# Bicyclist Behavior in San Francisco: A Before-and-After Study of the Impact of Infrastructure Investments

August 2016

A Research Report from the National Center for Sustainable Transportation

Dillon Fitch, University of California, Davis
Calvin Thigpen, University of California, Davis
Antonio Cruz, University of California, Davis
Susan Handy, University of California, Davis

National Center for Sustainable Transportation

ITS UCDAVIS INSTITUTE OF TRANSPORTATION STUDIES

## About the National Center for Sustainable Transportation

The National Center for Sustainable Transportation is a consortium of leading universities committed to advancing an environmentally sustainable transportation system through cutting-edge research, direct policy engagement, and education of our future leaders. Consortium members include: University of California, Davis; University of California, Riverside; University of Southern California; California State University, Long Beach; Georgia Institute of Technology; and University of Vermont. More information can be found at: ncst.ucdavis.edu.

## Disclaimer



## Acknowledgments

This study was funded by a grant from the National Center for Sustainable Transportation (NCST), supported by USDOT and Caltrans through the University Transportation Centers program. The authors would like to thank the NCST, USDOT, and Caltrans for their support of university-based research in transportation, and especially for the funding provided in support of this project.

# Bicyclist Behavior in San Francisco: A Before-and-After Study of the Impact of Infrastructure Investments

A National Center for Sustainable Transportation Research Report

August 2016

Dillon T. Fitch, Institute of Transportation Studies, University of California, Davis

Calvin G. Thigpen, Institute of Transportation Studies, University of California, Davis

Antonio N. Cruz, Environmental Science and Policy, University of California, Davis

Susan L. Handy, Institute of Transportation Studies & Environmental Science and Policy, University of California, Davis

NCST

[page left intentionally blank]

# TABLE OF CONTENTS

# Bicyclist Behavior in San Francisco: A Before-and-After Study of the Impact of Infrastructure Investments

## EXECUTIVE SUMMARY

This study explores bicyclist behavior in San Francisco using data collected before and after major bike infrastructure investments. From early 2011 to December 2013, investments of $3.3 million correlated with a 14% increase in counts of bicyclists, part of a 96% increase in bicyclist counts from 2006 to 2013 (San Francisco Municipal Transportation Agency, 2013a). To better understand the relationship between these investments and changes in bicycling behavior, we build on the successful GPS travel survey conducted in 2010 by the San Francisco County Transportation Authority (SFCTA) (Hood et al., 2011). We used data from the smartphone-based GPS data collection method (the CycleTracks application developed by SFCTA) which records bicyclists' routes. In addition, we administered a detailed web-based survey to CycleTracks users in order to better characterize the factors associated with their bicycling behavior. We examine the relationship between bicycle infrastructure and behavior of bicyclists so as to assess the effectiveness of existing investments, and to provide guidance on efforts that are effective at increasing bicycling.

## Introduction

There is growing evidence that bike-specific infrastructure positively affects the choice to bicycle as well as the routing decisions of bicyclists (Buehler and Pucher, 2012a; Furth, 2012; Heinen et al., 2010; Krizek, 2006; Winters et al., 2011). However, at the local level, the relationship between the type and placement of infrastructure and bicycling behavior varies. Many major cities in California have goals of increasing their bike mode share as a way to increase the sustainability of travel in their city. From a planning perspective, many questions emerge: Should particular groups of people be targeted for increasing bicycling (e.g. women, existing bicyclists, non-bicyclists)? Where should communities construct new bike infrastructure and what should be the focus of the infrastructure (e.g. connection to commercial services, schools, workplaces, etc.)? What type of infrastructure should communities build (e.g. bike lanes, off-street paths, sharrows, green pavements, bicycle parking, etc.)?

Any community looking to increase their bicycle mode share will need to answer these questions as they decide to plan for bicycling as a viable travel alternative. However, responding to the above questions is difficult in the absence of an understanding of the outcomes of infrastructure type and placement on people's travel behavior. The goals of this research are twofold:

(1) To examine the relationship between bicycle infrastructure and behavior of bicyclists so as to assess the effectiveness of existing investments.
(2) To provide valuable knowledge to governments and the public to use to implement strategic planning efforts that are effective at promoting sustainable travel.

While many advances in automobile technology and low-carbon fuels continue to decrease the environmental impacts of our transportation system (Madsen et al., 2010; Sperling and Gordon, 2008), efforts are needed to transform travel behavior in order to achieve a sustainable transportation system. This need has been particularly stressed in California with the passage of SB375 and the greenhouse gas (GHG) reduction targets set by the Air Resources Board (ARB, 2011). One key way travel behavior can be more sustainable is by shifting vehicle miles traveled to bicycle miles traveled. Evidence from national travel surveys shows that in many cases people make multiple trips of short length on a daily basis, especially for school, shopping, and personal errands (FHWA, 2009). These trips have the greatest potential to be taken by bicycle.

Bicycling is a low-impact mode of travel not only has the potential to reduce air pollution, congestion, GHGs, and noise pollution, but it also has tremendous potential for increasing the public health of communities (Garrard et al., 2012). Understanding the travel behavior of bicyclists is an important element of planning and designing sustainable communities. With growing attention to bicycle planning in many cities, the need for research to inform and direct these efforts is also growing. It has become generally accepted that the success of efforts to increase bicycling is not dependent on one strategy or measure but instead on an integrated approach incorporating both infrastructure and programs (Pucher et al., 2010). However, because infrastructure investments are often the most costly element of any integrated approach, they merit the most scrutiny. The following research aims to add to existing knowledge around the relationship between bicycle infrastructure investments and bicycle travel behavior, and to provide direct policy guidance for future investments.

San Francisco serves as an excellent case study to examine the impacts of bicycle infrastructure investments given the quirks of its recent bicycling history and its unique prioritization of non-automobile modes as part of its Transit-First Policy. In 2006, San Francisco's Bicycle Plan was served a court injunction as part of a CEQA (California Environmental Quality Act) challenge of the plan's environmental review (Superior Court of California - County of San Francisco, 2006), resulting in a dormancy in bicycling infrastructure investments until 2009, when the injunction was lifted (San Francisco Municipal Transportation Agency Bicycle Program Staff, 2010). The San Francisco Municipal Transportation Agency (SFMTA) thereafter rapidly made a variety of bicycle infrastructure investments across the city (San Francisco Municipal Transportation Agency, 2012). Simultaneously, in 2009 the San Francisco Country Transportation Authority (SFCTA) developed a smartphone application called "CycleTracks", which allowed users to record and upload their bicycle rides for SFCTA's use in improving their travel demand forecasting model. San Francisco bicyclists have continued to use the CycleTracks application into the present day, providing a steady stream of bicycle route data to the city.

This fortuitous confluence of events helped engender this study. We compare CycleTracks users' routes from late 2009 and early 2010, at the tail end of the bicycle infrastructure stasis, with routes in late 2013 and early 2014, allowing us to evaluate the influence and impact of the flurry of infrastructure investments made by San Francisco in the intervening years. We also attempt to control for two concurrent changes: the cohort of CycleTracks users has almost completely turned over between 2009 and 2013 and overall bicycling volumes in San Francisco have increased steadily since counts began in 2006, rising to 206% over 2006 levels in the most recently published counts in 2014 (San Francisco Municipal Transportation Agency, 2015).

This section provides a review of both the substantive and methodological elements featured in this report. We first discuss the different types of bicycle infrastructure and the existing evaluations of those infrastructure types before describing existing and planned bicycle infrastructure in San Francisco. Subsequently, we provide an overview of both the before-and-after research design we employ, and the data collection methodology we apply.

## Bicycle Infrastructure Types

Traditional American bicycle infrastructure classification features three types - bicycle routes, lanes, and paths. We briefly describe the sub-types within each classification.

### *Bicycle routes*

These facilities are shared, on-road bicycle infrastructure. Typically, the cost of implementing facilities for bicycle routes is more affordable than the more intensive work needed to construct bicycle lanes or bicycle paths (Bushell et al., 2013).


**Figure 1. A bicycle route marked by a sharrow (photo credit: Eric Fischer)**

***Sharrows***, also known as shared lane markings, are “road markings used to indicate a shared lane environment for bicycles and automobiles” that seek to “recommend proper bicyclist positioning” and may also “be configured to offer directional and wayfinding guidance” (see Figure 1) (National Association of City Transportation Officials, 2014). Though sharrows are perhaps the most affordable method of designating a street as a bikeway (Bushell et al., 2013), recent research indicates that sharrows are inferior both in attracting bicycle ridership and in increasing bicyclist safety (Ferenchak and Marshall, 2016).

**Figure 2. An example of infrastructure used to create a bicycle boulevard (photo credit: Richard Masoner)**

***Bicycle boulevards*** are "streets with low motorized traffic volumes and speeds, designated and designed to give bicycle travel priority" which "use signs, pavement markings, and speed and volume management measures to discourage through trips by motor vehicles and create safe, convenient bicycle crossings of busy arterial streets" (see Figure 2) (National Association of City Transportation Officials, 2014). As a relatively affordable bicycle facility option (Bushell et al., 2013, p. 12), some research has found that bicycle boulevards nonetheless can be more attractive to bicyclists than striped bicycle lanes (Broach et al., 2012; Winters and Teschke, 2010).

### *Bicycle lanes*

A variety of bicycle lane designs have been implemented. Though all bicycle lanes are located on-street, they vary by degree of separation from vehicular traffic and direction with respect to vehicular traffic flows.

**Figure 3. An example of a conventional bicycle lane**

***Conventional bicycle lanes*** "designate an exclusive space for bicyclists through the use of pavement markings and signage" (National Association of City Transportation Officials, 2014). Stated preference research indicates that bicyclists are more comfortable on, and would deviate substantially to ride on, a route with bicycle lanes either with or without parking, than on a route on unsigned streets with or without parking (McNeil et al., 2014; Tilahun et al., 2007). Aggregate studies also support the value of conventional bicycle lanes to bicyclists. One study found that every additional 1 mile of bicycle lanes per square mile was associated with a 1% greater bicycle mode share (Dill and Carr, 2003), while another found that bicycle lane supply per capita was associated more closely with increased bicycle mode share than was off-street bicycle path supply (Buehler and Pucher, 2012b). The latter result was echoed in the first round of route choice modeling using the San Francisco CycleTracks data, which also found that bicyclists had greater affinity for bicycle lanes than for bicycle paths (Hood et al., 2011). While they are shown to be effective magnets for bicyclists, bicycle lanes cost substantially more to implement, from $5,000 to $133,000, than bicycle routes (Bushell et al., 2013).

***Buffered bicycle lanes*** are "bicycle lanes paired with a designated buffer space separating the bicycle lane from the adjacent motor vehicle travel lane and/or parking lane" (National Association of City Transportation Officials, 2014). Evidence from intercept surveys in five US cities demonstrates that bicyclists are more comfortable on buffered bicycle lanes than conventional bicycle lanes (McNeil et al., 2014).

**Figure 4. An example of a buffered bicycle lane in San Francisco**

***Contraflow bicycle lanes*** are "bicycle lanes designed to allow bicyclists to ride in the opposite direction of motor vehicle traffic," turning "a one-way traffic street into a two-way street: one direction for motor vehicles and bikes, and the other for bikes only" (National Association of City Transportation Officials, 2014). Limited evaluations have been made of contraflow bicycle lanes, perhaps due to their relative rarity in the field. Of the many facilities evaluated in the five-city study by McNeil and his colleagues, two contained contraflow elements. These contraflow lanes, which were also buffered by soft-hit post pylons, did not receive substantially different ratings than their with-traffic peers (McNeil et al., 2014).


**Figure 5. An example of a contraflow bicycle lane (photo credit: Eric Fischer)**

***Physically separated bicycle lanes***, or ***cycle tracks***, marry “the user experience of a separated path with the on-street infrastructure of a conventional bike lane,” as they are “physically separated from motor traffic and distinct from the sidewalk” (National Association of City Transportation Officials, 2014). A strong and growing body of evidence shows that cycle tracks are indeed perceived as safer and more comfortable than the previously-mentioned flavors of bicycle lanes. Respondents in the five-city study rated physically-separated bicycle lanes well above buffered bicycle lanes and substantially above conventional bicycle lanes (McNeil et al., 2014). Cost estimates for cycle tracks range from $170,000 per mile on the low end (Greenfield, 2012) to $17.6 million per mile at the upper bound (Mobility Investment Priorities, 2014), an increase in several orders of magnitude over the cost of conventional bicycle lanes.


**Figure 6. An example of a physically separated bicycle lane in San Francisco**

### *Bicycle paths*

***Bicycle paths*** are off-street facilities, usually accommodating bicyclists as well as other non-motorized modes, that are typically eight feet or more in width (Bushell et al., 2013). While some studies indicate bicyclist preference for bicycle paths over conventional bicycle lanes (Tilahun et al., 2007), an ecological study shows that bicycle paths are less strongly associated with bicycle mode share than bicycle lanes (Buehler and Pucher, 2012b) and the San Francisco CycleTracks route choice model shows that while bicyclists do prefer routes with bicycle paths, they prefer bicycle lanes even more strongly (Hood et al., 2011). Bicycle paths cost between $65,000 per mile to $4.3 million per mile to implement (Bushell et al., 2013).

## San Francisco Bicycle Infrastructure

Thanks to its unique Transit-First Policy, a guideline that seeks to prioritize transit, bicyclist, and pedestrian modes (City and County of San Francisco, 2007), San Francisco's bicycle network in 2009 reached an extent of 208 miles. Of those 208 miles, 23 were separated bicycle paths, 45 were conventional bicycle lanes, and 132 were bicycle routes.

However, between 2006 and 2009, bicycle infrastructure in San Francisco languished. After the San Francisco Board of Supervisors unanimously passed the 2005 San Francisco Bicycle Plan (City and County of San Francisco, 2005), a judge placed an injunction on the plan due to a CEQA challenge of the plan's environmental review (Superior Court of California - County of San Francisco, 2006). Ultimately, the San Francisco Bicycle Plan was passed in August 2009, prior to a modification of the injunction in November 2009 (San Francisco Municipal Transportation Agency Bicycle Program Staff, 2010) and the full-scale resumption of infrastructure projects in August 2010 after the injunction was fully lifted (Gordon and Tucker, 2010).

Thereafter, implementation of bicycle infrastructure was swift: within six months after the August 2010 ruling, SFMTA had installed new bicycle lanes on 11 miles of city streets (James, 2011). And by the end of 2012, 20 miles of bicycle lanes and 41 miles of sharrows were installed over 2009 levels, increases of 45% and 178%, respectively (San Francisco Municipal Transportation Agency, 2012).

### *San Francisco Bicycle Plan*

The San Francisco Bicycle Plan (SFBP), developed in 2009 (San Francisco Municipal Transportation Agency, 2009), contains several elements: a list of the organizations and individuals that helped craft the plan, an introduction that helps to motivate the specific goals, and chapters devoted to each goal, describing in detail how the city plans to achieve those goals through specific action items. In particular, San Francisco's goals include improving bicycle infrastructure, including the bicycle network, bicycle parking, and transit and bridge access, as well as enhancing education, law enforcement, and promotion efforts. The plan lists how the city will enact these policies in accordance with the city's General Plan and the environmental review process and how it will secure funding for the programs listed in the plan.

San Francisco developed the SFBP for several reasons. Under the guidance of its Transit-First Policy, San Francisco is interested in promoting bicycling to reduce greenhouse gas emissions, improve the health of its residents, and achieve a host of other related goals. On a more practical level, the SFBP specifies how city agencies should coordinate on bicycle programs and the plan also contains all of the necessary elements to fulfill the requirements for funding from the state's Bicycle Transportation Account (California Department of Transportation, 2015).

The SFBP also coordinates with the San Francisco Municipal Transportation Agency's (SFMTA) Strategic Plan (San Francisco Municipal Transportation Agency, 2013b) and Bicycle Strategy (San Francisco Municipal Transportation Agency, 2013c) documents. The Strategic Plan helps define the agency's overarching goals, across all modes, for the next six years. For example, the Strategic Plan's Objective 2.3 targets a 50/50 split between single-occupancy vehicle (SOV) and non-SOV modes by 2018. Furthermore, the SFMTA Bicycle Strategy takes these high-level policies and applies them more directly to bicycling. In particular, the Bicycle Strategy document aims to help achieve the 50% non-automobile mode split, identified in the Strategic Plan, by "increas[ing] from 3.5% to 8-10% bike mode share". These strategies are then put into more concrete action items in the SFBP, which identifies (among other things) the specific bicycle infrastructure projects that the city plans to build.

