. This is worrisome only if the ratio of cyclists on A to B varies by time of day. In fact, this is not implausible—perhaps evening and weekend cyclists, traveling to shopping or dining destinations rather than work, favor the retail-oriented Shattuck Ave. over more the residential and institutional Milvia St. Regrettably, there is no way to determine whether this is true using available data. However, there would need to be quite a lot of off-peak bicycle travel and a large difference in route choice at those times in order to erode away the larger risk ratios found between arterials and bicycle boulevards.

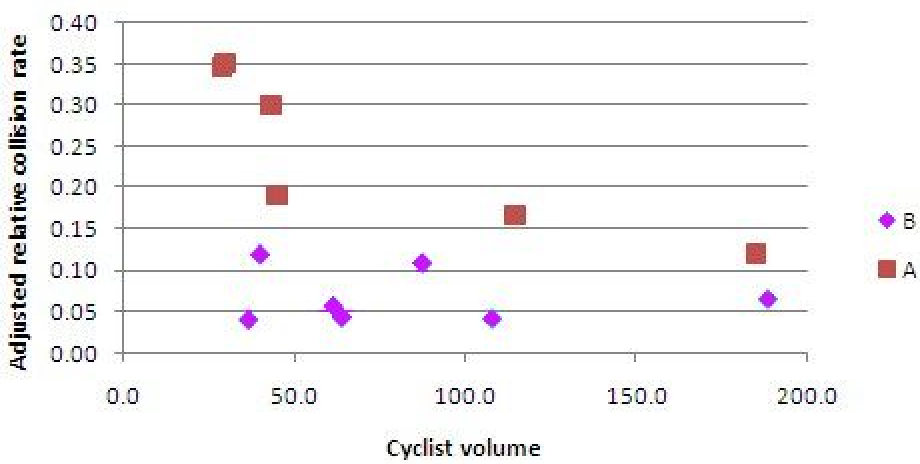
If one can be reasonably certain that the results in Table 8 reflect a real and significant difference in collision rate between bicycle boulevards and arterials, there is still room to ask what drives the difference. My hypothesis is based on literature findings about motor vehicle speed and volume and the presence of heavy vehicles. In other words, I would like to show that a difference in motor vehicle traffic characteristics is what creates the difference in safety for cyclists.

But as described in Methodology, the self-selection theory—that arterials simply attract dangerous cyclists, while bicycle boulevards attract safe cyclists—offers one other possible explanation.

Yet of the 1715 bicycle-motor vehicle collisions in the Berkeley dataset, police found the cyclist at fault in just 705 (41%). The motorist was found at fault in 925 collisions (54%) and the remaining 85 (5%) were no-fault. While this simple either-or assignment of fault does not tell the whole story of what causes a collision, these numbers are nonetheless hard to square with the idea that cyclist behavior wholly determines risk level, which is necessary in order for self-selection alone to create the apparent difference in collision rate.

Another theory is that of causative "safety in numbers"—perhaps, for any given level of exposure, a bicycle boulevard and an arterial would have an equal collision rate, but bicycle boulevards carry more bike traffic, achieving safety in numbers, and so have lower collision rates. In fact, this theory cannot explain away the difference in collision rates that I find. Figure 47 plots risk per cyclist versus exposure for each street. Risk per cyclist is an adjusted relative collision rate: to make all the streets at least roughly comparable to one another, I divided the relative collision rates from Table 8 by the length of street segment considered for each street. Volume, again, is simply the average two-hour PM peak count.

Figure 47. Collision rate vs. exposure for arterials (A) and bicycle boulevards (B) in Berkeley



The results overall are reasonably consistent with correlative safety in numbers: high collision rates on some low volume streets, low collision rates on high volume streets. The coloring of the dots for different street types lets us also ask about causation: do cyclists ride on streets that they (correctly) believe are safer for them, or does the presence of cyclists *make* a street safer for cyclists? Figure 47 suggests that street typology and design *do* make streets safer for cyclists, because at every point along the x-axis, the red dots (arterials) have higher collision rates than the purple dots (bicycle boulevards). Alongside that trend, it also appears that, for arterials at least, those with higher cyclist volume are safer for cyclists. Still, causation could flow in either direction. The two red dots towards the right are Telegraph and Shattuck. Perhaps the large number of cyclists there makes cycling safer, or perhaps the sharrows and intermittent bike lanes on Telegraph, and the low speeds and frequent stoplights on Shattuck are what make cycling safer, and thus draw more cyclists.

In sum, Figure 47 does not refute the idea that numbers cause safety, but it shows that, at a minimum, street typology matters as well. The results shown above, that bicycle boulevards have lower bicycle-motor vehicle collision rates than arterials, cannot be explained away by safety in numbers.

5.2 Cambridge

The results for relative collision rate are shown in Table 15. Collisions are over the whole period of study (5 years, 9 months) and exposure is the average two-hour count observed. Relative collision rate is simply collisions divided by exposure; its units are not meaningful, and comparisons should only be made within street pairs.

Table 15. Relative collision rate calculations for two street pairs in Cambridge

Street	Type	Collisions	Exposure	Relative Collision Rate
Green St.	side	3	24.25	0.12
Mass Ave. parallel to Green St.	arterial	49	311.5	0.16
Oxford St.	side	6	159	0.04
Mass Ave. parallel to Oxford St.	arterial	22	269.5	0.08

The results are not inconsistent with my hypothesis, as the relative collision rate is in each case higher for Mass Ave. than for the corresponding side street. The difference in rates, however, is much less dramatic than for Berkeley—whereas the risk ratio in Berkeley ranged from about 2 to 8, these pairings exhibit risk ratios of just 1.3 (Mass Ave. v. Green St.) and 2.2 (Mass Ave. v. Oxford St.).

As for severity, Table 16 shows that there is no discernible correlation between collision severity and location on or off a major road. The difference between 18 of 460 and 2 of 64 is clearly not statistically significant, and in any case, the data quality is questionable and the analysis crude. Accurately assigning collisions to the street on which the cyclist had been traveling would have required a careful reading of the summary of the police narrative included in the collision dataset. Instead, Table 16 is based simply on proximity: crashes within a 20-foot buffer of Cambridge's major roads (as depicted in red in Figure 30) are labeled as "On major roads," and the remaining collisions are labeled as "Non-major road locations."

Table 16. Severity vs. major road locations in Cambridge

Location	**Crashes**	**Severe injuries**	**Proportion severe**
On major roads	460	18	3.9%
Non-major road locations	64	2	3.1%
Total	524	20	3.8%

Assuming collision and count data to be accurate, I test whether random variation in an underlying Poisson process could have created the apparent difference in collision rates, as described in Appendix 2. The resulting p values are shown in Table 17 and are both above the .05 threshold, meaning the difference is not statistically significant. In other words, even assuming that collision and count data are accurate and representative, the difference shown here could well have been produced by random variation.

Table 17. Results of Poisson regression on collision rates in Cambridge

Arterial	Side street	Risk ratio	p value
Mass Ave	Green St.	1.3	0.686
Mass Ave	Oxford St.	2.2	0.094

Meanwhile, the exposure measure used here, average 2-hour volume, is based on far fewer observations than was the case for Berkeley. On the plus side, these were all paired counts, where the two parallel streets were counted at exactly the same time, but still, a fair amount of variation was apparent. Counts are included in Appendix 5. The paucity of count data means that the results in Table 15 must be taken with even less confidence.

There is another reason not to lightly conclude that traveling on side streets in Cambridge is safer than using the arterials. Since Cambridge's side streets are unconnective and may take cyclists far out of their way, a slightly lower risk per mile traveled on side streets may not mean a lower risk per trip.

Though I do not conclude that there is, in fact, a difference in safety between side streets and arterials in Cambridge, it is still worth raising the possibilities of safety in numbers or self-selection as playing a role in the streets' collision rates.

As for the self-selection theory, that side streets and arterials may attract different cyclists with different risk preferences, police records do not indicate that cyclists bring about all the collisions in Cambridge. Of the 524 collisions in the dataset, the cyclist was cited or warned in 54 (10%),the driver in 79 (15%), and both parties were cited or warned in 3 collisions (<1%). This leaves about three fourths of collisions where fault was either unclear to the officer at the scene or was not recorded. Based on these figures, it does not appear that cyclists are overwhelmingly the immediate cause of collisions, and so differences in cyclist risk preference would be unlikely to explain differences in safety between streets, if such a difference were observed.

Though I do not find a significant safety difference between arterials and side streets in Cambridge, the safety in numbers theory is worth addressing because a safety in numbers correlation can be roughly observed in Cambridge in time series on a city-wide basis. The number of cyclists counted nearly doubled between 2003 and 2008, yet the number of bicycle collisions reported each year was fairly flat, as seen in Table 18. Note that the 2009 figure is for January-September, though multiplying by 4/3 to extrapolate to the full year would probably overestimate total collisions for 2009, as Cambridge's bicycle collisions are seasonal, with August and September being the worst months (Parenti 2006).