## Research Designs to Evaluate Bicycle Infrastructure

### *Before-and-After Research Designs*

Studies show that different bicycle infrastructure types and placements affect behavior differently (Ibeas et al., 2013; Larsen et al., 2013). However, most studies of travel behavior and bike infrastructure are subject to various confounding factors that obscure causal relationships. One way to overcome this problem is to employ "before-and-after" methods in conjunction with data collection at a disaggregate level. In these quasi-experimental designs, control for the timing of the cause relative to the effect is naturally provided by surveying before and after the treatment, in this case the building of infrastructure. This is one of the methods that will be used to assess the relationship between bicyclist behavior and bike infrastructure in San Francisco.

This particular study employs two different before-and-after quasi-experimental designs. The first is a classic two group study in which the before group is the control group and the after group is the treatment group, and in which the two groups include different (but potentially overlapping) sets of individuals. This repeat cross-sectional design is complicated by the fact that where someone rides is also a determinant of whether a person is in the control or treatment group (i.e. if a person in the after period rides in an area where there is no new infrastructure nearby, they could be considered a control rather than a treatment observation).

The second research design is a within-subject design in which there is no traditionally-defined control group (because all subjects could be subject to the treatment), but in which individuals serve as their own control. This longitudinal design is executed by collecting repeated measures for many subjects (i.e. subjects that record 2 or more trips on different routes). However, some subjects only provided a route that would be considered a control (in the before condition), some only received a treatment (after), some both (internal control), and some recorded a route in the after period but are considered a control observation (see above).

Longitudinal data of bicyclists' routes has the potential to provide a richer explanation of the link between infrastructure and bicycling behavior because behavioral *changes* are observed at the individual level. However, distinguishing individual route variability from fundamental change can be difficult. We rely on both quasi-experimental designs to evaluate a set of alternative explanations of the changes in bicycling infrastructure use in San Francisco.

## Data Collection Methodology

### *Smartphone Applications*

In contrast to more traditional data collection methods such as surveys and travel diaries, the rise of smartphones with Global Positioning System (GPS) sensors has provided the potential for planners and academics to collect larger quantities of travel data with greater ease on the part of the participant. Smartphone applications that are designed to collect travel data generate much more route specific information than more traditional methods and with different data management needs and challenges.

### *SFCTA's CycleTracks Smartphone Application*

This study relies on the smartphone application "CycleTracks" for the collection of the core data: bicycle routes. This application was the first of its kind, designed by SFCTA in 2009 to collect bicycle route data for planning purposes. The development team focused on making the application free, easy to download and use, and energy efficient (so as to avoid draining the user's battery) (Charlton et al., 2011). This last item was particularly important, given the battery-intensive nature of GPS communications. CycleTracks notified users with a bicycle bell noise after the first 15 minutes of recording and every five minutes thereafter, and also instituted an automatic shutdown feature when a user's phone battery dipped below 10% charge remaining (Charlton et al., 2011).

In addition to collecting the GPS traces of users' bicycle routes, several additional pieces of information are requested in the CycleTracks application. It requests that users provide basic socio-demographic information, including age, email address, gender, home location ZIP code, work location ZIP code, school location ZIP code, and cycling frequency. At the end of each trip, users are asked to select a trip purpose from the following options: commute, school, work-related, exercise, social, shopping, errand, and other.

Using the first six months of data collected in San Francisco, Hood and his colleagues (Hood et al., 2011) estimated a bicyclist route choice model, which they then incorporated into the SFCTA travel demand model SF-CHAMP (Zorn et al., 2012). Since then, the open-source CycleTracks code has been used and modified in other regions, such as Atlanta and Reno.

### *Data Cleaning and Map Matching*

A recent report reviewed the state of the practice of collecting, cleaning, and analyzing GPS data for travel research (Wolf et al., 2014a, 2014b). Much of the recent literature focuses on the challenge of inferring travel behavior from GPS traces (e.g. trip purpose, trip origin and destination), to begin to replace traditional manual travel diaries with automatic diaries. However, for the purposes of using CycleTracks type data for assessing bicycling behavior, inferences about trip purpose are not necessary because users provide that type of information every time they record a trip. This makes using GPS data from CycleTracks much simpler for the analyst. The two main types of pre-processing that are needed for this data are general GPS cleaning (i.e. removal of noisy GPS points and traces), and map matching (i.e. determining which network links are used on a given trip).

Data cleaning algorithms are often simple rule-based screening algorithms based on GPS attributes such as position dilution of precision (PDOP) and instantaneous measures of speed. Such algorithms have been shown to be successful for traditional stand-alone GPS units in cars (Wolf et al., 2014a, 2014b). However, GPS data from smart phones often do not come with the same attributes as stand-alone GPS units (e.g. PDOP, number of satellites, etc), because positions are based on additional location data (e.g. wifi, cellular networks), and algorithms are proprietary. Therefore, alternative rules are used when satellite attributes are unavailable that only indirectly address the accuracy of positions. The rules are often based on context (e.g. removing points with vertical positions outside the study area's known elevation (Wolf et al., 2014a, 2014b)), or checks of realistic positions based on the time series of traces. In addition to screening algorithms, smoothing algorithms are also used to improve the accuracy of GPS positions (Wolf et al., 2014a, 2014b).

Map matching of GPS points to a given network has received considerable attention in the literature (Li et al., 2013; Schuessler and Axhausen, 2009; Yang et al., 2005). Algorithms vary in their methodology, but most attempt to identify deterministically the "true" path followed by the GPS trace. With large amounts of GPS data, analysts must often trade off accuracy and computational efficiency. The decision as to how to match GPS points to a network is based on the sparsity of GPS data, the complexity of the underlying network, and the ultimate use of the map matched data.

### Research Questions

In order to achieve the two main goals of this study (see Introduction), we focus on the following primary research questions:

(1) How do bicyclists' infrastructure use change over time?
   a. On what types of infrastructure do bicyclists ride? How much individual variability is there?
   b. Does infrastructure use vary by the socio-demographics and attitudes of riders?
   c. Does infrastructure use vary by trip purpose?

(2) Do changes in bicycling infrastructure cause changes in bicycling routes?

## Methodology

### Data Collection

Our data collection followed three main stages:

(1) Outreach (Fall 2013): This included passing out fliers for CycleTracks at coffee shops, bike shops, and transit areas. Also, outreach was conducted through various bike list serves and social media outlets (Facebook, Twitter, etc.).

(2) CycleTracks Data Collection (November 2013 – March 2014): Collection of CycleTracks data has been continuous from its creation. However, we will only use routes from the months following our outreach campaign in the fall that are seasonally matched to the original CycleTracks analysis. This process requires no additional work because SFCTA has built the CycleTracks system to be managed automatically.

(3) Online Survey (May2014): The online survey was created using SurveyGizmo.com's professional survey service. We constructed the survey, pre-tested the survey, administered the survey to all CycleTracks users through SFCTA, and extracted the data from SurveyGizmo.com.

## Data

This project utilized three main sources of data: (1) repeat cross sectional route choice data on a sample of bicyclists in San Francisco who have volunteered to download a smartphone application (CycleTracks) and record bike routes from either November 2009 – March 2010 (phase 1) , and/or November 2013 – March 2014 (phase 2); (2) an online survey of travel behavior and attitudes for a subsample of CycleTracks users; (3) a synthesis of existing GIS network information obtained from the SFCTA travel model network, the SFMTA bikeway layer, and extensive manual GIS review (see section below).

### *Network and Network Attributes*

We compiled bike network data with the aim of generating a database of routable network attributes thought to be representative of infrastructure investment in San Francisco from 2009-2014. These attributes included various categories of bike infrastructure (e.g. shared road markings, bike lanes, off-street paths, parking, etc.) by year of construction.

In order to generate a routable network with infrastructure information, we combined the geometry of SFCTA's network used in their 2009 travel demand model (this network includes links for bikes only such as off-street paths) with the infrastructure information from the SFMTA bikeway GIS layer. This data synthesis was largely manual because the two data layers were not topologically matched and because they did not have a unique tabular key between the two datasets. We supplemented this manual data synthesis by a review of Google Maps Streetview when inconsistencies between the layers were present, or when there were known errors due to personal knowledge of the study area. A detailed review of our data synthesis can be found in Appendix A – Detailed Data Cleaning Methods.

### *Network Data Limitations*

The most important limitation of our data synthesis is the lack of specific construction dates for past bike infrastructure. After extensive outreach with both SFCTA and SFMTA, among others, it was determined that no easily accessible digital record of construction date for bike infrastructure exists. This has implications for our analysis (as presented below), but importantly indicates that there is still a need for better local data management. We decided to rely on the year of installation for each bike facility as obtained from the SFMTA bikeway GIS layer.

The second limitation of our data synthesis is that we did not update the geometry of the SFCTA network to include any new links. There very well could have been small links which were either missing in the existing data, or were added after the final compilation of the 2009 SFCTA network. We expect any missing links to have only a minor consequence on route map matching given the exhaustive coverage of the 2009 SFCTA network (e.g. any bicyclist who used a road/path not represented in our network is likely to only have their route minimally detoured).

### *GPS Data Cleaning*

The data collected from the CycleTracks smartphone application included detailed bicyclist locations collected from the internal GPS of each participant's phone. The sampling frame for this study included all people who travel by bike within the county/city limits of San Francisco regularly and who possess an Android or iOS smartphone. The sample is not random, but was recruited through outreach efforts similar to the efforts in 2010 (Hood et al., 2011). It should be noted that although several opportunities for sample bias exist (e.g. smartphone users, cycling advocates), past comparisons with traditional travel diaries in the Bay Area (e.g. BATS) have indicated that the bias from CycleTracks samples is marginal (Sall, 2013). The original sample consisted of 696 participants with > 9,000 total recorded routes. Following the consolidation of GPS data, the following data cleaning algorithm, which was loosely based on Wolf et al. (2014a, 2014b), was used. This process, along with removal of trips for "exercise" or "other" resulted in a sample of 539 participants with 7,963 total utilitarian bike routes.

The cleaning algorithm excludes GPS points based on the following criteria (see Appendix A – Detailed Data Cleaning Methods for exact thresholds chosen based on the data context):

(1) Instantaneous speed was unreasonably large or negative
(2) Acceleration was unreasonably large
(3) Calculated speed from consecutive GPS points was unreasonably large
(4) Difference between instantaneous and calculated speed was unreasonably large
(5) Distance between consecutive GPS points was large (empirically justified by sporadic GPS positions within a short time interval)
(6) Horizontal accuracy of GPS was large

### *GPS Map Matching*

In order to determine which streets (and thus bicycle facilities) were used on each route, we needed to match GPS points to our GIS road network. GPS is accurate but imprecise, due to technology limitations and environmental conditions (e.g. satellite obstructions, nearby cell towers, etc.), while our GIS network is both accurate and precise, using graph theory to represent intersections and streets as nodes and links. We designed and implemented a new map-matching algorithm that automatically determines the links used for each trip given the GPS points and the network. The algorithm uses horizontal accuracy of GPS locations, distance between GPS locations and near network links, headings of consecutive GPS locations, heading of near links, and connectivity of trip routes (using network topology) to probabilistically weight links for their likely match to each GPS point. We sum the probabilities for each link, and generate a least cost path (based on maximizing link probability) to determine the final links used in each trip. The map-matching algorithm also acts as a secondary data cleaning procedure by removing trips which (1) have the same start and end location, (2) have very sparse GPS data, and (3) have an unreasonable trip distance (e.g. one block). This algorithm is only successful for destination-oriented travel (e.g. not round trips) because the least cost path between adjacent start and end locations is the adjacent link, not the traversed round trip path.

We implemented this algorithm in the R statistical computing language relying on several well-used R libraries. The current algorithm is in its first stage of development and is considerably inefficient. To process the nearly 8 million GPS points, it took roughly 80 hours on a 32 core processor. We expect that the speed of this algorithm could be improved by using alternative technology, or by rewriting the portions of the code that were not parallelized. However, because the goal of this project was the end analysis, we did not spend time improving the efficiency of the map-matching algorithm, and instead leave that task as potential future work.

The most important assumptions regarding the algorithm are listed below:

(1) Only the links within the horizontal accuracy of a GPS point are considered (we call this the link set for each GPS point). When no links within that horizontal accuracy exist, a search for links in consecutive 30 meter buffers is used until at least one link is found.

(2) The probability that a link in the link set for a GPS point is matched is:

$$\left(\frac{\frac{1}{d^2}}{\sum_i \frac{1}{d^2}}\right)\left(\frac{\frac{1}{(180\ -\ ||Hg - Hl|\ -\ 180\ |)}}{\sum_i \frac{1}{(180\ -\ ||Hg - Hl|\ -\ 180\ |)}}\right)$$

Where *d* is the Euclidean distance between GPS point and link, *Hg* is the GPS heading, *Hl* is the link heading, and subscript *i* is each link in the link set.

(3) Trips with the same start and end location, distance < 200 meters (as measured from the diagonal of the bounding box (envelope) of gps points), or fewer than 4 GPS points per km were discarded.

(4) If the combined link sets for each GPS do not result in a connected network (topographically correct), we split the trip into the two longest topographically correct stretches and generate two routes for that trip. Through visual inspection, we found this method to cover the spatial domain of the majority of trips without having to make assumptions where GPS data are sparse or have great positional uncertainty.

The result of the map matching algorithm is exemplified in Figure 7 which demonstrates the selection of network links associated with one bicyclist trip.

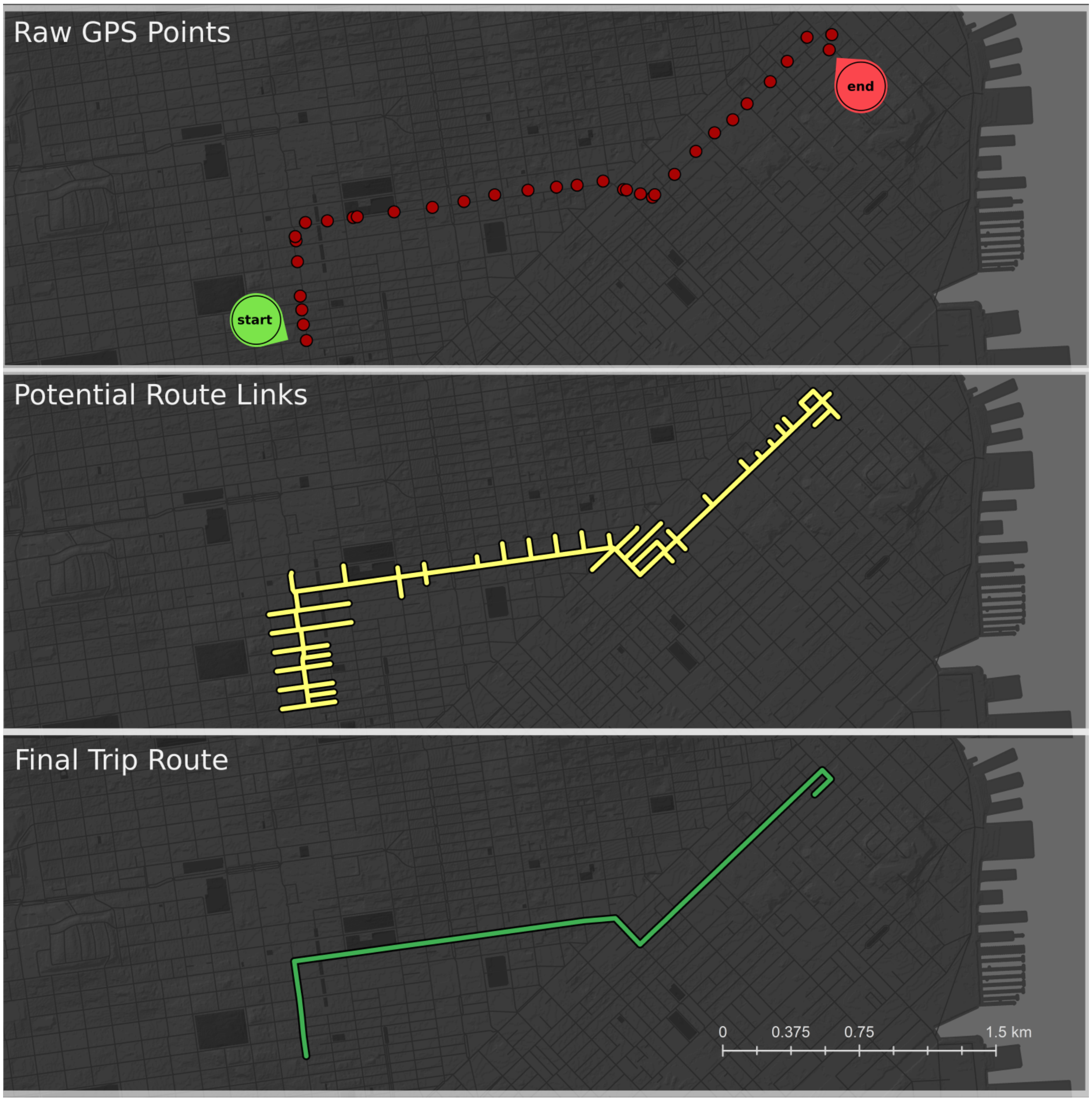


**Figure 7. Example of map matching algorithm for a single bicyclist's trip**

### *On-Line Survey*

We conducted an online survey of historical CycleTracks users from its inception (November 2009) through April 2014. The sampling frame for this survey included approximately 1,200 people (exact number unknown due to uncertainty in valid email addresses and SFCTA's handling of multiple email requests) who both recorded at least one route on CycleTracks and provided an email address to SFCTA. In coordination with SFCTA, we administered a travel survey using SurveyGizmo.com's online survey service. The survey included questions about socio-demographics, travel characteristics, as well as travel preferences and attitudes with particular emphasis on bicycling comfort and experience (see Appendix B). SFCTA anonymized the data by removing email addresses and in their place linked each survey response to a unique identifier that corresponded to a user id in the CycleTracks raw data. In this way, survey responses are associated with a subsample of CycleTracks data in our study. After removing duplicate survey responses (we assumed these were CycleTracks users who forwarded their individual specific web link to others) we had a total of 188 survey responses.