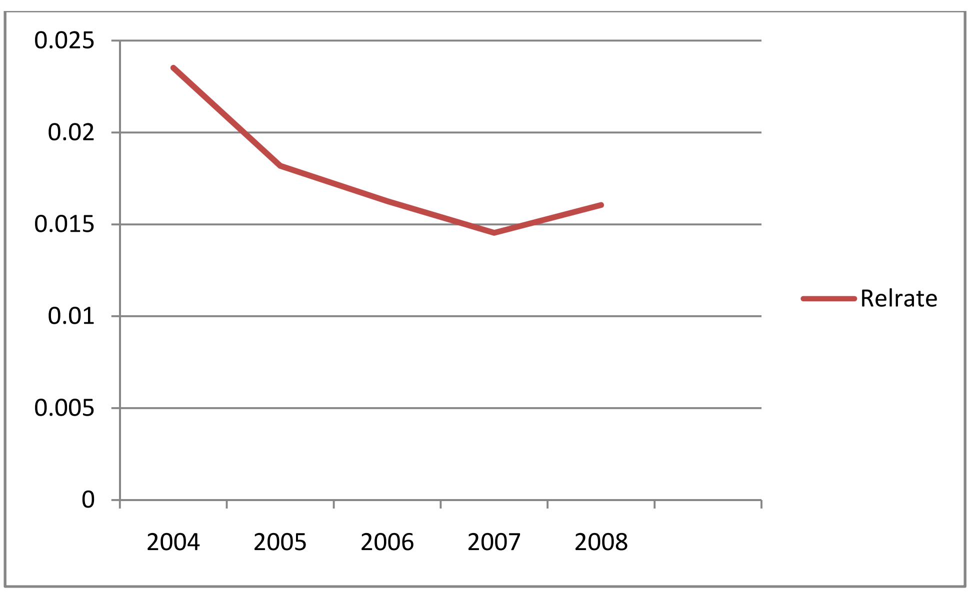
Table 18. Police-reported bicycle collisions per year and cyclists counted at 16 intersections in one hour each of AM and PM peak in Cambridge.

Year	Collisions	Cyclists
2003		3286
2004	90	3827
2005	83	4563
2006	85	5226
2007	76	
2008	102	6353
2009*	86	

*Collisions for 2009 represent only January - September

Using the 2006 cyclist count for 2007, which is missing a count, and simply dividing the number of collisions by the total cyclist count, it is possible to graph a relative collision rate over 2004-2008 for Cambridge, as shown in Figure 48. The units of the rate are not meaningful, but the trend is clear: it appears that risk to each cyclist has dropped over this five-year period.

Figure 48. Relative collision rate per cyclist in Cambridge 2004-2008



This could be because numbers cause safety, or because Cambridge has made its streets safer and thereby enticed more cyclists onto the road. Either way, the fact that correlative safety in numbers is observable in Cambridge over this time period makes it natural to ask why Mass Ave., with so many cyclists, is not found to be safer than Green St. or Oxford St. In fact, it is possible that numbers do make Mass Ave. safer—not safer than side streets with few cyclist, but safer than Mass Ave. itself would have been with few cyclists. The safety in numbers effect of high cyclist volume may be small, and thus outweighed by the "danger in numbers" effect of high motor vehicle volume. On the other hand, it is also possible that the safety in numbers effect is quite large, and that without it, Mass Ave. would have rated as much more dangerous than the side streets I study here.

6 Conclusions

The results shown above allow not only for a testing of my hypothesis in Berkeley and Cambridge, but also for recommendations about cities' collection of data on cyclists and for a broader discussion of city policy and cycling. This section will address all of these points.

6.1 Hypothesis tested

Findings in the literature led me to hypothesize that side streets are safer for cyclists than arterials, with lower bicycle-motor vehicle collision rates (due to lower motor vehicle volume) and a smaller proportion of those collisions being severe (due to lower motor vehicle speeds and fewer heavy vehicles). I hypothesized that this would hold both in Berkeley and in Cambridge, two cities with very different strategies, where Berkeley has accommodated cyclists on side street bicycle boulevards while Cambridge has striped bike lanes on its arterial streets.

Berkeley has created, as the centerpiece of its bicycle plan, a network of seven bicycle boulevards—traffic-calmed side streets designated and improved for bicycle use—that stretches over the city. Compared to the arterials that they parallel, the bicycle boulevards have low motor vehicle volumes and speeds and few heavy vehicles. Using police-reported collision data and the city's cyclist count data, I find strong evidence that bicycle boulevards do indeed have lower collision rates for cyclists than their parallel arterial routes. This is true for all six bicycle boulevard-arterial pairs for which data are available, with risk ratios ranging from 1.8 to 8.0. This is true whether only reported bicycle-motor vehicle collisions or all reported crashes are examined. Poisson regression reveals that the difference in collision rate is highly statistically significant. Cyclist count data provides a good enough measure of true exposure as to not call these findings into doubt. The ratios of cyclist volume on arterials to adjacent bicycle boulevards exhibit consistency over space when measured at different cross streets. The ratios of north-south and east-west counts at the intersections counted repeatedly over multiple years likewise exhibit a consistency suggesting that route choice is consistent over time, even if total volume fluctuates. I am unable to rule out the possibility that off-peak travel patterns may be different than the peak hour patterns measured in my count data, but the difference would have to be rather large in order to erode away the differences in collision rate that I have observed. Cyclist self-selection is unlikely to explain the poorer safety record of some streets, because cyclists are found at fault in only a minority of collisions with motor vehicles. Neither can safety in numbers explain away the differences I observe, because at all levels of exposure, bicycle boulevards prove to have a lower collision rate than arterials. I conclude that street typology must account for the difference in collision rate. As for the proportion of collisions that are severe, I find no significant difference between the two street types. This may reflect a threshold of severity needed in order for a collision to be reported, which would mean that the difference in total risk of severe injury is actually comprised of a smaller difference in collision rate than I have observed, multiplied by some difference in the proportion severe. Either way, it is true that bicycle boulevards carry a lower risk than arterials of severe injury to cyclists.

Cambridge has pursued the rather different strategy of adding bike lanes to most of its arterial streets, with relatively few accommodations on side streets. Given Cambridge's non-Cartesian street network, arterials are the most direct routes and side street alternatives do not always exist. I address two segments where an untreated side street runs parallel to a major arterial treated or partially treated for cyclist use. Analyzing police-reported cyclist collision data and my own count data, I find lower collision rates on the side streets, but the difference is small, with risk ratios of 1.3 and 2.2, and Poisson regression shows that the difference is not statistically significant. Moreover, my analysis of Cambridge is based on very limited count data which may not accurately reflect cyclist volume over the period of study. There is no statistically significant difference in the proportion of collisions that are severe on side streets versus on arterials. The fact that arterials are not found to be safer than side streets in Cambridge is interesting in light of the much higher cyclist volume on arterials. I speculate that safety in numbers may play some role in improving safety on arterials, but not enough to outweigh the effect of street typology.

6.2 Data collection practices

Current cyclist count data collection practices do not especially enable this type of research. It is useful for researchers to be able to know collision rates on individual streets, and this requires collisions grouped by street as well as a good measure of exposure—something approximating bicycle miles traveled—for each street. This is necessary not only for comparing parallel routes, as I have done, but also for before-after studies of the same street when a bicycle facility has been added or other intervention undertaken.

Collision data ideally should contain attributes that allow collisions to be assigned to streets of interest without personally reading text about each one. Cambridge's collision data does not allow this, and though reading the text for tens of collisions as studied here is not especially burdensome, a city-wide analysis, especially with more years of data, would be taxing. Alameda County's SWITRS data, on the other hand, embeds the necessary information in the Direction of Travel and Movement Preceding Collision attributes, and while extracting the information requires a bit of thought, it can be done with SQL code rather than manually. A best practice would be to simply have one column indicating the street on which the cyclist was traveling. A second column could indicate the street, if any, that the cyclist turned onto, so that some collisions, as appropriate, could be assigned to two streets.

Count data ideally should be collected at several points along each street so that counts provide a good estimate of cyclist traffic across the street's entire length. Counts should be collected at identical points along parallel routes so that counts are highly comparable, and a variety of collection times and seasons can help control for temporal variations. Count data should be available for each street rather than just by intersection.

Of course, all this is easier said than done: counts are expensive to collect. Conducting many more counts would be a boon to safety research such as this, but would probably require automatic counting technologies, which cost a few thousand dollars per unit (Alta Planning and Design 2009). Then again, this might eventually pay for itself in staff hours saved from manual counts. In any case, at a minimum, collecting cyclist counts for streets rather than just intersections is crucial for this sort of analysis. In fact, officials in Berkeley and in Portland, Oregon already collect this data, but aggregate it by intersection, throwing out direction of travel data, when entering data into electronic format. I originally planned to include Portland in this study but eliminated it for this reason. Direction of travel data—or more precisely, data on *which streets the cyclists were on*—is needed in order to establish collision rates for each street. Since this data is collected to begin with, just preserving it when counts are typed up should not be too costly.

6.3 Policy implications

The comparison of Berkeley and Cambridge lends itself to a variety of different interpretations. Since Berkeley's bicycle boulevards are much safer than its arterials, while Cambridge's side streets appear to be just a bit safer, if at all, one interpretation is that side streets *can* be safer alternatives to arterials, but only if treated correctly with traffic calming, signage and many of the other treatments that make Berkeley's bicycle boulevards what they are.

On the other hand, maybe Cambridge's side streets are similar to its arterials in risk not because Cambridge's side streets are more dangerous than Berkeley's, but because Cambridge's arterials are safer than Berkeley's. I have no way to compare streets between these two cities in any meaningful way, but the possibility is worth considering. The section of Mass Ave. adjacent to Green St. has bike lanes as well as a number of traffic calming features, such as neck downs for pedestrian crossings, and it is imaginable that these treatments have made Mass Ave. nearly as safe as a side street.

If bicycle boulevards really are the safest place for cyclists, this is still not enough to conclude that creating bicycle boulevards improves safety. A road treatment may be said to improve cyclist safety if it makes existing cyclists on that route safer or if it induces more cyclists to choose the safer route. So a recommendation in favor of turning side streets into bicycle boulevards would require knowing that this treatment either improves safety on those streets, or induces more cyclists to use those streets. To address the first possibility, that particular streets became safer after being treated as bicycle boulevards, would require a before-after study, which I have not conducted here, but would be a valuable area for future research. As for the second possibility, in Appendix 4 I examine the second possibility, that cyclists have shifted from arterials to bicycle boulevards, and find that data are too sparse to give any firm conclusions.