## Study Design

We use two quasi-experimental designs when analyzing the CycleTracks data (i.e. repeat cross sectional and longitudinal) to evaluate four alternative explanations of infrastructure use change. It is worth noting that we began this study with the hopes of a third design (a longitudinal before-and-after); however, only one CycleTracks user recorded routes in phase 1 and 2, thus we limit our analyses to the two designs as presented below.

### *Repeat Cross Sectional (Before and After)*

Two time periods were used to evaluate our four alternative hypotheses regarding the changing use bike specific infrastructure. The first time period, collected during the first phase of CycleTracks, was from November 2009 through March 2010. The second time period was four years later, from November 2013 through March 2014. The seasons were matched to attempt to control for weather characteristics which have been shown to affect the decision to walk and bicycle (Iacono et al., 2010). Between these two time periods, as stated above, the City of San Francisco made significant improvements in bicycling infrastructure. We did not attempt to study the effect of specific infrastructure through intercept surveys, instead our goal was to get a broad sample of bicyclists across the city to find differences in travel behavior between these two time periods, especially in regions where infrastructure improvements were made. This data subset includes only data collected directly from the CycleTracks app (i.e. GPS locations, trip purpose, gender)

### *Longitudinal*

This analysis focuses on the online survey participants and examines these individuals' infrastructure use over time. Because many of the survey participants have recorded many trips on CycleTracks, this subset of the data will be used to examine two of our alternative hypotheses regarding the influence of context as well as socio-demographics and attitudes on bicyclists' route behaviors. We compiled all the CycleTracks trips for the survey participants from 2009-2014 (including the period between phase 1 and 2) in order to expand our sample size. This data subset includes data from CycleTracks (i.e. GPS locations, trip purpose, gender) as well as data from the online survey (e.g. attitudes, comfort, income, race, etc.).

## Alternative Explanations

In addressing the main behavioral research question—"what is the relationship between bicyclist route behavior and bike infrastructure?"—we structure our results and discussion around alternative explanations for our findings with respect to changes in infrastructure use. These alternative explanations arise primarily because of the nature and limitations of the data collected. Our ability to draw strong inferences about the causal effect of infrastructure on bicycling from the cross-sectional and longitudinal subsets of our crowd-sourced data is limited. In general, we measure infrastructure-use as the share of bicycle travel using a given type of infrastructure. For example, a person recording one 10 mile route with 5 miles on no bike infrastructure, 2 miles on buffered bike lane, and 3 miles on "sharrowed" bike routes would have infrastructure-use shares of 50%, 20%, and 30% for those respective infrastructure types (see Results and Discussion for a detailed description of how these shares are calculated and aggregated for our samples). Using these type of infrastructure-use metrics over time, we consider four primary alternative explanations of bicyclist infrastructure-use change:

- Travel context explanation: same or different riders, with different trip characteristics (such as different origins and destinations and trip purpose), cause observed differences in infrastructure use (between phase 1 and 2, and/or over time).
- Attitudinal cohort explanation: different riders, with different attitudes about comfort and general route preferences, cause observed differences in infrastructure use (between phase 1 and 2, and/or over time).
- Targeted planning explanation: bicycle infrastructure is being installed where the same or similar bicyclists are already riding and thereby cause observed differences in infrastructure use (between phase 1 and 2, and/or over time).
- Route change explanation: same or similar riders switching their routes cause observed differences in infrastructure use (between phase 1 and 2, and/or over time).

Evidence for the last two explanations would be encouraging from a planning perspective, suggesting that the investments in the San Francisco bicycle network since the 2009 Bicycle Plan improved the comfort and safety of bicyclists. Evidence for the first two explanations would be more equivocal.

In addressing the second overall research question—"How can we evaluate the effectiveness of bicycle infrastructure investments including type, placement, scale, and magnitude?"—We draw on both the behavioral evidence from this study, and evidence of bicycling change from prior studies in San Francisco. We choose to address this question in the discussion only, as our empirical analyses only indirectly address the evaluation of infrastructure investments.

## Results and Discussion

We present our results in two formats. First, we summarize our sample population and some trip characteristics directly with univariate and bivariate statistics. Second, we summarize the majority of trip characteristics and infrastructure use through weighted summary univariate and bivariate statistics. We weight most of the statistics on trips and infrastructure use because of the wide range of data for each person (i.e. some people recorded only a single trip, some recorded multiple trips to and from similar origins-destination pairs, while others recorded multiple trips to and from multiple origin-destination pairs). Direct summary statistics will give much more weight to users with repeated trips. This can be particularly problematic for users who have recorded hundreds of commute trips on nearly the same route. In an attempt to balance the influence of heavy CycleTracks users who record repeated trips of similar characteristics, we weigh infrastructure-use based on the number of trips a person records to the same origin-destination neighborhoods[1]. The weighting scheme reduces the influence of people with repeated trips to and from the same neighborhoods, but at the same time gives more weight to people who record many routes to and from different neighborhoods. This is a balance between overweighting the behavior of heavy CycleTracks users with underweighting those same users who provide a diverse set of routes and thus diverse route behavior. In the following results and discussion, we refer to these metrics of infrastructure-use as "weighted" to distinguish them from raw aggregate infrastructure-use metrics.

## How do Bicyclists' Infrastructure Use Change Over Time?

### *Sample Characteristics*

Table 1 provides summary statistics of the phase 1 and 2 samples as well as the longitudinal sample of online survey participants. Although our phase 1 and 2 samples include different people (repeat cross-sectional design), their overall characteristics are relatively consistent: the two samples are predominantly male bicyclists who ride daily and for commute purposes. Our phase 2 sample has a slightly higher representation of female bicyclists than in phase 1, and the bicyclists in phase 2 are also somewhat more likely to record commute trips and report bicycling on a daily basis. However, phase 1 and 2 samples can be generally thought to exhibit characteristics consistent with the population of San Francisco bicyclists (Sall, 2013).

Of the sampling frame of CycleTracks users, those who chose to participate in our survey were more likely to be male, bicycle daily, and record commute trips than their prevalence in the phase 1 and phase 2 samples of CycleTracks users. These survey participants were also predominantly white, reported earnings below the San Francisco median income, felt uncomfortable on four-lane roads without a bicycle lane, and had 5 or more years of bicycling experience. In particular, our survey sample held remarkably similar perceptions regarding the comfort and safety as the participants in the interviews conducted for the 2011 San Francisco Bicycling Study Report, in which only 13% of the 1,063 respondents reported feeling safe from traffic on a bicycle and 19% reported feeling safe on roads with no designated bikeway (Corey Canapary & Galanis, 2011).

---

1

$$\forall\, n\colon \sum_i \sum_j \frac{1}{R_{nij}}$$

Where $R$ is a trip for person ($n$) from origin neighborhood ($i$) to destination neighborhood ($j$). If person $n$ has only a single trip from $i$ to $j$, their weight reduces to 1, otherwise it is their fraction of trips from $i$ to $j$.

**Table 1. Summary Statistics of CycleTracks Sample**

| | Phase 1 | | Phase 2 | | Survey Participants* | |
|---|---|---|---|---|---|---|
| *Total Sample Size* | | | | | | |
| *% Male* | 81.8% | n=363 | 76.6% | n=77 | 84.3% | n=89 |
| *% Daily Bicycling Frequency* | 59.2% | n=336 | 64.1% | n=78 | 64.5% | n=62 |
| *% Work Related Trips* | 65.9% | r=4,864 | 68.2% | r=1,498 | 77.6% | r=3,832 |
| *% White* | - | | - | | 78.6% | n=98 |
| *% Below S.F. Median Income* | - | | - | | 80.2% | n=86 |
| *% Comfortable on 4-Lane Road with No Bicycle Lane* | - | | - | | 13.5% | n=104 |
| *% > 5 years bicycling experience* | - | | - | | 92.4% | n = 92 |

*some survey participants are included in either the Phase 1 or 2 samples. “n” refers to the sample size by individual, while “r” refers to the sample size by trip.

## *Network and Infrastructure Use*

As noted in the introduction, the amount of bike specific infrastructure in the overall San Francisco network increased substantially between phase 1 and 2. We’ve summarized the distance of each type of infrastructure as a percentage of the overall network in Table 2. In general, we see the largest percentage point increases in conventional bike lanes and sharrows, with much smaller increases for innovative facilities (even though *percent changes* of innovative facilities are substantial). These city-wide changes describe the aggregate availability of infrastructure for bicyclists. However, they should not be interpreted as the availability for an individual bicyclist, as individual travel contexts vary widely. We will address individual infrastructure availability in the four competing explanation sections below.

**Table 2. Infrastructure Availability**

| | Length of Infrastructure as a Percentage of Total Network Length | | Percentage point change |
|---|---|---|---|
| | Phase 1 | Phase 2 | |
| *Route (without Sharrow)* | 3.3% | 3.4% | 0.1 |
| *Route (with Sharrow)* | 2.3% | 6.5% | 4.2 |
| *Bike lane* | 3.8% | 5.5% | 1.7 |
| *Buffered bike lane* | 0.2% | 0.6% | 0.4 |
| *Safe-hit bike lane* | 0.1% | 0.4% | 0.3 |
| *Concrete curb bike lane* | 0.0% | 0.1% | 0.1 |
| *Parking-protected bike lane* | 0.0% | 0.1% | 0.1 |
| *Bike path* | 2.5% | 2.5% | 0.0 |
| *No infrastructure* | 87.9% | 80.9% | -7.0 |

Bicyclists' recorded routes follow the main patterns of bicyclist activity as noted in the annual bike count reports produced by SFMTA (San Francisco Municipal Transportation Agency, 2015). A few main corridors exist through the Mission district which run north/south connecting to Bernal Heights. Several east/west corridors connecting the Castro to Potrero Hill, a northeast/southwest corridor following Market and parallel streets, and the "Wiggle" (connecting Golden Gate Park and Height Ashbury down to Castro/Mission and downtown neighborhoods), are frequently used. Less dominant corridors exist such as the north/south connections between Golden Gate Park and the Presidio, and Polk street corridor connecting Russian Hill and Downtown. Finally, the Esplanade is a major corridor that connects all the bayside neighborhoods starting with South of Market through North Beach and to the Marina. Figure 8 aggregates the bicyclist routes as a weighted volume of percentage of link use. This map best describes the above-mentioned corridors, but also variability in routes across the city. Route variability is greatest in the non-core bicycling areas of the city where origin-destination pairs are likely most diverse. However, the presence of parallel corridors particularly those parallel to Market Street and the numerous east/west streets through the Mission district show that even when origin-destination pairs are similar, bicycling route choices have variability.

Differences between routes in phase 1 and 2 are presented in Figure 9. Like Figure 8, this map is an aggregation of weighted network link use, but as a percentage difference between phases. This map shows that although the main corridors of use between phases are quite similar, small differences in link use are observed. The northeast/southwest Market street corridor is less used in phase 2, as is the "Wiggle." This is most likely due to differences in the phase 1 and 2 samples as evidenced from count reports suggest bicycling in both of these areas has increased over time (San Francisco Municipal Transportation Agency, 2013a, 2008). Also, since alternative parallel routes were not shown to have increased bicycle volumes, it is likely that the trends from phase 1 to 2 are not indicative of a shifting of route behavior. Similarly, the increased use of the Esplanade and more bicycling activity in Western Addition and Pacific Heights in phase 2, since not accompanied with parallel route decreases, suggests that these difference may be due to variability in phase samples.


**Figure 8. Weighted bike volume, proportion of link use.**

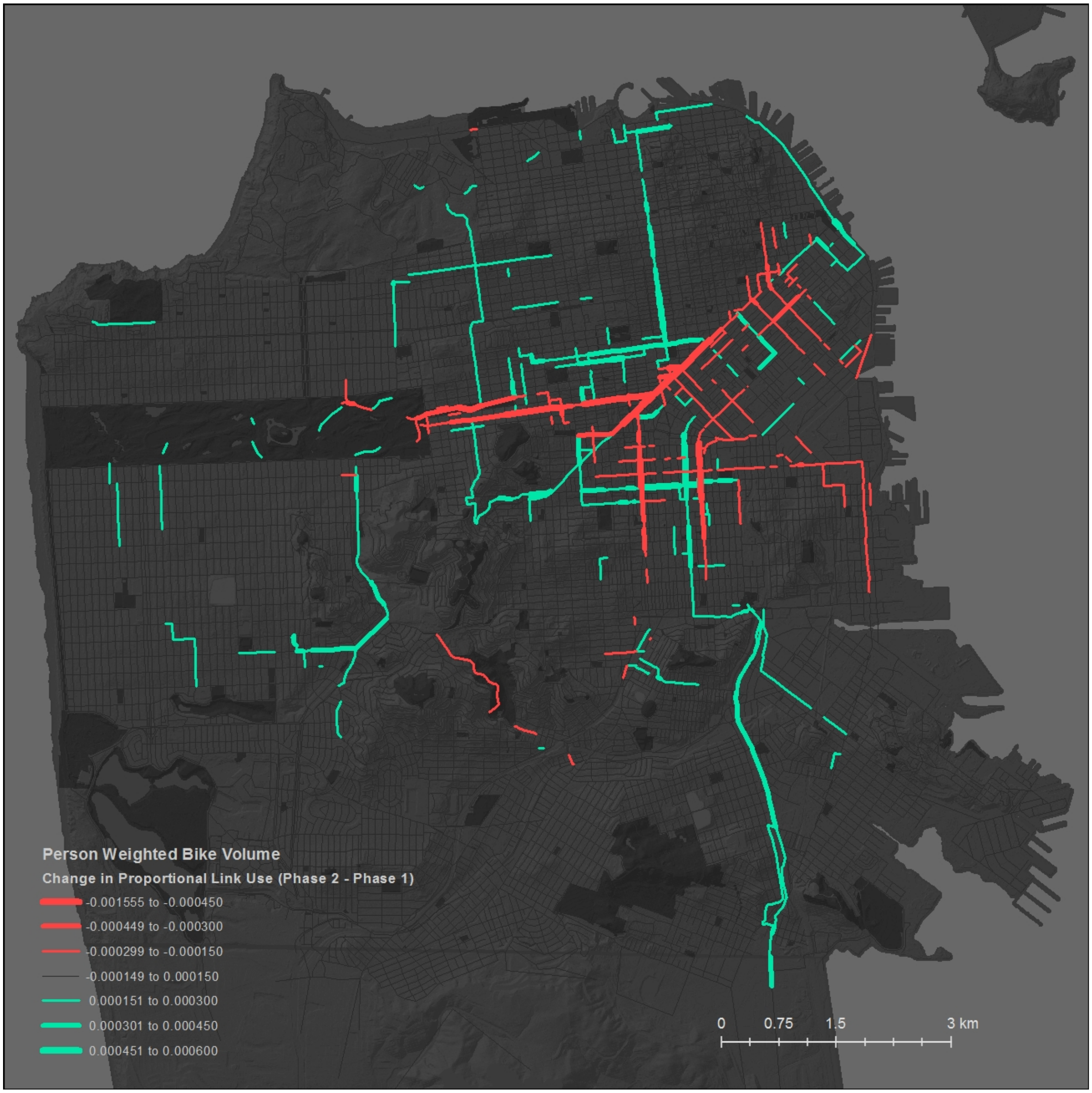


**Figure 9. Weighted bike volume, differences in proportion of link use by phase.**

The maps of weighted bicycling volume across the city help describe the overall patterns of route stability and change, though the changes displayed in Figure 8 and Figure 9 do not indicate whether the patterns of change are associated with changes in infrastructure use. In order to examine infrastructure use, we use the same weighting scheme and compare aggregate infrastructure use between phases in Figure 10. Overall, infrastructure use in phase 2 is slightly greater than that of phase 1. Though most infrastructure types are used in similar proportions in phase 1 and phase 2, there appears to be some switching between sharrows and bike routes (thanks to city efforts to paint sharrows on bicycle routes), and our sample of bicyclists also seems to have begun using the new parking protected and concrete curb bike lanes to a small extent in phase 2.