However, I have written from the standpoint of a city making an exogenous decision about bicycle planning, when in fact the decision clearly depends on input from cyclists and other citizens. It may be that increased cycling on side streets brought about bicycle boulevards, rather than that bicycle boulevards brought about increased cycling on side streets. In this case, perhaps the main function of the bicycle boulevards was to improve cyclist convenience (through the replacement of stop signs with traffic circles and the addition of priority crossings of major streets) and cyclist comfort (through the reduction of motor vehicle volumes and signage to indicate that cyclists take priority).

This raises the valuable point that safety is never the only goal of a city's bicycling policy. In online literature describing what motivates their bike strategies, Berkeley and Cambridge both mention safety alongside points such as cyclist speed and convenience, comfort, and inducing more people to take up cycling (Wilbur Smith Associates 1998 Berkeley Bicycle Plan and City of Cambridge, Bike Lanes). They also both address concerns about the impact on other road users from a bicycle plan. All of these may be major reasons why stakeholders make up their minds in favor of bicycle boulevards, bike lanes, or neither.

Critics of bicycle boulevards with whom I have spoken in the course of this research have expressed concern about bicycle boulevards being indirect and inconvenient. Relatedly, Cambridge's Community Development Department explains that it promotes cycling on arterials in part because cyclists "want to get to their destinations by the most direct route available" (City of Cambridge, Bike Lanes). In Cambridge, it is clear that arterials are the quickest route for cyclists. In general, though, cyclists on arterials must contend with congestion and red lights, while a bicycle boulevard, properly implemented with traffic circles instead of stop signs, priority for crossing major streets, and located just one block away from the arterial, actually has the potential to be the quicker through street for cyclists, though it would still provide less direct access to shops on the arterial. There is no easy answer here: public policy must weigh the potential for providing cyclists with convenience on either route, along with the safety implications of the two routes.

Another fear that I have heard cyclists express about bicycle boulevards is "ghettoization"—the idea that cyclists will be confined to one street rather than allowed to use any street. This does not appear to be the case in Berkeley, where many cyclists still use arterials, particularly the retail-oriented ones such as Shattuck and Telegraph. Related to the "ghettoization" concern is the notion that cyclists will be less visible to motorists on a side street, and so much of the benefit of "safety in numbers" will be lost. Yet my results are suggestive that side streets are safer than arterials regardless of where the numbers are.

As I mention above, the impact on other road users is another important factor in the designation of a bike route, and likely to be part of any public dialogue on the subject. On Mass Ave., where bike lanes are located on the right side of the roadway yet buses must pull right to make stops, conflicts between cyclists and buses are frequent. Bicycle boulevards offer the possibility of accommodating cyclists without interfering with bus travel. Bicycle boulevards might also be more acceptable to motorists, since they keep cyclists out of the key routes for cars, though this idea in turn is offensive to vehicular cyclists who believe in sharing the road, and might even be worrisome to pedestrians who do not want the cars to speed on arterials. Then again, some pedestrians are afraid of being hit by cyclists themselves and may prefer to have the cyclists off of arterials. In any case, it is clear that cyclists are not the only stakeholders here, and that is another reason why a specific policy recommendation in favor of bicycle boulevards does not emerge from my findings.

However, a more general policy recommendation does emerge: route choice does carry safety implications, and is therefore worth considering when developing a bike plan. The two- to eightfold collision rate difference observed between bicycle boulevards and arterials in Berkeley is evidence that some routes are, or can be made, much safer than others, and this surely merits a place in any public discussion about bicycle network planning. Debates over *what* should be done for cyclists ought to also include *where* it should be done.

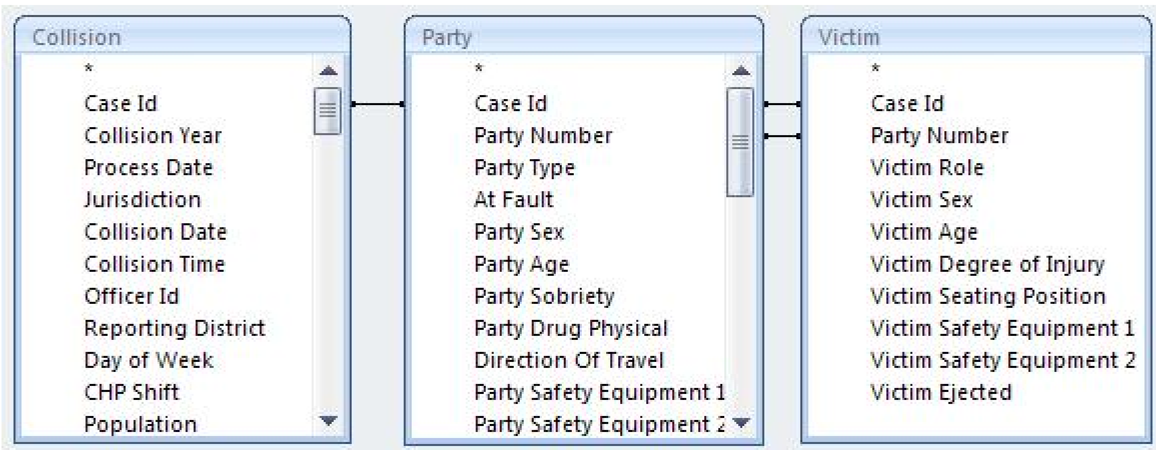
In any event, the decision of whether to create a bike network on arterials or side streets is just an initial step for cities that have not yet done anything to accommodate cyclists. In the long run, this need not be an either-or choice. The side street network proposed by Cambridgebikes.org shows that, even after bike lanes are installed, there is still a desire from some cyclists to have an alternative to arterials. Meanwhile, count data in Berkeley indicate that many cyclists are still choosing arterials over bicycle boulevards: Shattuck carries nearly as many cyclists as Milvia, and Telegraph carries more cyclists than Hillegass-Bowditch. The City of Berkeley, concerned about sidewalk riding and lack of cyclist last-block access to destinations, plans to address bicycle access on commercial arterials in its 2011 Berkeley Bicycle Plan Update (Anderson, personal email, 2010).

It appears, then, that having started with rather different strategies, Berkeley and Cambridge may both wind up with parallel side street and arterial bike networks. While this research has sought to help cities beginning to accommodate cyclists answer the question of "where", it seems that in time, cities where cycling becomes popular may converge on the same answer: "everywhere."

Appendix 1. SWITRS database schema and data processing

As indicated under Methodology, the SWITRS database contains three tables, whose relationships are shown in Figure 49. Each row of Collision represents one incident, while each row of Party represents a person involved. Victim also represents a person, but only injured persons are included, so a record in Party will not necessarily have a match in Victim.

Figure 49. Relationship diagram of SWITRS tables



This three-table diagram is a simplification, though, because the dataset I obtained in fact contained two tables of collisions—Collision, shown above, as well as CollisionWGeocodingWPlace, which like Collision would be linked to Party and Victim by Case Id, and contains all the attributes of Collision plus several geographic attributes.

The Collision table contains 346,458 records and the CollisionWGeocodingWPlace table contains 136,912 records, so it is tempting to assume that the latter is a subset of the former. However, an inner join of the two tables returns just 116,574 rows, meaning that the two tables include an overlapping set of collisions, but that each contains some collisions that the other does not. I ran some tests to examine, for instance, the range of dates of bicycle collisions included, and did not see any obvious answer as to on what criteria some collisions were included in one table and not the other.

The non-geocoded collisions in the Collision table contained Primary Rd and Secondary Rd attributes which would be sufficient for geocoding as long as the city were known. However, the only information included on city location was a Jurisdiction number. The SWITRS database includes a City table which provides supposedly one-to-one matches of Jurisdiction with Alameda County cities. However, on examining the CollisionsWGeocodingWPlace table, I determined that most matches were in fact many-to-many—for instance, Jurisdiction 105 contained collisions matched to addresses in Fremont, Newark and Union City, while collisions matched in Union City might be from Jurisdiction 105, 106 or 113. Berkeley, thankfully, was almost perfectly one-to-one: all of the geocoded collisions in

Berkeley were listed as Jurisdiction 103, and Jurisdiction 103 contained almost no locations in other cities (just a handful in neighboring Oakland and Emeryville).

I selected all of the bicycle collisions in Berkeley from CollisionWGeocodingWPlace, and found 1989 records. Next I selected from Collision those bicycle collisions which occurred in Jurisdiction 103, and found 1954 rows, 1823 of which were also among the 1989 geocoded collisions. This left 131 records for me to geocode.

I selected these 131 records from the Collisions table and used ArcGIS to geocode based on Primary Rd and Secondary Rd as the intersection and assuming Berkeley as the city. In the end I threw out six which either I could not match or which I discovered were in Oakland or Emeryville. This left me with 125 collisions to add to the dataset. I created a new table from them, named bColBerk, and appended the original 1989 records for a total of 2114 rows. Geocoding was labor-intensive, and would have been even moreso for the many Alameda County cities which match SWITRS Jurisdiction codes on a many-to-many basis, so I did not bother to geocode collisions in the rest of the county. This is why my comparative analysis of general severity trends in Alameda County bicycle collisions (Table 3) is limited to the 7707 collisions that were already geocoded.

Next I set about collapsing the relevant information from the relational database into a single table so that the resulting table could be imported into GIS and all of the relevant information would be associated with a point on the map. Since my approach was initially exploratory, it was important to be able to see party-level data while working in GIS. I created tables, t_cyclAgg and t_drivrAgg by selecting all important attributes from entries in PARTY for cyclists and drivers respectively and aggregating these by Case Id, using FIRST() as the aggregator function for most attributes. I created tables, rather than just queries, because next I had to throw out much information from the handful of cases where more than one cyclist or more than one driver were involved, recording 'Multiple' in the affected cells. Since many attributes of interest were nonnumeric, for instance, Direction of Travel ('N','S','E','W','-'), it would have required considerable effort to give meaningful special treatment to multi-cyclist or multi-driver collisions rather than just using the FIRST() aggregator. Since only five bicycle-motor vehicle collisions (0.3%) involved more than one cyclist, I didn't bother. Later, I did re-join my final table to the original PARTY table to investigate the circumstances of, among other things, bicycle-bicycle collisions.

I joined t_cyclAgg and t_drivrAgg to bColBerk and created a new table, t_compiledData, which I brought into GIS and subsequently used for most of the analysis of Berkeley in this thesis. While selecting comparable street segments, I added two columns to t_compiledData: NS_St and EW_St, indicating what, if any, streets of interest (i.e. selected segments of bicycle boulevards and arterials) the collision had occurred *near*—I did not yet determine whether the cyclist had been in fact traveling along that street.

Once the collisions had been labeled by their proximity to streets of interest, I added two more columns, cyclNS and cyclEW, indicating with what level of certainty it could be believed the cyclist had been traveling along the north-south and east-west street respectively. The numbers 0, 1, and 2 indicate respectively no inclusion, generous inclusion and conservative inclusion. Table 19 shows the SQL code used to implement these conservative and generous inclusion criteria for north-south streets—the codes ‘D’, ‘E’ and ‘F’ for cyclMove all indicate turns. Of course, the equivalent was done for east-west streets as well.

Table 19. SQL implementation of generous and conservative inclusion criteria

```
-- Criteria for generous inclusion:
UPDATE t_compiledData SET cyclNS = 1
WHERE cyclDir='N' Or cyclDir='S' Or cyclDir='-' Or cyclDir='Multiple' Or
intCol='Y';

-- Criteria for conservative inclusion:
UPDATE t_compiledData SET cyclNS = 2
WHERE (cyclDir='N' Or cyclDir='S') Or ((cyclMove='D' Or cyclMove='E' Or
cyclMove='F') And intCol='Y');
```

Finally, to compute the number of collisions on each comparable street segment, I simply counted the collisions with a 2 in the cyclNS (or cyclEW) column for conservative inclusion, and counted those collisions with a 1 OR 2 for generous inclusion.

I ran this process of counting collisions on each street segment three separate times: first, with all 2114 records from t_compiledData included—a measure of “All Crashes”, then with the one-party falls deleted—1782 records of “All Collisions”—and finally with the bicycle-bicycle and bicycle-pedestrian crashes deleted as well—1715 “Bicycle-Motor Vehicle Collisions Only.” I copied each set of crashes into a separate table before deletion, so that I could analyze them separately. The SQL “WHERE” clauses used to identify crashes by type are shown in Table 20. Recall that t_compiledData already only contained collisions involving a bicycle, so a test for bicycle involvement was not needed in the WHERE clause.

Table 20. SQL WHERE clauses used to identify crash types

```
-- Bicycle-pedestrian collisions
WHERE pedCol = 'Y' AND numDrivr IS NULL;

-- Bicycle-bicycle collisions
WHERE numCycl > 1 AND pedCol IS NULL and numDrivr IS NULL;

-- Falls
WHERE pCount = 1;
```

This leaves any collision not captured by the above criteria to be considered as a bicycle-motor vehicle collision. Of the 1715 bicycle-motor vehicle collisions thus left over, 1698 were clearly bicycle-motor vehicle collisions because they contained party-level information about the driver. The other 17 had no party-level record for the driver. 3 of those were hit-and-runs where, presumably, no information had been gathered on the driver. For others, it was unclear why no driver data was included. I examined the Collision Type attribute, and some of the 17 collisions were of types

(“Sideswipe,” “Broadside,” and so on) that clearly involved a real collision. “Hit fixed object”, on the other hand, may have involved a parked car, possibly a dooring, or could indicate a one-party cyclist fall for which the party count was erroneously recorded as 2 instead of 1. So a handful of these 17 collisions might not be true bicycle-motor vehicle collisions. In any case, these 17 collisions make up less than 1% of the bicycle-motor vehicle collision data, so even if the inclusion of some subset of them is in error, it does not seriously affect my results.

Appendix 2. Methodology for Poisson regression on collision rates

The number of bicycle-motor vehicle collisions occurring on a given street in a given period of time may be considered as an event count. In particular, such collisions are rare events—the probability of any one cyclist being involved in a collision on any one trip is extremely low, but the number of cyclist trips taken on a given street over nearly six years is enormous.

As a back of the envelope calculation, take the average two-hour weekday PM peak volume on the Milvia St. Bicycle Boulevard, 188 cyclists. On the conservative assumption that PM peak represents 1/3 of daily volume and that there are 250 work day equivalents in a year (no cycling takes place on weekends), Milvia carries 188*3*250 = 141,000 cyclist trips each year. Over five years and ten months, this is 822,500 cyclist trips. On the more generous assumption that PM peak is 1/5 of volume and there are 365 work day equivalents per year (people cycle as much on weekends as weekdays), that number could be 188*5*365*5.83 = 2 million cyclist trips.

Out of so many trips on Milvia came only 26 police-reported bicycle-motor vehicle collisions. So for any one cyclist trip, the probability of experiencing a collision is on the order of one in a hundred thousand.

If each cyclist trip is a trial, and the two possible outcomes are one collision or no collison, then collisions follow a binomial distribution. In the limit where the probability of an event per trial is small but the number of trials huge, the total number of events occurring is described by the Poisson distribution. Cameron and Trivedi (1998) observe that the count of such events that have occurred is poorly modeled by Gaussian linear regression because a Gaussian model assumes unbounded support for the count data—that the number of events could be any real number from negative infinity to positive infinity. Obviously, the number of events occurring must be an integer, and cannot be less than zero. They describe Poisson regression, a form of a generalized linear model, as a model which respects the underlying rare event characteristics of the data.

Poisson regression is better than Gaussian linear regression for rare event counts, but still not a perfect model for the real world—many studies of rare events, for instance in occupational safety, exhibit more zeroes than would be expected based on a Poisson distribution (Carrivick, Lee and Yau 2003). This phenomenon is known as over-dispersion, since it violates the Poisson assumption that each cyclist is exactly equally risky as the next. Negative binomial and zero-inflated Poisson (ZIP) models offer an improvement over Poisson regression in cases of over-dispersion, but for simplicity I conduct Poisson regression on my data. I use cyclist trips as the exposure measure over which risk is distributed.

In plain language, then, my assumption is that bicycle-motor vehicle collisions occur as a Poisson process over cyclist trips, with each street's Poisson process having its own parameter, or frequency of such collisions per cyclist trip. I assume that each street's parameter is invariant over time during the period of study, January 2003 through October 2008. I wish to test whether the parameters of two streets A and B are significantly different—in other words whether, if the parameters were in fact the same, the probability would be very small of observing at least as large a difference in collision rate as I have observed.

One way to do this would be to simply test the difference of Poisson means. Assuming, again, that yearly traffic is 750 times the average PM peak observed volume, then during the entire period of study, Milvia had 26 collisions over 822,500 cyclist trips while neighboring Shattuck saw 47 collisions over 809,000 cyclist trips. Thus the two Poisson means are 26/822,500 and 47/809,000. Ng, Gu and Tang (2007) compare several different methods of conducting a test for difference of Poisson means. However, since my collision data are labeled by date, it is possible to do better than this. By breaking the collision totals—26 and 47—into subtotals for each year of study, I can use Poisson regression over each year's totals to determine the statistical significance of the difference between the two streets' collision rate parameters in the context of checking how well the Poisson model holds up.

The Poisson regression model I assume for my data is:

Equation 4. Poisson regression model with exposure offset

$$\ln(collisions) = \ln(exposure) + a + b * sidestreet + \varepsilon$$

Where the error term ε is Poisson distributed. Exposure is an *offset*, so its coefficient is constrained to be 1. Since ln(collisions) – ln(exposure) = ln(collisions/exposure), Equation 4 is equivalent to the more intuitive Equation 5:

Equation 5. Poisson regression model with rate as dependent variable

$$\ln\left(\frac{collisions}{exposure}\right) = a + b * sidestreet + \varepsilon$$

While Equation 5 is more elegant, Equation 4 represents the format in which this regression must be inputted in the statistical software package R.

To conduct the regression, I break the total collisions up into subtotals of collisions by year, for each street. I choose years because if there is actually seasonal variation in exposure, and therefore in collisions, then any smaller unit of time would introduce the confounding variable of seasonality. Since I do not have separate exposure data for different seasons, I wish to avoid this. Annual collision data are shown in Table 21 and Table 22.

Table 21. Subtotals of bicycle-motor vehicle collisions in Berkeley, by street and year

Street	2003	2004	2005	2006	2007	2008*	Total
Hillegass-Bowditch	2	1	1	0	1	0	5
Telegraph	3	3	5	2	5	3	21
College	4	5	2	8	5	6	30
Milvia	4	3	4	8	4	3	26
Shattuck	8	8	7	7	5	12	47
MLK	6	2	3	4	2	1	18
California	1	1	0	2	3	2	9
Sacramento	2	3	1	4	3	2	15
Ninth	1	0	1	0	1	0	3
San Pablo	2	1	4	5	10	4	26
Virginia	0	1	2	1	1	1	6
University	1	2	2	6	2	4	17
Channing	2	2	1	3	4	7	19
Dwight	3	1	3	2	0	2	11
Russell	1	0	1	3	0	2	7
Ashby	6	3	6	2	7	2	26

*2008 data represent collisions for January – October only.