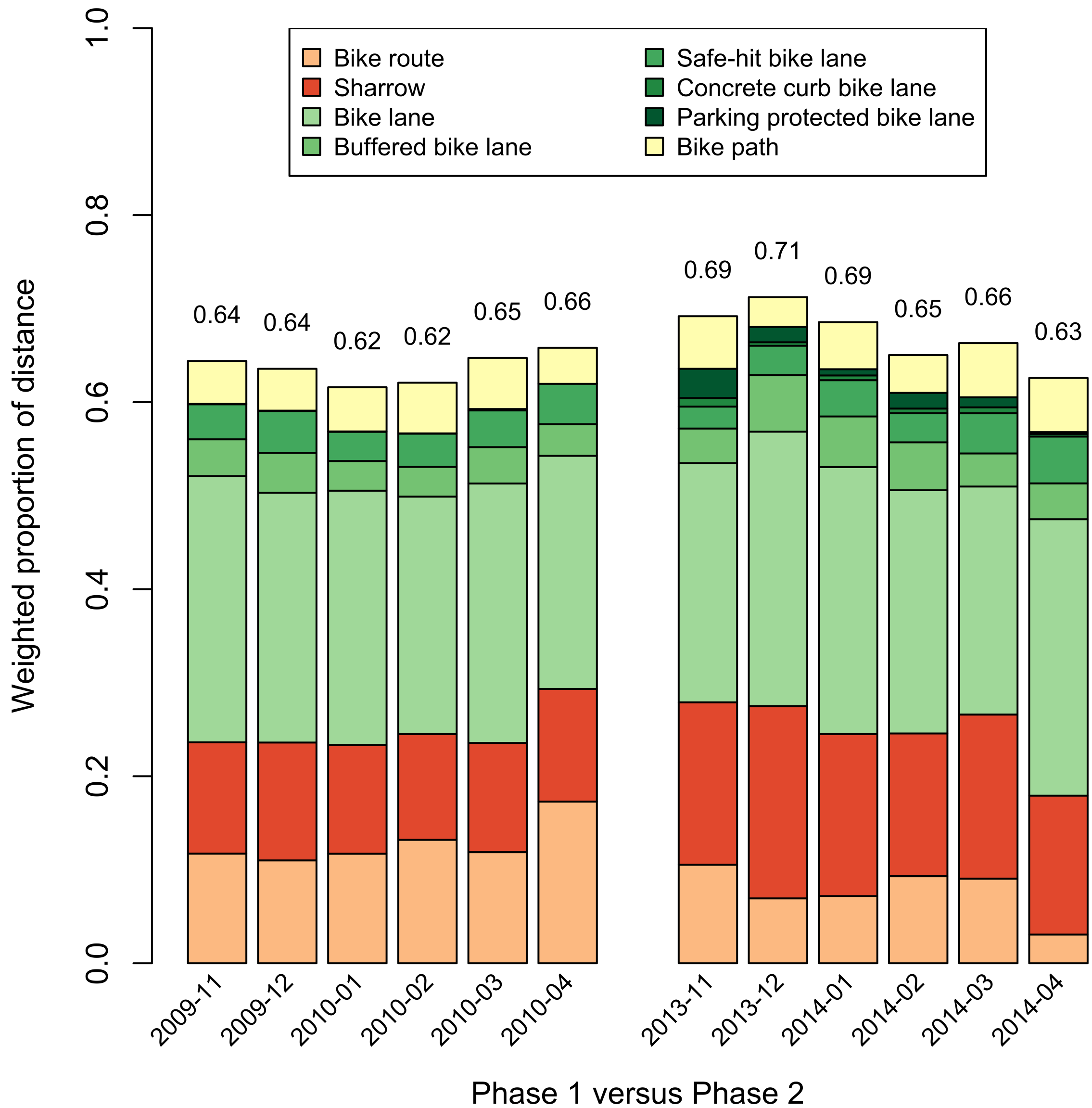


**Figure 10. Infrastructure use by month (in phase 1 and 2)**

The average monthly infrastructure use displayed in Figure 10 masks much of the variation between individuals. In reality, individuals exhibited a wide array of infrastructure use patterns, as shown in Figure 11. For example, while most CycleTracks users in our sample used bike lanes at higher rates than sharrows or bike routes, there were some individuals whose trips showed a different pattern altogether – riding most or all of their trips on bike routes or sharrows and very little on bike lanes. However, as indicated by their lines' light shade and narrow width, these riders typically only recorded a few weighted trips.

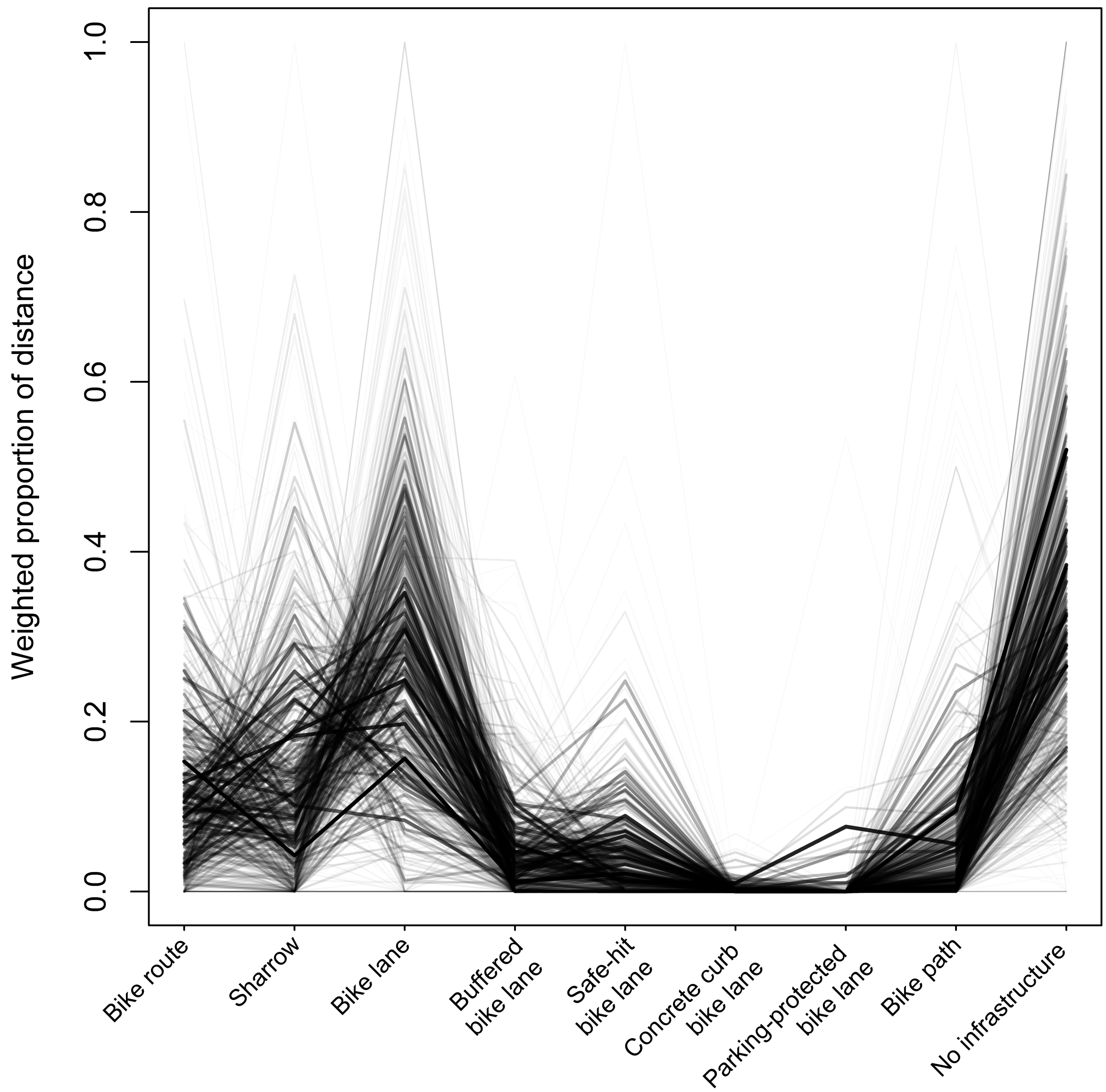


**Figure 11. Weighted infrastructure use by person. CycleTracks users with more weighted trips recorded are represented by darker and thicker lines.**

We further examined the differences in infrastructure use by gender and by trip purpose – two of the characteristics available from most CycleTracks users, rather than our survey. We observed that bicycle infrastructure use is higher for work and school trips than for non-work trips (Figure 12) and that women utilize bike infrastructure at higher rates than men (Figure 13), particularly sharrows and bike paths.

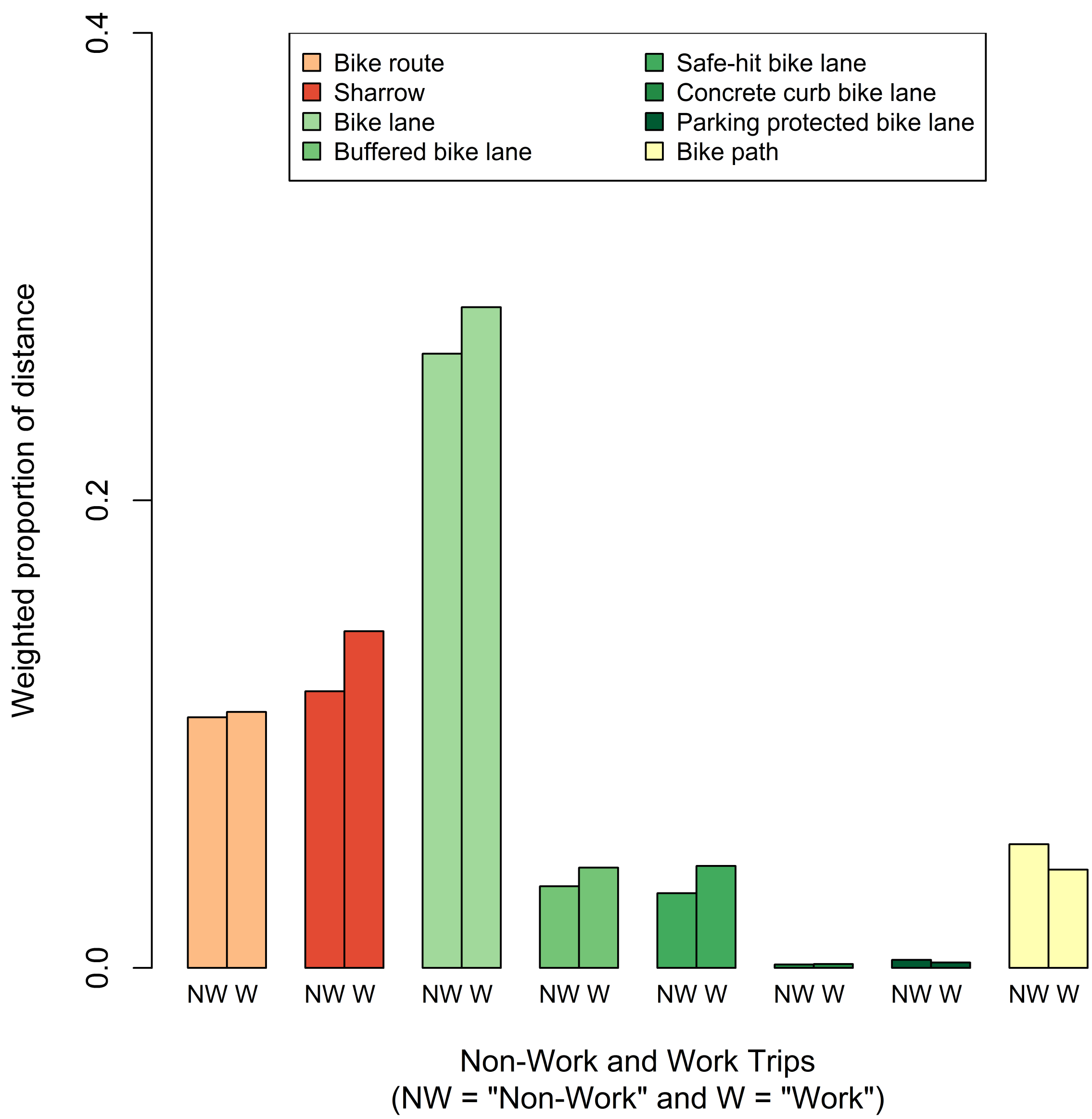


**Figure 12. Infrastructure use by trip purpose (work vs. non-work). We group commute, school, and work related trips into “work”, and social, shopping, and errand as “non-work”.**

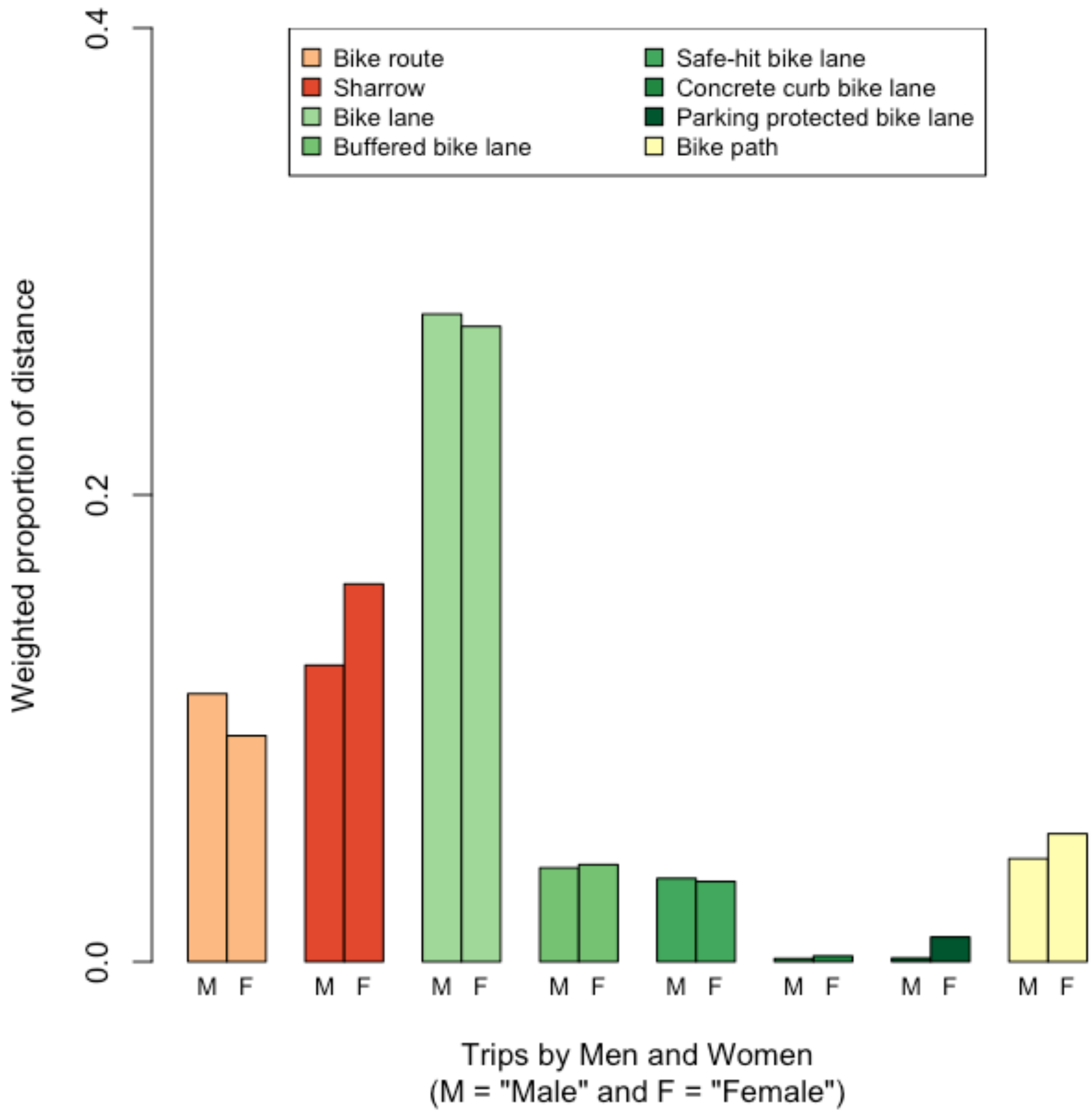


**Figure 13. Infrastructure use by gender**

## Do Changes in Bicycling Infrastructure Cause Changes in Bicycling Routes?

### *Explanation 1: Travel context explanation*

In an attempt to disentangle the causal mechanisms of route choice change, we first hypothesize that an individual's travel context (i.e. same or different riders with different trip characteristics, such as different origins and destinations and trip purpose) explains the observed differences in aggregate infrastructure use (phase 1 and 2, and/or over time).