Table 22. Subtotals of bicycle-motor vehicle collisions in Cambridge, by street and year

Street	2004	2005	2006	2007	2008	2009*	Total
Green St.	0	2	0	0	1	0	3
Mass Ave.	6	7	11	7	9	9	49
Oxford St.	1	0	1	1	3	0	6
Mass Ave.	6	4	2	2	4	4	22

*2009 data represent collisions for January – September only.

Next I need an exposure measure for each year. I use the average two-hour count (PM peak only in the case of Berkeley, AM and PM peak for Cambridge), inflated by a multiplier to represent total yearly volume. Conservatively, I assume that two hours of PM peak, or one hour each of AM and PM peak, represents 1/3 of total daily volume, and that there are just 250 work day equivalents per year. Therefore the multiplier is 3 * 250 = 750. Since the average count on Milvia was 188.1 cyclists, then, I obtain a yearly exposure of 750 * 188.1 = 141,075 cyclists. This conservative assumption may not be accurate, but the exact multiplier is not important. The statistical significance obtained by Poisson regression is not especially sensitive to the scale of the exposure measure, so I could have used a multiplier of 5*365 = 1825, as described above, and gotten the same results. Using some sort of plausible multiplier, though, is important for consistency with my assumption of an underlying Poisson distribution. To use no multiplier and simply conduct the regression on 26 collisions out of 188 cyclist trips would be unfaithful to this assumption, since in that case, collisions would not be “rare”, as Poisson requires, but in fact rather frequent—indeed, so frequent that more than one collision could happen per cyclist trip, which complicates the mathematical formulation of the model.

Note that I have not adjusted for street segment length or made any assumption about the trip length for each cyclist trip, so it is still the case that each bicycle boulevard can be compared only to the arterials that it parallels, since I use identical segments of these street pairs.

I apply the same exposure measure for every year because separate count data are not available for every street in every year, and even if they were, this would introduce daily variability in count volumes since each year’s count would be from one or at most two days. Since there appears to be no trend in cyclist volumes in Berkeley over the time period of study, as shown in Figure 46, averaging over multiple counts in different years smoothes out daily variations without obscuring any important year-on-year trends. In Cambridge, I have cyclist count data from only one year (2010), so the question is moot.

The model in Equation 4 takes the log of collisions, which means it is not acceptable to have zero collisions in any record. Therefore, I lump years with zero collisions in with years that have at least one collision and adjust the exposure measure accordingly. For instance, Ninth St. in Berkeley had one collision in 2003 and no collisions in 2004, so I create a row with one collision and a two-year exposure measure. I similarly adjust the exposure measure for 2008 in Berkeley and 2009 in Cambridge to reflect that collision data are from less than a full year.

Table 23 shows an example of the final data table created for each street pair. In Table 23, Ninth (a bicycle boulevard) and San Pablo (an arterial) are being compared. I choose this pairing as an example because it required grouping of years due to zero collisions in some years on Ninth. Each row represents a “street-year”. Since Ninth only had three years with collisions, pairs of years are lumped together. The exposure of 54,560 listed for 2003-4 and for 2005-6 is the average two-hour PM peak count on Ninth St., 36.4 cyclists, times the multiplier 750, times 2 for the two years. The exposure for 2007-8 is simply 22/24 of that figure, to account for only 10 months of 2008 being included. San Pablo had at least one collision every year, so it has six separate records, with the exposure for full years being 32,375, which is again the average count, 43.2, times 750.

Table 23. Records created for Poisson regression for Ninth v. San Pablo

		coll	exposure	bblvd
Ninth	2003-2004	1	54650	1
	2005-2006	1	54650	1
	2007-2008	1	50096	1
San Pablo	2003	2	32375	0
	2004	1	32375	0
	2005	4	32375	0
	2006	5	32375	0
	2007	10	32375	0
	2008	4	26979	0

The rightmost three columns of the above table were then read into the statistical software program R and regressed using the glm function for generalized linear models. By default, glm would treat my "exposure" column as an explanatory variable and compute a coefficient for it, but as shown in Equation 4, I need to constrain the coefficient of "exposure" to be 1 so that the model is equivalent to Equation 5, where "bblvd" is the one explanatory variable and collision *rate* is the dependent variable. Introducing "exposure" as an *offset* constrains the coefficient to be 1. Code and output are shown in Figure 50 with two points of interest bolded.

The coefficient estimated for the bblvd flag is about -1.9899. This means that the ratio of collision risk per cyclist trip on San Pablo to Ninth is estimated to be $1/e^{-1.9899} \approx 7.3$. I could calculate a risk ratio much more simply by finding a collision rate (collisions/exposure) for each street and then dividing San Pablo's collision rate by Ninth's, and get essentially the same answer.

The bblvd flag is found to be highly statistically significant, with a p value of .0011. This means that a difference at least as large as that observed in collision rate between "street-years" on Ninth and "street-years" on San Pablo would occur only with probability .0011 if the two streets actually had the same underlying collision rate.

Figure 50.Code and output of Poisson regression for Ninth v. San Pablo using R

```
> p<-read.table("nvsp.csv", sep=",", header = TRUE)
> p
  coll exposure bblvd
1    1    54650     1
2    1    54650     1
3    1    50096     1
4    2    32375     0
5    1    32375     0
6    4    32375     0
7    5    32375     0
8   10    32375     0
9    4    26979     0
> m1<-glm(p$coll~offset(log(p$exposure))+p$bblvd, family=poisson)
> summary(m1)

Call:
glm(formula = p$coll ~ offset(log(p$exposure)) + p$bblvd, family = poisson)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.9812  -0.2204  -0.0283   0.1464   2.2530

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  -8.8906     0.1961 -45.334   <2e-16 ***
p$bblvd      -1.9899     0.6097  -3.263   0.0011 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 28.068  on 8  degrees of freedom
Residual deviance: 10.849  on 7  degrees of freedom
AIC: 39.631

Number of Fisher Scoring iterations: 5
```

I repeated this process for all street pairs. Results for Berkeley and Cambridge are shown in Table 24 and Table 25. (These tables are copied in Analysis and Results as Table 10 and Table 17). Risk ratio is the estimated ratio of the collision rate per cyclist trip on the arterial to that on the side street. P values given are for the significance of the “bblvd” flag in explaining the difference in collision rates between “street-years.” Differences for every street pair in Berkeley are statistically significant, with p values well below .01 for all of them except for Milvia v. Shattuck, which anyway is still well within the .05 level of significance. In Berkeley, it is safe to reject the null hypothesis that there is no difference in collision rate between each bicycle boulevard and its corresponding arterial. Of course, concluding that the rate is really different requires a reexamination of the collision and count data themselves, and this is the purpose of much of the discussion in Analysis and Results above. Differences between the two street pairs studied in Cambridge are not statistically significant, and it is not possible to reject the null hypothesis in Cambridge. This means that even if collision and count data are completely reliable estimates of total collisions and total exposure, the difference observed could still be due to chance.

Table 24. Poisson regression results for Berkeley

Arterial	Bicycle Boulevard	Risk ratio	p value
Telegraph	Hillegass-Bowditch	3.9	0.0058
Shattuck	Milvia	1.8	0.0128
MLK	Milvia	2.9	0.0005
San Pablo	Ninth	7.3	0.0011
University	Virginia	6.0	0.0002
Ashby	Russell	8.0	< 0.0001

Table 25. Poisson regression results for Cambridge

Arterial	Side street	Risk ratio	p value
Mass Ave	Green St.	1.3	0.686
Mass Ave	Oxford St.	2.2	0.094

Since it is possible that a negative binomial model would have fit the data better than a Poisson model, I conducted a quick test for over-dispersion using the odTest function (in R's pscl library). It performs a likelihood ratio test to accept or reject the null hypothesis that the Poisson model, rather than negative binomial, is the correct specification (R Documentation, R: Likelihood ratio test for over-dispersion in count data). For all eight street pairs, it was not possible to reject this null hypothesis.

Appendix 3. Expanded analysis of collision rates for Berkeley

The results that I found for Berkeley are quite robust in that changing the methodology a bit still gives essentially the same results. In particular, the risk difference between bicycle boulevards and arterials holds up even if different criteria are used to assign collisions to streets, and even if all crashes, rather than just bicycle-motor vehicle collisions, are included.

Conservative vs. generous inclusion criteria

Recall that I developed two formulae, which I have called conservative and generous, for determining whether a collision should be assigned to a given street of interest. SQL code used to implement these formulae is shown in Appendix 1.

I believe the conservative criteria provide a better measure, so I used them in computing the relative collision rates in Table 8 and conducting Poisson regressions in Appendix 2. However, I also computed relative collision rates using the generous criteria, and found essentially the same results: again, bicycle boulevards are always safer than their corresponding arterials, and the risk ratios range from 2.0 to 6.1 rather than 1.8 to 8.0. Results from conservative criteria and generous criteria are compared side by side in Table 26. The risk ratio for each arterial is its collision rate (if available) divided by that of the corresponding bicycle boulevard.

Table 26. Collision rates and risk ratios in Berkeley under conservative vs. generous inclusion criteria for collisions.

Street	Conservative Inclusion			Generous Inclusion		
	Collisions	Relative Collision Rate	Risk Ratio	Collisions	Relative Collision Rate	Risk Ratio
Hillegass-Bowditch	5	0.05		5	0.05	
Telegraph	21	0.18	3.9	26	0.23	4.9
College	30			33		
Milvia	26	0.14		33	0.18	
Shattuck	47	0.25	1.8	66	0.36	2.0
MLK	18	0.40	2.9	27	0.60	3.4
California	9	0.23		12	0.30	
Sacramento	15			22		
Ninth	3	0.08		6	0.16	
San Pablo	26	0.60	7.3	34	0.79	4.8
Virginia	6	0.10		10	0.16	
University	17	0.59	6.0	25	0.87	5.3
Channing	19	0.22		28	0.32	
Dwight	11			23		
Russell	7	0.11		14	0.22	
Ashby	26	0.88	8.0	40	1.35	6.1

Bicycle-motor vehicle collisions vs. all crashes

I made the decision to include only bicycle-motor vehicle collisions in my study because I had formulated my hypothesis based on evidence that motor vehicle traffic characteristics have an important impact on collision rates and severity for cyclists. Bicycle-motor vehicle collisions do account for 81% (1715 of 2114) of Berkeley's reported bicycle collisions for the time period of study, but the dataset includes three other types of crashes as well: 51 bicycle-pedestrian collisions, 16 bicycle-bicycle collisions and 332 one-party bicycle crashes, which I term "falls" though some of them may have involved a fixed object. These crashes represent part of the full risk profile for cyclists as well, so it is important to examine them and make sure that my conclusions about safety are not undone by these other types of crashes. Table 27 compares bicycle-motor vehicle collisions alone (the same figures used throughout this paper) with "all collisions" (bicycle-motor vehicle, bicycle-pedestrian and bicycle-bicycle) and "all crashes" (the above plus falls).

Table 27. Comparison of collisions, collison rates and risk ratios for Berkeley for different sets of crashes

Street	Bicycle-motor vehicle collisions only			All collisions			All crashes		
	Collisions	Relative Collision Rate	Risk Ratio	Collisions	Relative Collision Rate	Risk Ratio	Collisions	Relative Collision Rate	Risk Ratio
Hillegass-Bowditch	5	0.05		5	0.05		7	0.07	
Telegraph	21	0.18	3.9	23	0.20	4.3	24	0.21	3.2
College	30			31			32		
Milvia	26	0.14		27	0.14		33	0.18	
Shattuck	47	0.25	1.8	49	0.26	1.8	50	0.27	1.5
MLK	18	0.40	2.9	19	0.42	3.0	21	0.47	2.7
California	9	0.23		9	0.23		10	0.25	
Sacramento	15			15			15		
Ninth	3	0.08		3	0.08		4	0.11	
San Pablo	26	0.60	7.3	26	0.60	7.3	26	0.60	5.5
Virginia	6	0.10		6	0.10		7	0.11	
University	17	0.59	6.0	19	0.66	6.7	23	0.80	7.0
Channing	19	0.22		19	0.22		21	0.24	
Dwight	11			11			15		
Russell	7	0.11		7	0.11		12	0.19	
Ashby	26	0.88	8.0	26	0.88	8.0	28	0.95	5.0

The results for all collisions and for all crashes are not substantially different from those for bicycle-motor vehicle collisions alone. Including all crashes does narrow some risk ratios a bit, leaving

Shattuck with just 1.5 times the risk of Milvia. Re-running the Poisson regression with all crashes for Milvia v. Shattuck gives a p-value of .054, just outside the .05 level of statistical significance. All of the other street pairs are still significant at the .01 level or better.

I examined the three types of non-motor vehicle crashes to look for any interesting patterns. Of the 16 bicycle-bicycle crashes, under conservative inclusion criteria, none occurred on arterials or bicycle boulevards. Of the 51 bicycle-pedestrian crashes, under conservative inclusion and only counting those since 2003, eight occurred on arterials and just one on a bicycle boulevard. This makes sense, since many of the arterials are commercial, and so have more pedestrians walking around than the bicycle boulevards. Moreover, sidewalk riding takes place on arterials (Anderson, personal email, 2010), perhaps because cyclists are intimidated off the road by the heavier traffic and lack of bicycle facilities. Of the 332 falls, under conservative inclusion and since 2003, 14 occurred on arterials and 16 on bicycle boulevards. It is unclear what, if anything, the occurrence of falls might have to do with street typology.

Interestingly, though, reported falls are more likely to be severe than reported bicycle-motor vehicle collisions, as shown in Table 28. The p value column shows the results of a difference of proportions test for each crash type's proportion severe versus bicycle-motor vehicle collisions, the reference category. Bicycle-bicycle crashes and bicycle-pedestrian crashes are not distinguishable from bicycle-motor vehicle collisions in severity rate, but falls are about twice as likely to be severe, and the difference is significant at the .01 level. This may just reflect a difference in the threshold of severity needed to report an incident. Indeed, it is hard for me to imagine why any cyclist would bother to report a fall, unless they were so severely hurt as to need emergency treatment, or perhaps wanted to report poor pavement conditions.

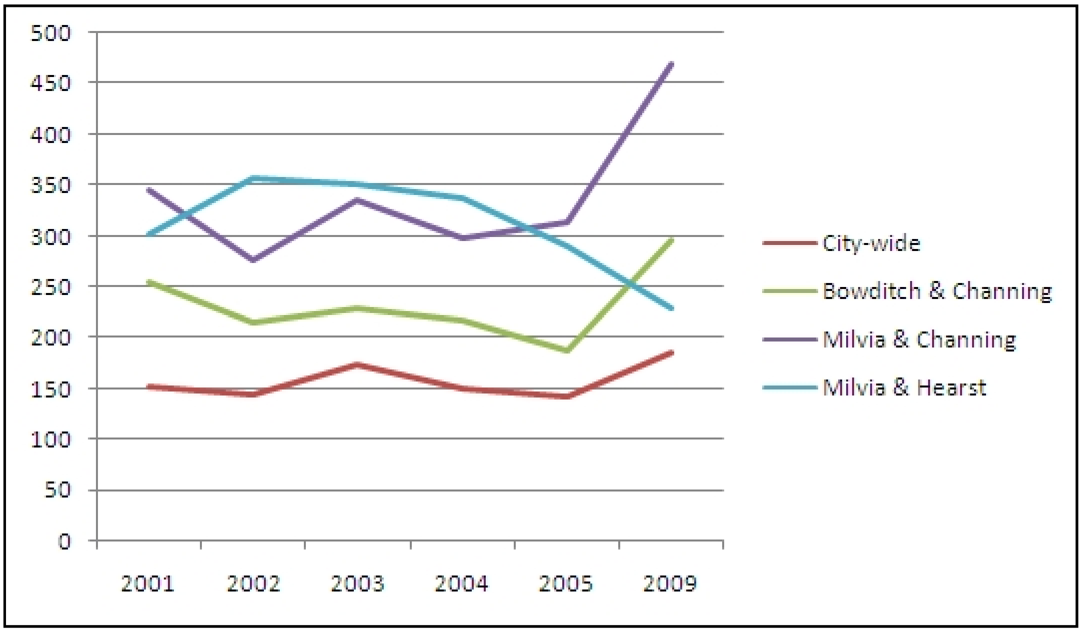
Table 28. Severity of different reported crash types in Berkeley, 1996-2008

Type of crash	Number of collisions	Number severe	Proportion severe	p value
Bicycle-motor vehicle	1715	71	4.1%	
Bicycle-bicycle	16	2	12.5%	.31
Bicycle-pedestrian	51	2	3.9%	.78
Falls	332	27	8.1%	.0029

Appendix 4. Effects of bicycle boulevard implementation on cyclist route choice in Berkeley

In Analysis and Results I touch on the question of whether the creation of bicycle boulevards has increased overall cycling in Berkeley, has shifted cycling from arterials to bicycle boulevards, or neither. Figure 51 below (identical to Figure 46 above) suggests that neither has occurred.

Figure 51. Change in Berkeley cyclist counts at selected intersections, 2001-2009

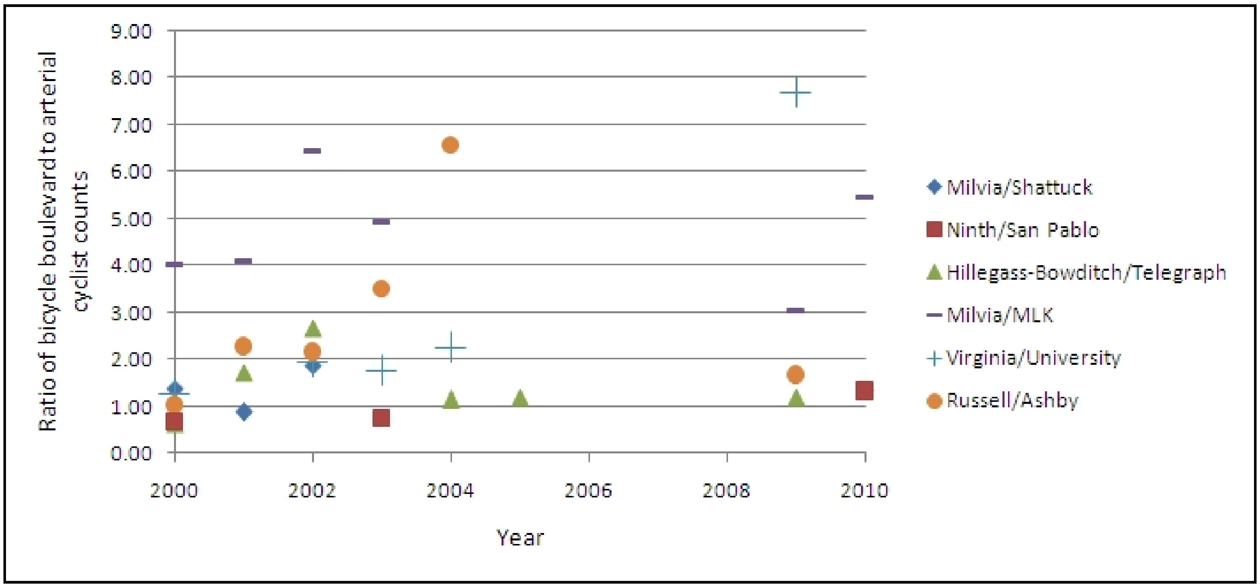


Source: Berkeley 2000-2009 Bicycle Counts – Chart Update II – Trend Line Data

The question of whether cyclist traffic has shifted from arterials to bicycle boulevards has important implications for the city-wide safety effects of the bicycle boulevard strategy as a whole, as discussed under Policy Implications, so it is worth examining a bit further. Unfortunately, as this appendix shows, data do not allow for any meaningful conclusions.