Evidence for this hypothesis would suggest that sample differences between phases, or general variability over time, rather than behavioral change, explain route differences.

We find that the geographic distribution of trips from Phase 1 and 2 are strikingly similar when we plot and regress paired neighborhood weighted trip productions and attractions (Figure 14 and Figure 15). Our regressions show strong correlations between Phase 1 and 2, with $R^2$ values of 0.87 and 0.91, respectively (Figure 14 and Figure 15). This indicates that in the aggregate, trips are originating and terminating in similar areas between phases, counter to the travel context explanation.

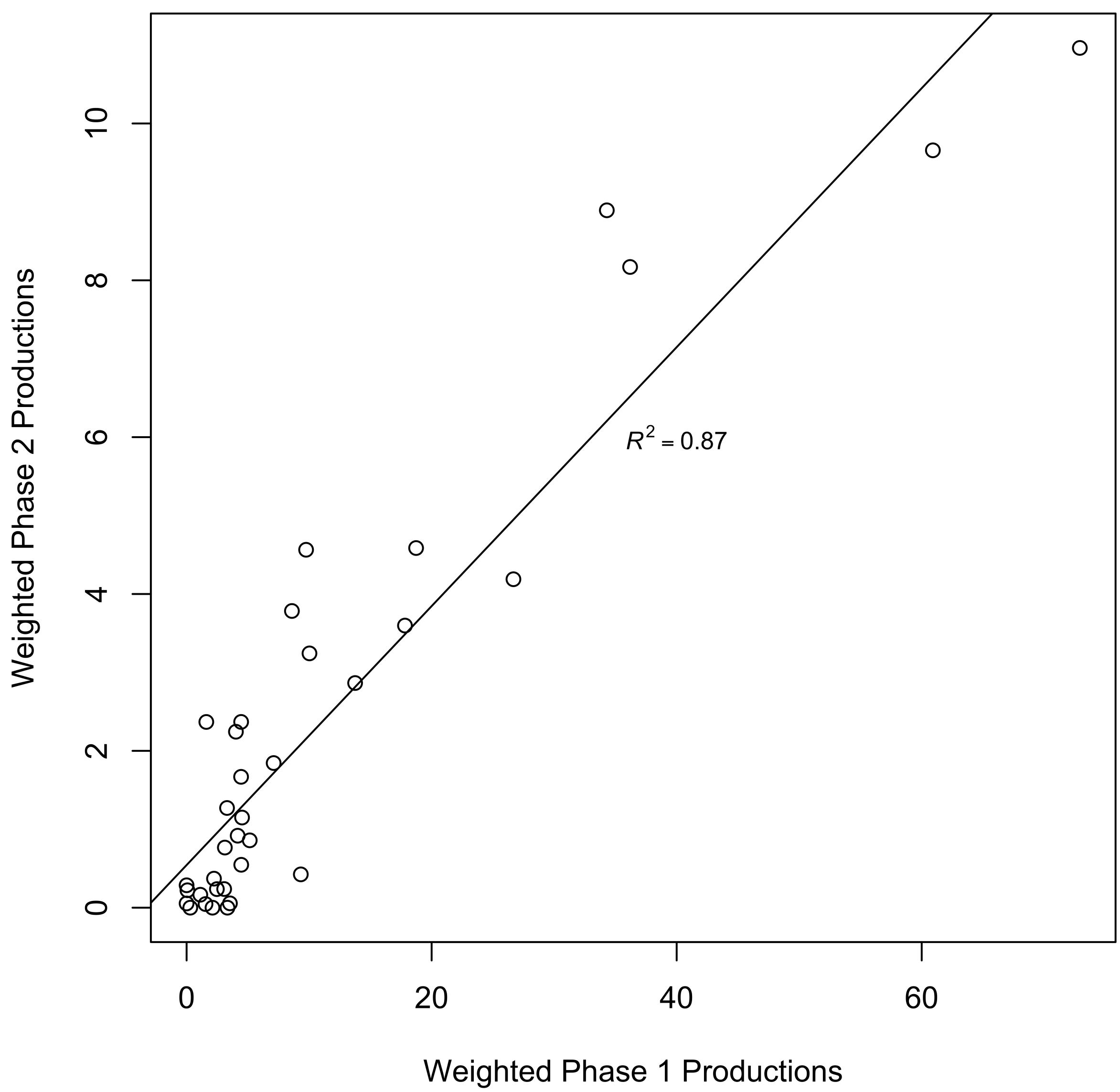


**Figure 14. Scatter plot of weighted trip productions by neighborhood for Phase 1 and 2**

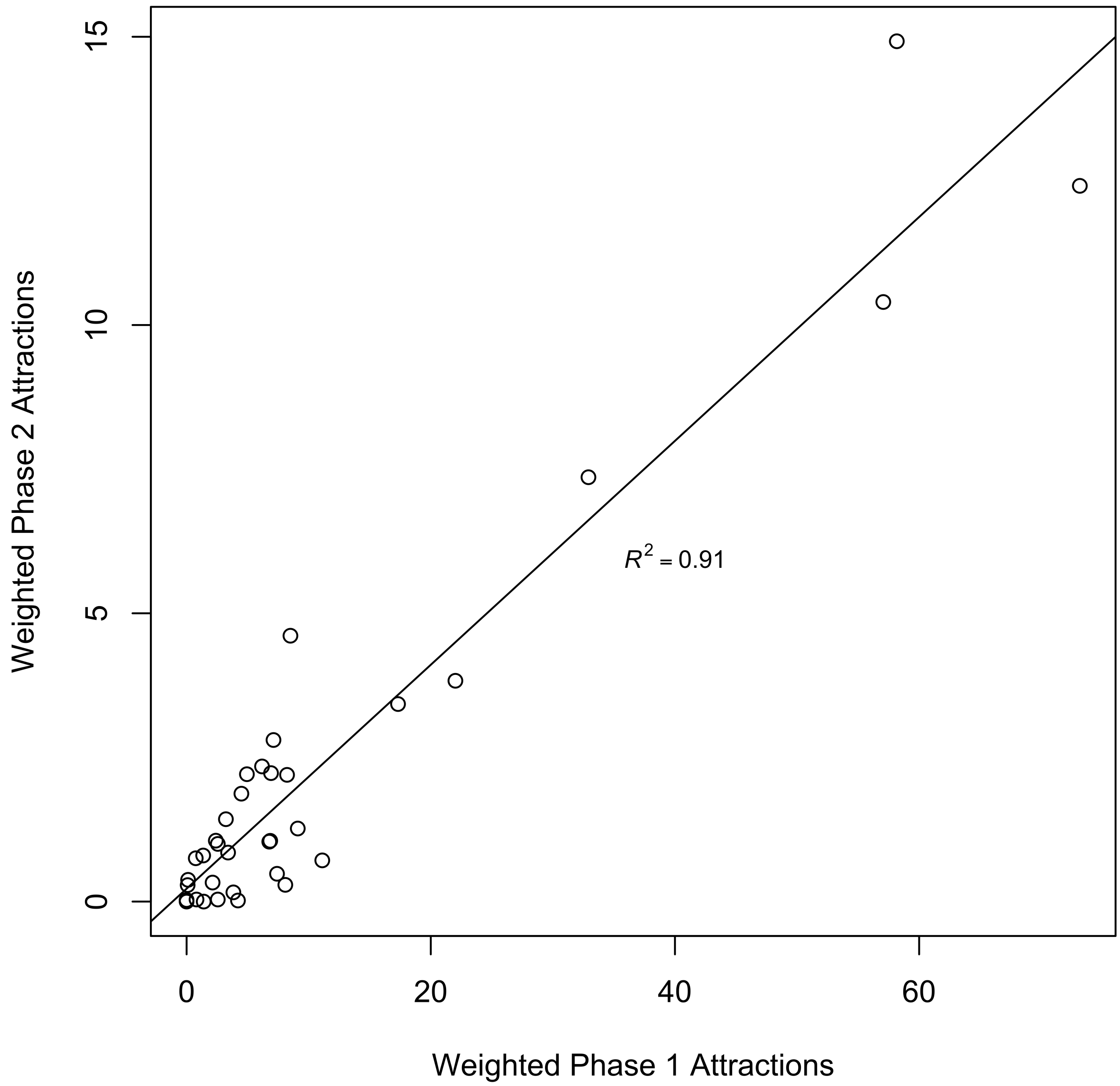


**Figure 15. Scatter plot of weighted trip attractions by neighborhood for Phase 1 and 2**

The combined weighted productions and attractions show a clear focus of bicycling in the neighborhoods of the Castro, Mission, Western Addition, South of Market, Downtown, and Financial District (see Figure 16). This map provides further detail regarding the previously-discussed bicycling core area of the city (see background), indicating that not only is bicycling *through* this core very prevalent, but that trips commonly start and end in the core.

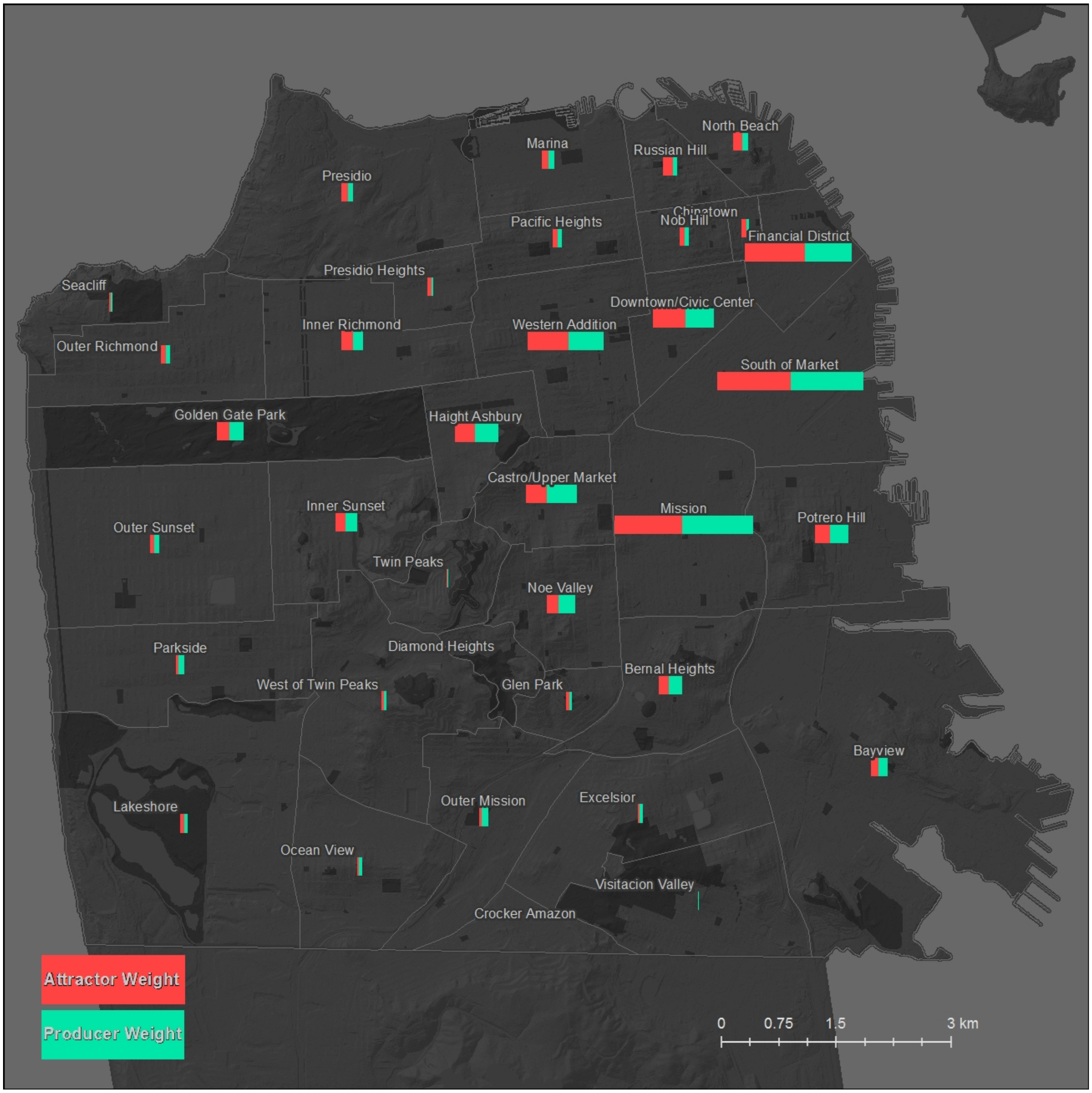


**Figure 16. Map of weighted productions and attractions for all trips**

In addition to examining trip productions and attractions, we also analyzed the correlations between weighted trip distributions between phases. We find that the weighted trip distribution between phase 1 and 2 is well correlated ($R^2$ of 0.64, see Figure 17), though not at the high levels we observed for productions and attractions. This provides some evidence in favor of the travel context explanation, as the variability in trip distribution may play a role in determining infrastructure use change. However, comparing a map of weighted trip distribution (Figure 18) with a map of weighted productions and attractions (Figure 16), evidence suggests that much of the variability in trip distribution is located within the bicycling core. We can infer that even when the origin-destination neighborhoods differ between phases, bicyclists are likely to be traveling through the same general neighborhoods in phase 2 as they did in phase 1.

Given that the general trip distribution trend is consistent between phases, and that bicyclists with differing origin-destination pairs between phases are likely to be riding through the same neighborhoods, we suggest that the travel context explanation only weakly accounts for differences in aggregate infrastructure use. This conclusion is further corroborated by the similarity in weighted trip distance distributions between phases (Figure 19). If the distance distributions had been quite different by phase, infrastructure use changes could be purely driven by differences in potential exposure to infrastructure. In reality, however, we observe only a slight shift in the density curve for phase two (suggesting slightly longer trips on average). This is not surprising given the similarity in trip distribution between phases (Figure 17). And though it is difficult to estimate the magnitude of influence a slight increase in trip distance might have on infrastructure exposure given the citywide variation in infrastructure availability, it seems unlikely that this change has an overwhelming influence on infrastructure use, relative to the other explanations.

It is important to distinguish between aggregate and individual infrastructure use changes. As evidenced by the jagged nature of individuals' longitudinal infrastructure use patterns in Figure 11, a particular individual's travel context may play an instrumental role in their infrastructure use. In other words, an individual who uses sharrows at a high rate has a correspondingly low use of bike lanes, a consistent pattern across individuals which likely emerges due to travel context rather than attitudinal or preference differences. Given these observations, we suggest that the travel context explanation, while perhaps relevant to particular individual differences on the margin, only weakly accounts for differences in aggregate infrastructure use between phases.