One approach is to calculate an exposure measure for each street for each year, and then divide the exposure on each bicycle boulevard by the exposure on the corresponding arterial and check whether that ratio is increasing over time. I computed exposure measures using average PM peak count(s) on each street in each year. The ratio of bicycle boulevard to arterial traffic is graphed in Figure 52. If bicycle traffic had been shifting from bicycle boulevards to arterials, one would expect a positive trend in the ratios.

Figure 52. Ratio of PM peak cyclist volume on Berkeley bicycle boulevards to arterials, by year, 2000-2010



Data points are sparse, and cannot even be reasonably plotted as lines since so many years are missing. Though the ratio appears to be rising for some street pairs, the primary impression the graph conveys is one of noise. This probably stems from a difference in the cross streets at which the counts were conducted, and perhaps from daily variations due to factors such as weather. For most years, each street's volume is based on just one or two observations.

A better approach is to examine paired, or nearly paired, counts, where parallel streets were counted at the same cross street and at least in the same season if not on the same day. Unfortunately, there are not many count data usable for this purpose. Only two street pairs had paired counts in more than two years. Counts and ratios are shown in Table 29.

Viewed this way, there is some indication of an upward trend in the ratios. This becomes more apparent when the same cross street is examined over years: Figure 53 graphs all of the count ratios measured at Channing St. from Table 29, and Milvia and Ninth both appear to have gained traffic relative to their respective arterials since 2000.

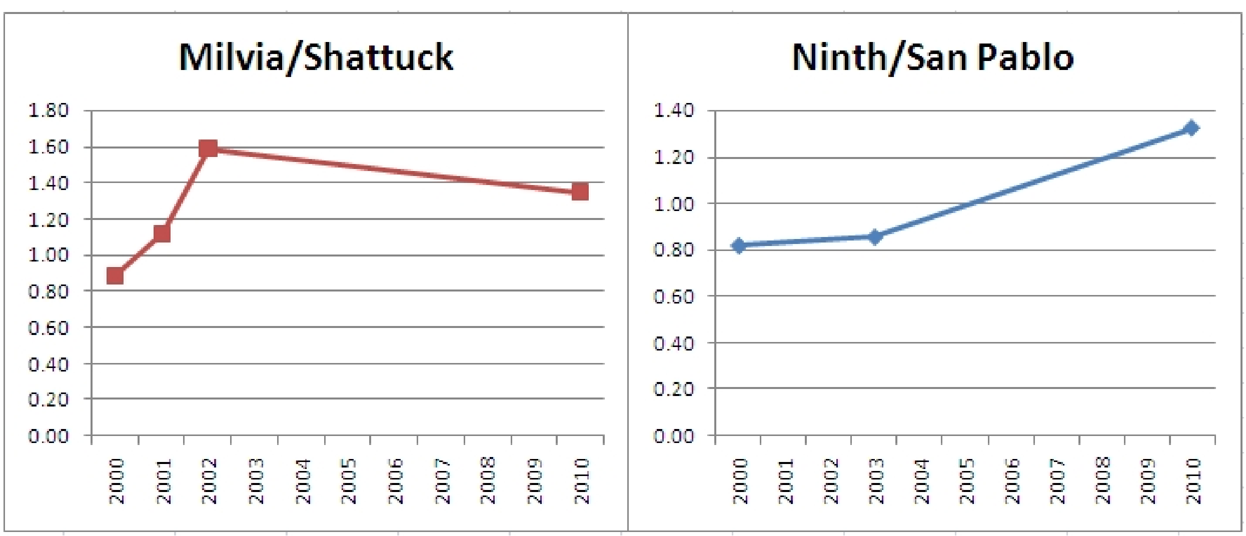
Table 29. Ratio of counts on selected bicycle boulevards to arterials in Berkeley, 2000 - 2010

Ninth v. San Pablo						
	Cross Street	**BBlvd Date**	**Arterial Date**	**BBlvd Vol**	**Arterial Vol**	**Ratio**
2000	Parker	2000/11/8	2000/11/8	16	52	0.31
	Channing	2000/11/14	2000/11/1	27	33	0.82
	University	2000/11/7	2000/11/7	29	46	0.63
2003	Parker	2003/5/27	2003/5/27	41.5	43.5	0.95
	Channing	2003/6/3	2003/6/3	32.5	38	0.86
	University	2003/5/29	2003/5/29	34.5	48	0.72
2010*	Channing	2010/3/25	2010/3/24	43	32.5	1.32

Milvia v. Shattuck						
	Cross Street	**BBlvd Date**	**Arterial Date**	**BBlvd Vol**	**Arterial Vol**	**Ratio**
2000	Channing	7/12/2000	7/12/2000	110	124	0.89
	Hearst	7/13/2000	7/13/2000	139	56	2.48
2001	Channing	10/9/2001	10/16/2001	212.5	190.5	1.12
2002	Channing	10/9/2002	10/18/2002	152	96	1.58
2010	Channing	3/22/2010	3/22/2010	249	185	1.35

*Counts in 2010 on Ninth and on San Pablo were one hour PM peak counts

Figure 53. Ratios of arterial to bicycle boulevard cyclist traffic, 2000-2010.



Still, even Figure 53 does not tell an extremely clear story, and is based on very few data. Also, a caveat is that the city assigned volumes in 2000 counts according to cyclist direction of approach, whereas my method for computing volume from 2001 onward involves counting turns as a half cyclist for each street, so the figures are not strictly comparable.

In sum, the analysis in this appendix can neither demonstrate that cyclist traffic has shifted somewhat from arterials to bicycle boulevards, nor can it rule out the possibility. Certainly, any shift that has occurred has not been dramatic in scale or speed.

Appendix 5. Expanded analysis of collision rates for Cambridge

Full data from the counts I conducted in Cambridge are given in Table 30. Not surprisingly, there is AM/PM variation in the volume on Green St., with more cyclists in the PM peak than AM, presumably because the one-way Green St. points away from downtown Boston (as well as MIT and Kendall Square).

Table 30. Complete data from 2010 cyclist counts in Cambridge

				Adjusted volumes		Ratio
Cross street	**Weather**	**Date**	**Time**	**Green**	**Mass**	**Mass/Green**
Pleasant	46F clear sunny	4/14 Wed	8:03a-9:03a	11	157	14.3
Pleasant	68F clear sunny	4/14 Wed	4:30p-5:30p	19	162	8.5
Sidney	52F clear windy	4/15 Thu	8:01a-9:02a	7.5	162	21.6
Sidney	59F clear windy	4/15 Thu	4:30p-5:29p	11	142	12.9
			Total	48.5	623	12.8
			Avg 2-hour	24.25	311.5	12.8
				Adjusted volumes		**Ratio**
Cross street	**Weather**	**Date**	**Time**	**Oxford**	**Mass**	**Mass/Oxford**
Wendell	50F clear sunny	4/20 Tue	8:04a-9:04a	91	141.5	1.6
Wendell	65F clear sunny	4/20 Tue	4:29p - 5:29p	68	128	1.9
			Total	159	269.5	1.7
			Avg 2-hour	159	269.5	1.7

Including all crashes, rather than only bicycle-motor vehicle collisions, in the analysis for Cambridge, makes little difference. It adds three bicycle-pedestrian collisions on Mass Ave. parallel to Green St. as well as one bicycle-pedestrian collision and one fall on Mass Ave. parallel to Oxford St. Collisions on Green St. and Oxford St. are unaffected. Therefore this analysis favors side streets more than only including bicycle-motor vehicle collisions, though the risk ratio for Mass Ave. vs. Green St. still rounds off to 1.3, as seen in Table 31.

Table 31. Collision rates for Cambridge streets including all types of crashes

Street	**Crashes**	**Exposure**	**Relative Collision Rate**	**Risk Ratio**
Green St.	3	24.3	0.12	
Mass Ave. parallel to Green St.	52	311.5	0.17	1.3
Oxford St.	6	159.0	0.04	
Mass Ave. parallel to Oxford St.	24	269.5	0.09	2.4

The fact that all four bicycle-pedestrian collisions added to this analysis occurred on Mass Ave. may be because, as a commercial street, it has plenty of pedestrians, but then again it also has more cyclists than the side streets observed here, and anyway with so few data points it is hard to draw any conclusions.

Bibliography

Allen, John S. 2003 Cambridge bike lanes: political statement or road improvement? http://www.bikexprt.com/massfacil/cambridge/lanes.htm (accessed May 16, 2010).

Allen-Munley, Cheryl, Janice Daniel and Sunil Dhar. 2004. Urban bicycle route safety rating logistic model. Transportation Research Record 1878: 107-115.

Alta Planning and Design. 2009. National Bicycle and Pedestrian Documentation Project: Automatic Count Technologies. http://bikepeddocumentation.org/index.php/download_file/-/view/22

Americawalks.org. 2008. Prevention/APMA 2008 list of Best Walking Cities http://www.americawalks.org/2008/03/walking-cities/ (accessed November 29, 2009)

Anderson, Eric. Associate Planner, City of Berkeley Department of Public Works. Personal email, April 19, 2010.