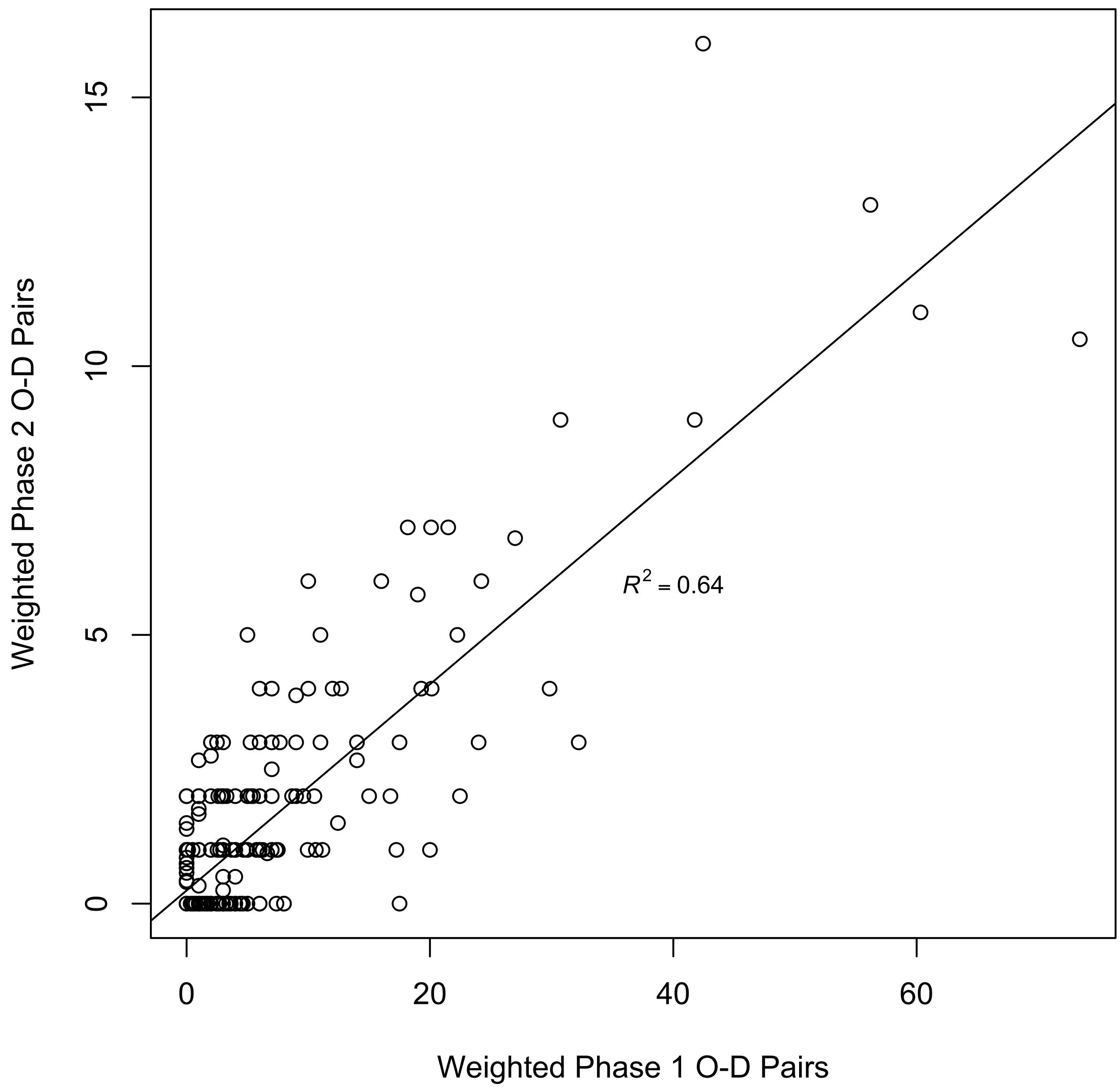


**Figure 17. Scatter plot of weighted trip distribution for Phase 1 and 2**


**Figure 18. Map of person weighted trip distribution for all trips**

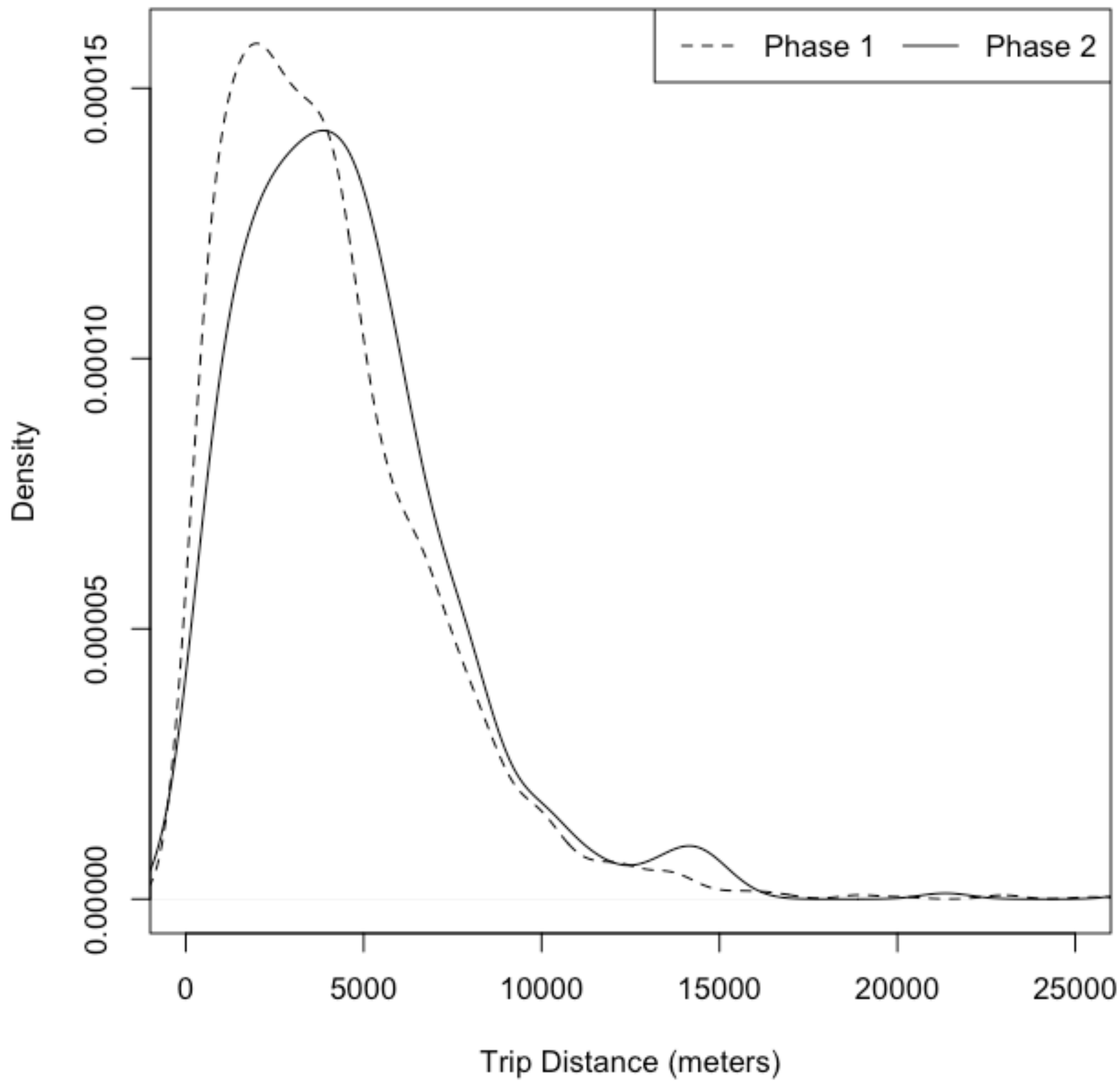


**Figure 19. Distance distribution by phase**

### *Explanation 2: Attitudinal cohort explanation*

We next examine a complementary conjecture to the travel context explanation: Given the repeat cross-sectional phase comparisons, infrastructure use changes are attributable to our sampling of riders with different attitudes and preferences about comfort and general route characteristics at the two points in time. We term this the attitudinal cohort explanation because it would suggest that our phase 1 and 2 samples have systematic differences in attitudes. From Table 1 it is clear that the percent male, bicycling frequency, and percent work trip purpose are similar between phases. This is important, because these are well-established socio-demographic and travel behavioral factors that are often associated with bicycling attitudes and thus behavior. This suggests that the phase 1 and 2 samples may therefore have consistent attitudes and preferences for infrastructure. In this section we take a closer look at the sample characteristics that might indicate systematic differences in bicycling attitudes and preferences.

The age distribution of CycleTracks participants remained strikingly similar between phases, bolstering the argument that the phase samples are relatively comparable (Figure 20). In contrast, the phase 1 and 2 samples differ on reported comfort on a four lane road with no bike lane (Figure 21). Fewer phase 2 survey participants reported that they would not ride on this particular facility, indicating that our phase 2 sample of CycleTracks users may be more comfortable on any bicycle infrastructure, or lack thereof. However, survey participants' reported facility comfort may have little impact on their observed route choices, as the "Wouldn't Ride" and "Uncomfortable" respondents' infrastructure use is strikingly similar, and the "Comfortable" respondents had small deviations from there on bike lane and parking-protected bike lane use (Figure 22).

To further complicate the evidence base for the attitudinal cohort hypothesis, the phase differences in bicycling experience run counter to the pattern established by the comfort responses. The survey participants in phase 2 tend to have less bicycling experience, yet also report being comfortable on busy streets with no bicycle infrastructure.