Bandyk, Matthew. 2009. 15 Cities for People Who Hate Driving and Long Commutes. *Yahoo! Real Estate,* November 12. http://realestate.yahoo.com/promo/15-cities-for-people-who-hate-driving-and-long-commutes.html (accessed November 29, 2009)

Bay Area Census. City of Berkeley, Alameda County http://www.bayareacensus.ca.gov/cities/Berkeley.htm (accessed April 1, 2010)

Broach, Joseph. 2008. B-SMaRT ~ Bike Safety Monitoring and Reporting Tool http://bikeportland.org/closecall/home.php (accessed April 14,2010)

Cambridgebikes.org. Cambridge Bike Route Network http://cambridgebikes.org/Network/ (accessed April 12, 2010)

Cameron, A. Colin and Pravin K. Trivedi. 1998. *Regression Analysis of Count Data*. Online Edition <http://cco.cambridge.org/uid=1895/extract?id=ccol0521632013_CCOL0521632013_root> Cambridge, UK: Cambridge University Press.

Carrivick, Philip J.W., Andy H. Lee, Kelvin K. W. Yau. 2003. Zero-inflated Poisson modeling to evaluate occupational safety interventions. *Safety Science* 41, no. 1:53-63.

City of Berkeley. Bicycle Boulevard History http://www.ci.berkeley.ca.us/ContentDisplay.aspx?id=6748 (accessed April 1, 2010)

———. Bicycle Boulevard Network http://www.ci.berkeley.ca.us/ContentDisplay.aspx?id=6694 (accessed April 2, 2010)

———. Speed Limits http://www.ci.berkeley.ca.us/ContentDisplay.aspx?id=8244 (accessed May 15, 2010)

———. What is a Bicycle Boulevard? http://www.ci.berkeley.ca.us/ContentDisplay.aspx?id=6690 (accessed April 14, 2010)

———. Bicycle Boulevards: Current Projects. http://www.ci.berkeley.ca.us/ContentDisplay.aspx?id=6692 (accessed May 1, 2010)

———. Average Total Daily Traffic Volume. http://www.ci.berkeley.ca.us/uploadedFiles/Public_Works/Level_3_-_General/TrafficVolumeMajor_Collector.pdf (accessed May 9, 2010).

City of Cambridge Community Development Department. 2009. Bicycle Lanes. http://www.cambridgema.gov/~CDD/et/bike/bike_lanes.html (accessed April 17, 2010).

City of Cambridge Department of Traffic, Parking and Transportation. 2010. City of Cambridge, MA 24 Hour Average Daily Traffic Counts. http://www.cambridgema.gov/CityOfCambridge_Content/documents/tpat_On_Line_Traffic_Count_Report.pdf (accessed May 15, 2010)

Decima Research, Inc. 2000. City of Toronto 1999 Cycling Study: Final Report on Quantitative Research Results. http://www.toronto.ca/cycling/reports/pdf/decimareport.pdf

Dill, Jennifer and Theresa Carr. 2003. Bicycle Commuting and Facilities in Major U.S. Cities: If You Build Them, Commuters Will Use Them – Another Look. *Transportation Research Record* 1828: 116-123.

Elvik, Rune and Truls Vaa. 2004. *The Handbook of Road Safety Measures*. Oxford, UK: Elsevier Ltd.

Forester, John. 1993. *Effective Cycling*. 6th ed. Cambridge, MA: MIT Press.

Hauer, Ezra. 2001. Computing and interpreting accident rates for vehicle types or driver groups. *Transportation Research Record* 1476:69-73.

Heron, Melonie, Donna L. Hoyert, Sherry L. Murphy, Jiaquan Xu, Kenneth D. Kochanek and Betzaida Tejada-Vera. Deaths: Final Data for 2006. Centers for Disease Control and Prevention (CDC) *National Vital Statistics Reports (NVSS)* 57, no. 14.

Jacobsen, PL. 2003. Safety in numbers: more walkers and bicyclists, safer walking and bicycling. *Injury Prevention* 9:205-209.

Jensen, Soren Underlien. 2008. Bicycle tracks and lanes: a before-after study. Paper presented at the Transportation Research Board 87th Annual Meeting, January 13-17, Washington D.C.

———. Personal email. November 10, 2009.

Kaplan, Jerrold A. 1975. Characteristics of the regular adult bicycle user. M.S. Thesis, University of Maryland.

Kathleen Murphy. 2010. Sergeant, Cambridge Police Department. Personal interview, February 4.

Kim, Joon-Ki, Sungyop Kim, Gudmundur F. Ulfarsson and Luis A. Porrello. 2006. Bicyclist injury severities in bicycle–motor vehicle accidents. *Accident Analysis and Prevention* 39, no. 2: 238-251.

Klop, Jeremy R. and Asad J. Khattak. 1999. Factors Influencing Bicycle Crash Severity on Two-Lane, Undivided Roadways in North Carolina. *Transportation Research Record* 1674:78-85.

Lombardi, Kristen. 2002. A dangerous & now deadly bicycle policy: A host of good intentions, but a lack of common sense. *The Boston Phoenix*, August 1-8. http://www.bostonphoenix.com/boston/news_features/top/features/documents/02379848.htm (accessed May 1, 2010)

Marshall, Wesley Earl. 2009. Community Design, Road Safety & Transportation Sustainability. PhD Dissertation, Center for Transportation and Urban Planning, University of Connecticut.

Metropolitan Transportation Commission. MTC -- Planning -- Bicycles/Pedestrians -- Safety Toolbox: Problem identification. http://www.mtc.ca.gov/planning/bicyclespedestrians/safety/data.htm#switrs (accessed April 3, 2010)

Mionske, Bob. 2007. *Bicycling & the Law: Your Rights as a Cyclist*. Boulder, CO: Velo Press.

Moritz, William E. 1997. Survey of North American Bicycle Commuters: Design and Aggregate Results. *Transportation Research Record* 1578:91-101.

———. 1998. Adult Bicyclists in the U.S. Paper presented at the Transportation Research Board 77th Annual Meeting, January 11-15, Washington D.C.

Nelson, Arthur C. and David Allen. 1997. If You Build Them, Commuters Will Use Them: Association Between Bicycle Facilities and Bicycle Commuting. *Transportation Research Record* 1578:79-83.

Ng, H.K.T., K. Gu, M.L. Tang. 2005. A comparative study of tests for the difference of two Poisson means. *Computational Statistics & Data Analysis* 51, no. 6:3085-3099.

Noland, Robert B. 1995. Perceived Risk and Modal Choice: Risk Compensation in Transportation Systems. *Accident Analysis and Prevention* 27, no. 4:503-521.

Parenti, Jeff. Principal Traffic Engineer, City of Cambridge Department of Traffic, Parking and Transportation. 2006. Detailed Bicycle Crash Reports in Cambridge. Presented to the Cambridge Bicycle Committtee, January 11.

Pucher, John and Lewis Dijkstra. 2000. Making walking and cycling safer: lessons from Europe. *Transportation Quarterly* 54, no. 3:25-50.

Pucher, John. 2001. Cycling Safety on Bikeways vs. Roads. *Transportation Quarterly* 55, no. 4:9-11.

R Documentation. R: Likelihood ratio test for over-dispersion in count data. http://bm2.genes.nig.ac.jp/RGM2/R_current/library/pscl/man/odTest.html (accessed May 9, 2010)

Reynolds, Conor C. O., M. Anne Harris, Kay Teschke, Peter A. Cripton and Meghan Winters. 2009. The impact of transportation infrastructure on bicycling injuries and crashes: a review of the literature. *Environmental Health* 8, no. 47. http://www.ehjournal.net/content/8/1/47

Robinson, Dorothy L.,. 2005. Safety in numbers in Australia: more walkers and bicyclists, safer walking and bicycling. *Health Promotion Journal of Australia* 16, no. 1.

Rodgers, Gregory B. 1997. Factors associated with the crash risk of adult bicyclists. *Journal of Safety Research* 28, no. 4:233-241.

Seiderman, Cara. City of Cambridge Bicycle Coordinator. Personal email, April 13, 2010.

Smith, Robert and Thomas Walsh. 1988. Safety impacts of bicycle lanes. *Transportation Research Record* 1168:49-56.

Tinsworth, Deborah K., Suzanne P. Cassidy and Curtis Polen. 1994. Bicycle-related injuries; Injury, hazard and risk patterns. *International Journal for Consumer Safety* 1, no. 4: 207-220.

Turner, S.A., A.P. Roozenburg, T. Francis. 2006. Predicting Accident Rates for Cyclists and Pedestrians. *Land Transport New Zealand Research Report* 289.

United States Census. 2000. American FactFinder http://factfinder.census.gov/home/saff/main.html?_lang=en (accessed May 1, 2010)

Van Houten, Ron and Cara Seiderman. 2005. How Pavement Markings Influence Bicycle and Motor Vehicle Positioning: A Case Study in Cambridge, MA. Paper presented at the Transportation Research Board 84th Annual Meeting, January 9-13, Washington D.C.

Walker, Lindsay, Mike Tressider and Mia Birk. 2009. Fundamentals of bicycle boulevard planning & design. Prepared for the Portland State University Initiative for Bicycle and Pedestrian Innovation.

Weigand, Lynn. 2008. A Review of Literature: Intersection Treatments to Improve Bicycle Access and Safety. Prepared for the Initiative for Bicycle and Pedestrian Innovation, Center for Transportation Studies, Portland State University.

Wilbur Smith Associates (WSA). 1998. Berkeley Bicycle Plan: Draft for Inclusion in the General Plan. Prepared for the City of Berkeley. http://www.ci.berkeley.ca.us/uploadedFiles/Public_Works/Level_3_-_Transportation/berkeley%20bike%20plan.pdf

Winters, Meghan, Gavin Davidson, Diana Kao, Kay Teschke. 2010. Motivators and deterrents of bicycling: factors influencing decisions to ride. (submitted)

Winters, Meghan, Melissa C. Friesen, Mieke Koehoorn, Kay Teschke. 2007. Utilitarian Bicycling: A Multilevel Analysis of Climate and Personal Influences. *American Journal of Preventative Medicine* 32, no. 1:52-58.

All photos by author.