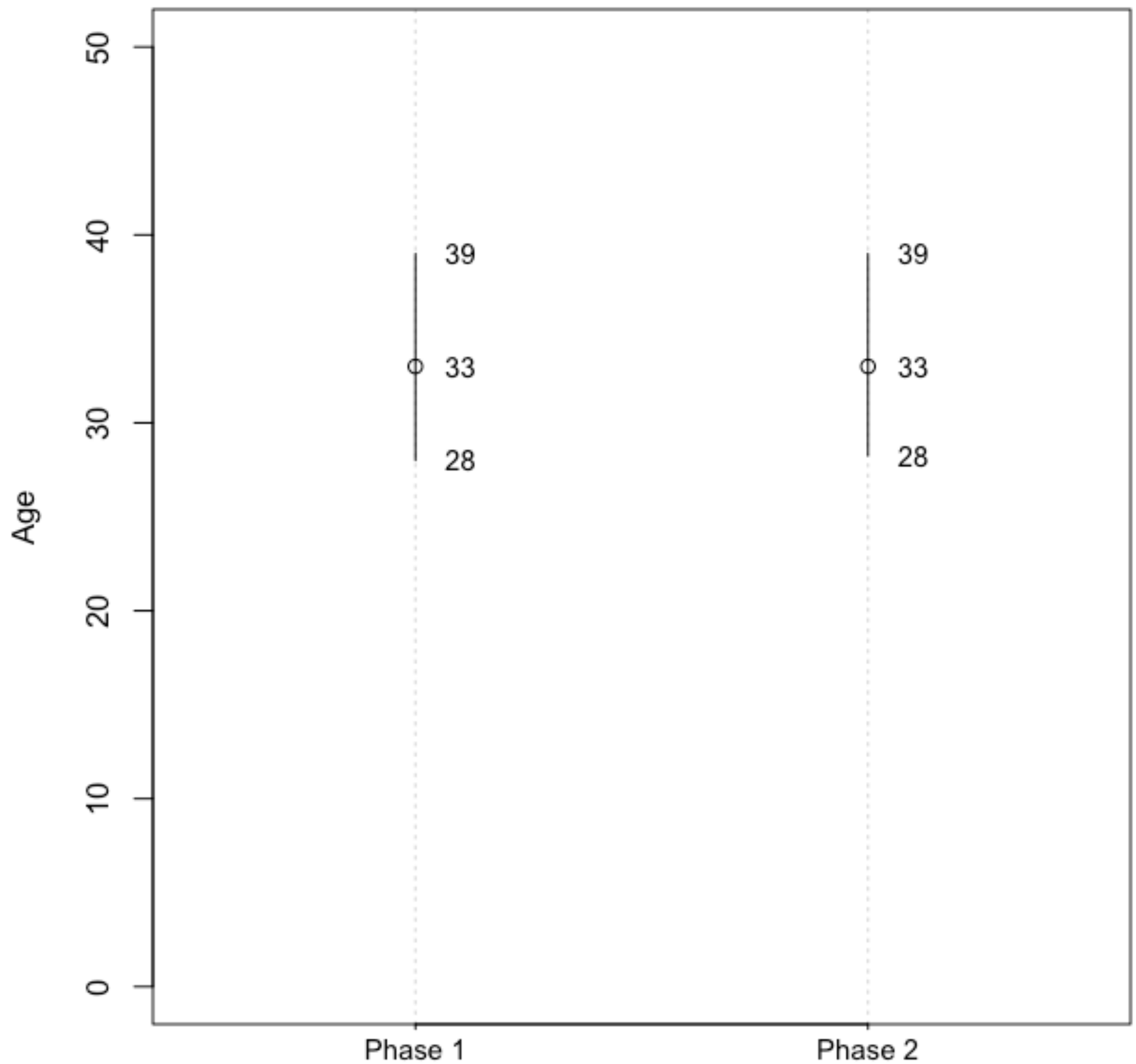


**Figure 20. CycleTracks participant age by phase**

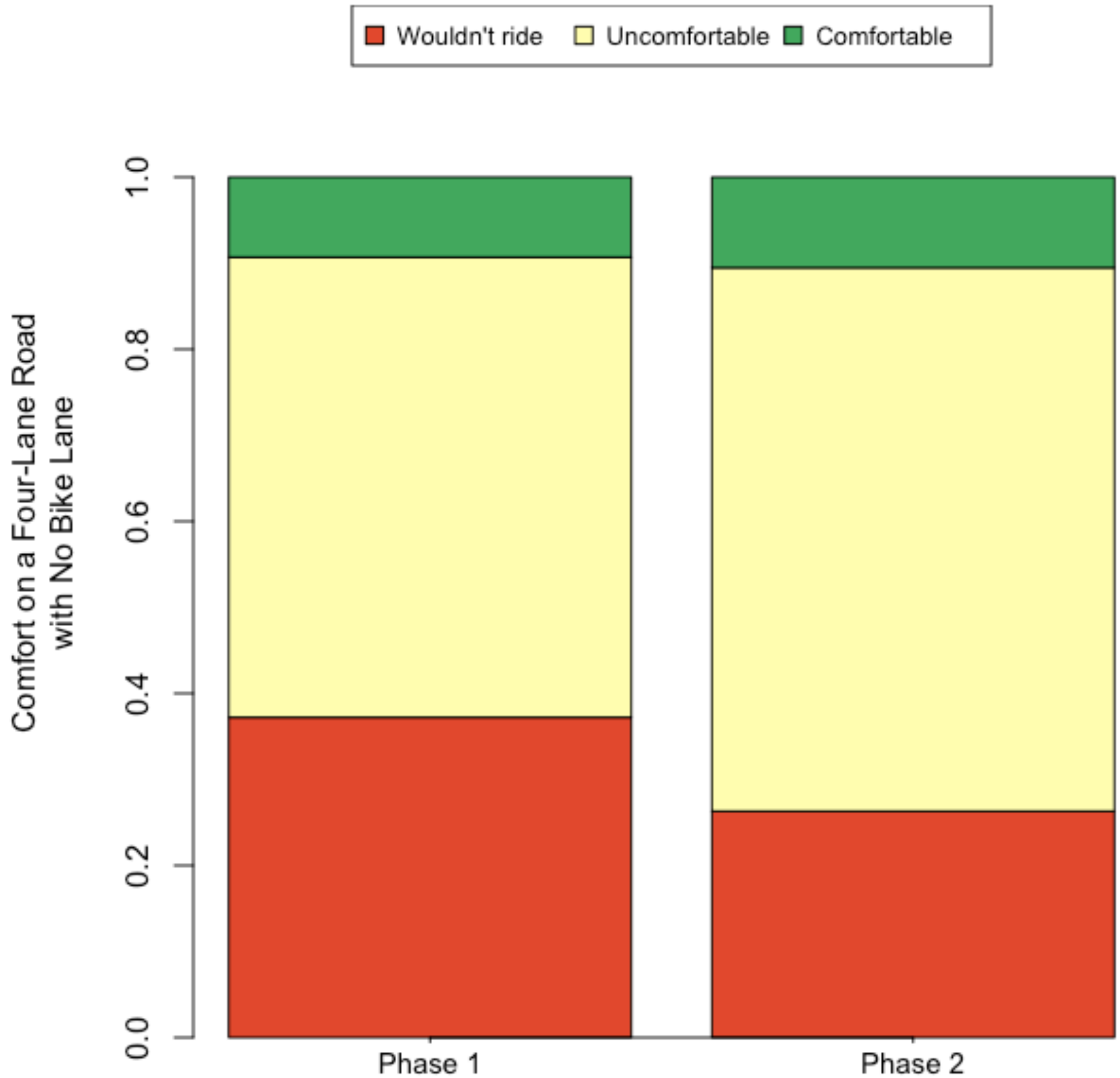


**Figure 21. Survey participant comfort by phase**

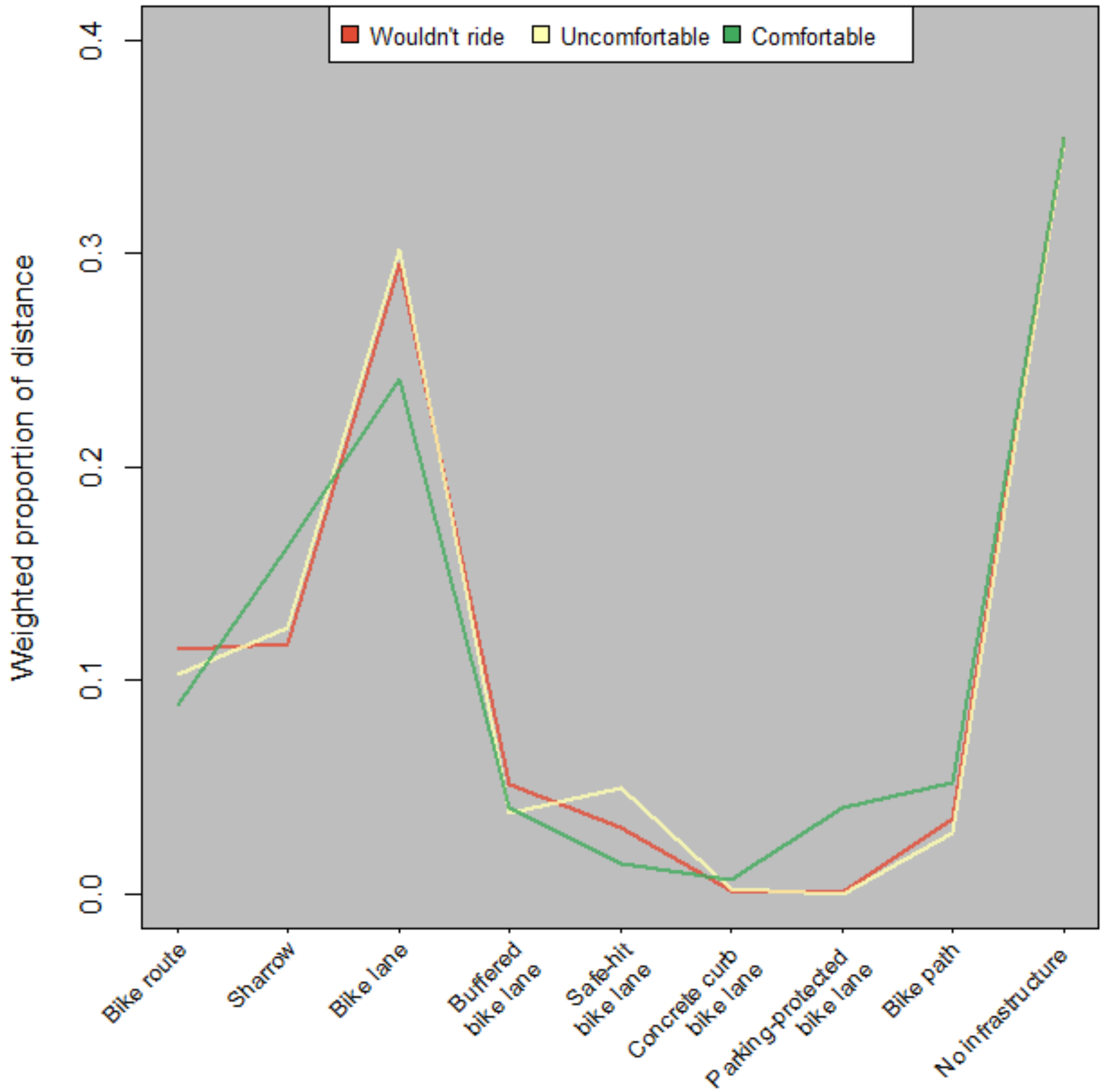


**Figure 22. Survey participant infrastructure use by perceived comfort on a four lane road with no bicycle lanes**

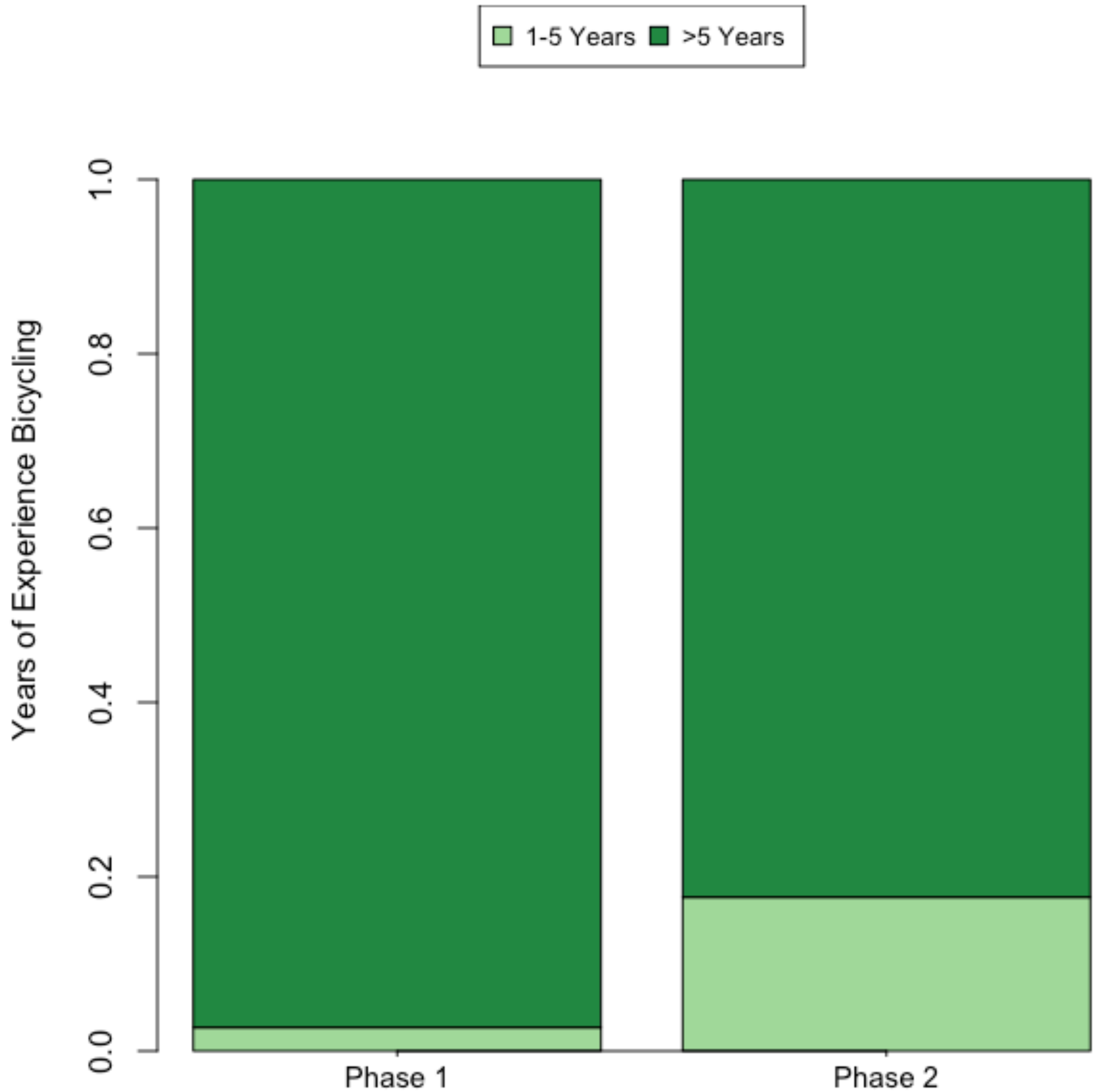


**Figure 23. Survey participant bicycling experience by phase**

Another indicator of bicycling attitudes and preferences is the amount of distance deflected from the Euclidean distance from origin to destination. Distance deflected describes how far "out of the way" (some of this mandated by physical barriers, some by route choices) a person rode to get to their destination. Numerous studies have established the principle of trade-off between bicycling comfort and route directness; to bicycle on safer infrastructure, a bicyclist will often require a substantial detour (Krizek, 2006). Though there is substantial between-individual variation in distance deflected in our sample, phase 2 has only a minor, negligible increase in deflection over the phase 1 average (Figure 24). Though any conclusions drawn from this plot are confounded by the infrastructure changes being implemented around the city between phases, the similar deflection ratios between phases lends credence to the idea that the phase 2 bicyclists do not hold substantially stronger or weaker attitudes regarding comfort or general preferences for bicycle infrastructure compared to our phase 2 sample. To shed further light on this indicator, we conducted further bivariate analyses of distance deflected by gender, purpose and comfort, with no divergent results (not shown).

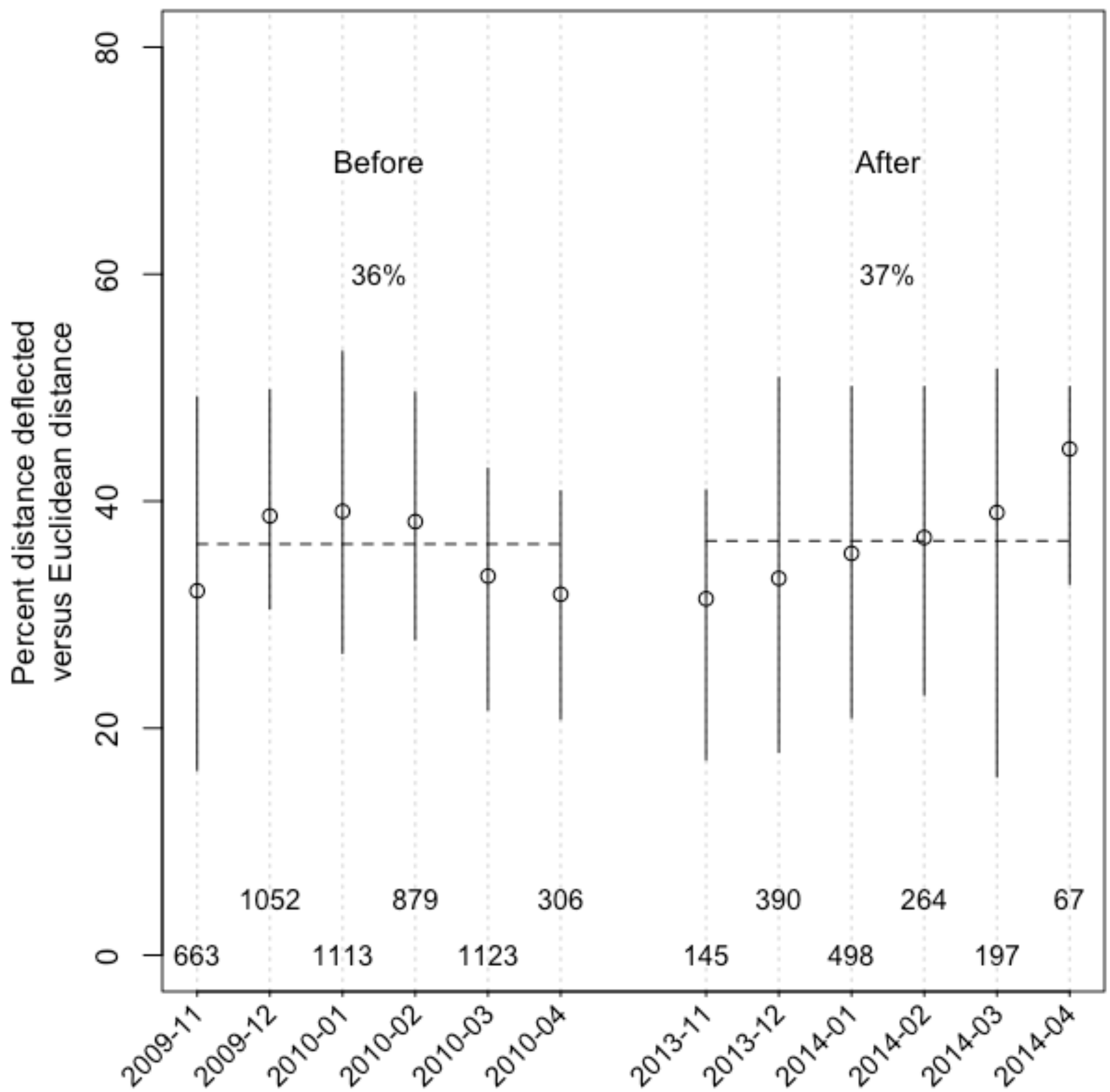


**Figure 24. Distance deflected by month for phase 1 and 2. The points represent the median distance deflected and the bars represent the 1st and 3rd quartiles for that month.**

### *Explanation 3: Targeted planning explanation*

The third explanation we propose is that increases in infrastructure use are driven by planning efforts, with planners placing infrastructure investments along the same roads that bicyclists already use. Though at first blush it might seem unnecessary to build infrastructure where bicyclists already ride, planners might be attempting to rectify a safety or traffic flow issue along a busy bicycle corridor through improved, safer bicycle infrastructure. Since safety is a major concern in planning for bicycling, this type of scenario is likely common. However, although San Francisco has a core bicycling area (noted in the introduction and clear form Figure 8), bicycling investments have been widespread around the city (Figure 25). This geographic mismatch between existing bicycling volumes and investment therefore provides evidence counter to this third explanation at the city scale. Nonetheless, San Francisco's widespread infrastructure investments could still contain some evidence of targeted planning within a given neighborhood. In the survey subsample, over 30% of respondents say a bike lane, safe-hit bike lane, or sharrows were installed on their primary route to their most common bicycling destination, among smaller percentages of other infrastructure (Figure 26, lower third). This indicates that there is likely some neighborhood scale targeted planning of bike infrastructure.

We also analyze CycleTracks users' distance deflected from the Euclidean distance to try to understand if we miss local route level details in our geographic aggregations above. Figure 24 demonstrates that distance deflected is nearly constant between phases. When pairing this with the results that infrastructure use is slightly higher in phase 2 (Figure 10), we do find evidence for the targeted planning explanation through observed route behavior. That is, if distance deflected did not rise substantially while infrastructure use did increase, then some of the infrastructure use change must be due to planners placing infrastructure where many bicyclists already ride.

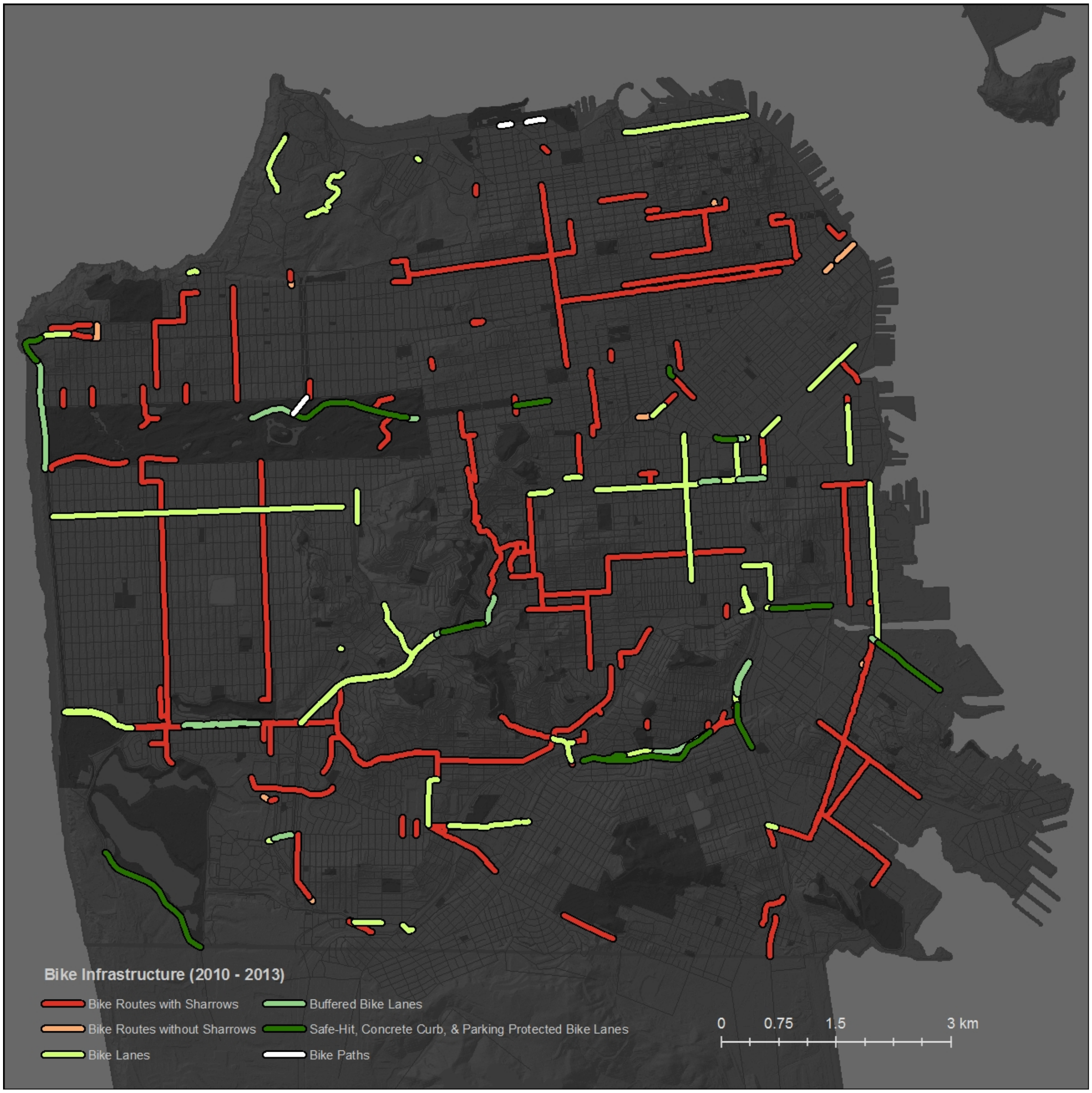


**Figure 25. Infrastructure investments between Phase 1 and 2.**

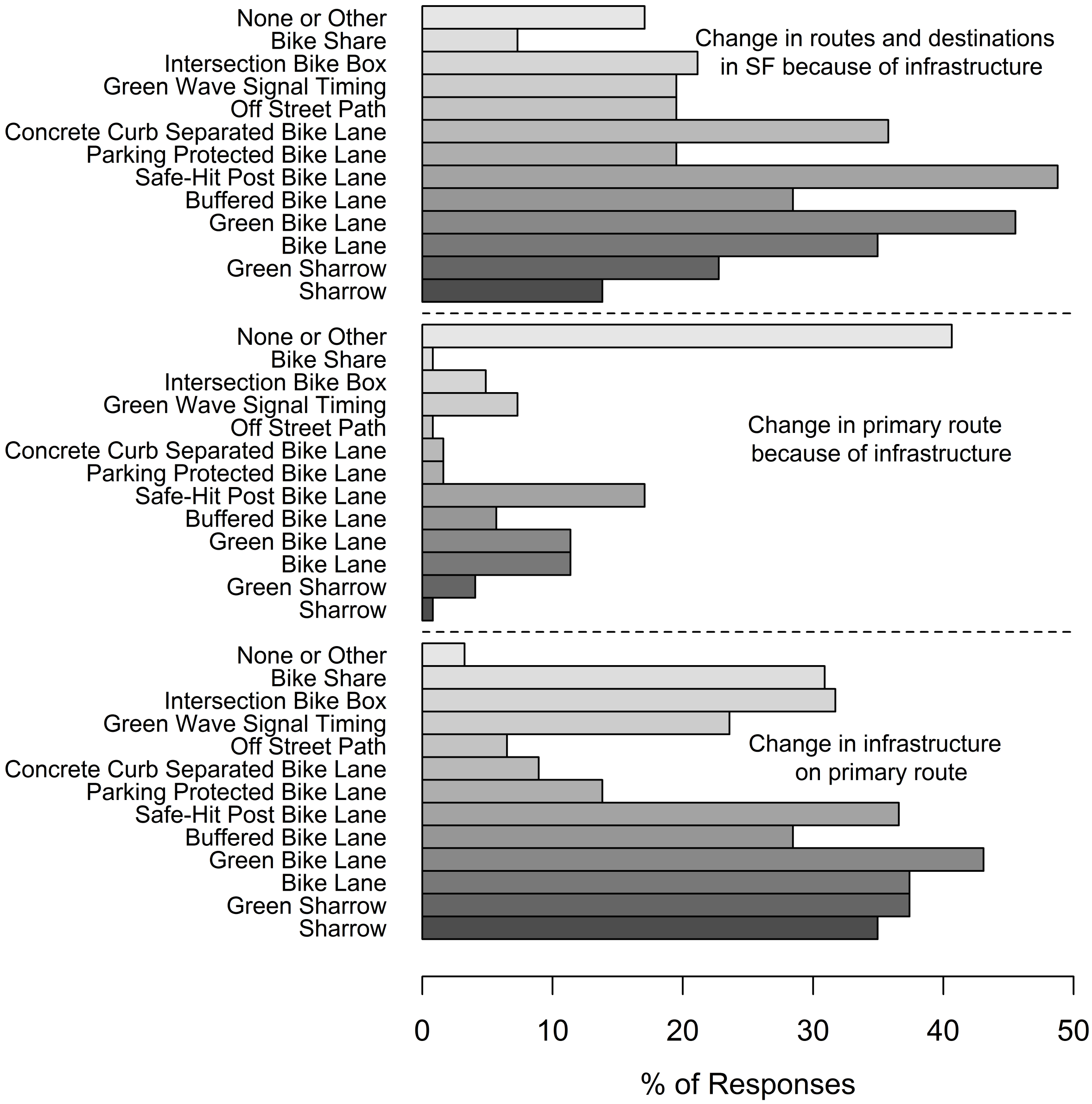


**Figure 26. Survey responses about infrastructure installation and change of routes.**

### *Explanation 4: Route change explanation*

Our final explanation posits that infrastructure use changes between phases are a result of bicyclists changing their routes in order to access new bicycle infrastructure. Good evidence for this explanation comes from the areas of the city where parallel routes have opposing shifts in bicycling volume. For example, the east/west connection between the Castro and Potrero Hill neighborhoods becomes more focused on 17$^{th}$ Street in phase 2 while the many parallel streets (16$^{th}$, 15$^{th}$, and 14$^{th}$ Streets) see declines in their relative use between phases (Figure 9).

Similarly, the north/south connection through the Mission neighborhood is slightly more concentrated on Folsom Street in phase 2, and the short corridor connection between the Castro and Western Addition neighborhoods is more focused on Sanchez Street (Figure 9). Another corridor of increased use is on Bayshore Boulevard entering the core bicycling area from the south (Figure 9). However, it is possible that Bayshore Boulevard's volume increase is due to sampling variability (i.e. it does not have corresponding parallel decreases), thanks to a new rider or new riders in our phase 2 sample.

Each of these four streets saw infrastructure investments between phases. Conventional bicycle lanes were installed on $17^{th}$, Folsom, and Bayshore, buffered bicycle lanes were installed on $17^{th}$ and Bayshore, safe-hit posts were installed on Bayshore, and sharrows were painted on Sanchez.

Further support for the route change explanation comes directly from the survey responses. We asked survey participants explicitly if infrastructure has changed their routes both "on their primary route to their most common destination" and "in general on routes and destinations around San Francisco". Results demonstrate that many bicyclists consider infrastructure influential in their route and destination choices when bicycling (Figure 26). As we would expect, infrastructure investments have a larger impact on their general bicycling choices than their specific route choice to their most common destination (i.e. most likely because of the base rate of infrastructure available in each scenario). Besides this magnitude difference, the results for both general and specific routes have very similar relative differences between infrastructure types. For example, many more people indicate that various types of conventional and innovative bike lanes cause them to change their routes compared to sharrows. Green painted sharrows seem to cause more route change compared to regular sharrows, perhaps indicating that bicyclists consider them to be a better indicator for drivers to share the road. In addition, many of the innovative types of bike lanes (e.g. buffered, safe-hit posts, curb protected) are reported by many (~20-50%) to be influential in their bicycling choices even though they are only a small fraction of the network in the city (Table 2).

We also report some infrastructure investments which we were not able to measure with observed route data due to the difficulty in measuring the impact of the types of features (e.g. green wave signal timing, bike boxes at intersections, and bike share) (Figure 26). Less than 10% of respondents report bike share to be influential in their route choices, suggesting this sample does not use bike share; while the other two innovations (green wave, bike boxes) influenced about 20% of the respondents. Although we don't have evidence for observed behavior and these infrastructure types, these results suggest further study of CycleTracks routes near these types of infrastructure may show route change.

### Comparing Infrastructure Use Explanations

Of our four explanations for changes in infrastructure use, we find weak evidence for the travel context and attitudinal cohort explanations and stronger evidence for the targeted planning and route change explanations. This indicates that the increase in bicycle infrastructure use in phase 2 is a true planning success rather than an artifact of our data, in particular the variation in sampling between phases 1 and 2. By establishing the relative unlikeliness of the first two explanations, we can then extend our analysis further to investigate the relative influence and importance of the different infrastructure types installed in San Francisco between 2010 and 2013.

### Evaluating Infrastructure Types in San Francisco

Transportation planners must balance a number of objectives when considering infrastructure investments, such as safety, cost, and other attributes. In this section, however, we evaluate investments on a network-wide level to determine which types of bicycle infrastructure appear to attract and retain bicyclists. To do so, we compare phase 1 and phase 2 proportional weighted link use and evaluate whether particular bicycle infrastructure types are associated with increases or decreases in use on a network-wide level (Figure 27).

Of all of the infrastructure types examined in this report, bicycle lane installations had the highest average increases in weighted link use, followed by sharrows. When comparing these two types of infrastructure based on their ratio of use to availability city-wide (i.e. the relative difference in infrastructure use given its availability), we see that bike lanes are much more likely to be used (Figure 27). However, our results in Figure 11 demonstrate that individuals can have different travel contexts where perhaps sharrows are the only reasonable infrastructure to use in getting from origin to destination, and vice versa for bike lanes. Nonetheless, given (a) that the nature of adding sharrows to predetermined bike routes is much less strategic than installing new bike lanes and (b) that there is evidence that sharrows may not provide enough comfort or protection to cause route behavior change (Figure 26 and Ferenchak and Marshall, 2016), we suggest that our evidence supports the argument that in general bike lane use is more likely an explanation of targeted planning and behavior change, while sharrow use is more likely influenced by travel context of our samples and targeted planning. Without a more sophisticated empirical model (e.g. multi-level route choice), these results remain suggestive.

As for the remaining innovations depicted in Figure 27, buffered bike lanes saw increases in weighted proportional link use, on average, while the newest innovations – safe-hit post bike lanes, concrete curb bike lanes, and parking-protected bike lanes – actually saw small decreases in their relative overall use. However, we would warn against inferring too much from these patterns, especially in light of the survey results demonstrating their influence (Figure 26). Given the small relative availability of all four infrastructure types and the small increases and decreases in average link use among them, it is likely that either a much larger sample or a much larger investment in these infrastructure types would be needed to provide an assessment of their use. Further analysis of these infrastructure types will be warranted as they continue to be installed around San Francisco, and may benefit from more targeted intercept studies to evaluate.

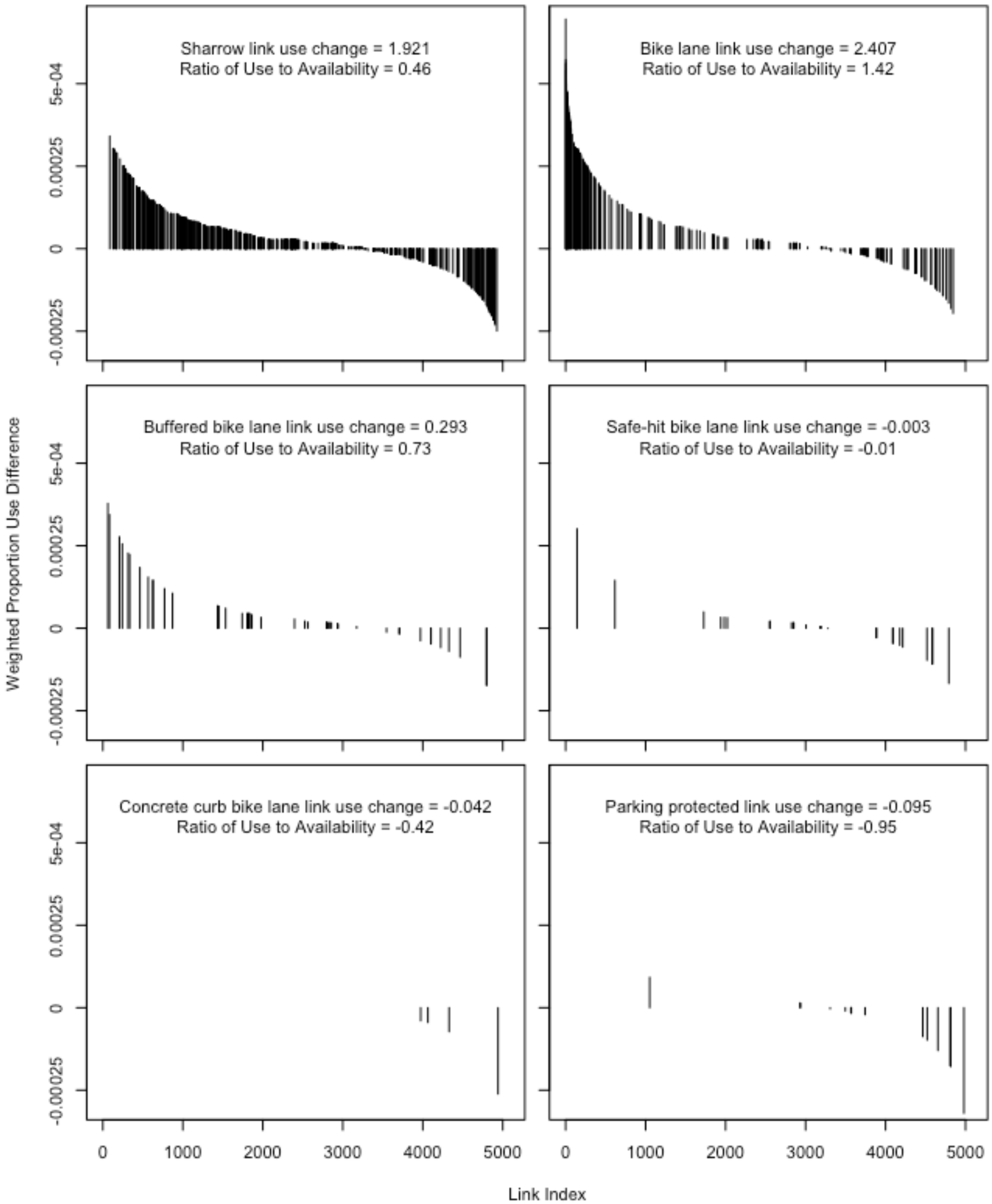


**Figure 27. Weighted difference between phase 1 and 2 link use proportions by infrastructure type.**

## Policy Implications

### *Data Needs and Management*

The efforts in collecting and measuring bicycling in San Francisco are considerable. SFCTA's creation of CycleTracks has provided an incredibly rich description of bicycling in San Francisco, and has had a major impact through spinoffs employed in numerous other cities (Figliozzi and Blanc, 2015). Additionally, our report could not have proceeded without SFMTA's database on bicycling infrastructure. Given that these agencies are at the forefront of local governments in measuring and modeling bicycling, we provide the following comments about needs, not to detract from the existing effort, but to enhance it.

Perhaps the most glaring limitation of our analyses in this report is the uncertainty in the date of bike infrastructure construction. Without specific dates, it is difficult to assess what type of infrastructure existed on a given person's recorded route (on a given day). While there is surely a record of construction completion dates in the city (perhaps as a paper record of work orders), neither SFMTA nor SFCTA had ready access to this information. We see this as a very important data gap for evaluating the relationship between bike infrastructure and bicycling behavior. While our approach of aggregating infrastructure in time (annually) is a working solution for use of CycleTracks data to evaluate infrastructure use, our results would be considerably less noisy and more defendable had we access to infrastructure dates. We suggest that management of bike infrastructure (and all road investments that benefit from a later evaluation) data include specific temporal information along with its geographic information.

In addition, considerable time was spend marrying infrastructure data managed by SFMTA in a reference GIS layer (good for mapping), to a modeling network managed by SFCTA (good for routing). We suggest a coordinated effort to join these layers could result in considerable efficiency gains for future infrastructure evaluations as more CycleTracks data is collected and more infrastructure investments are made.

### *Infrastructure Evaluation Approaches*

One way to evaluate infrastructure investments is to examine the cost/benefits for each investment (Krizek et al., 2006). In the case of past investments, the costs are given, but the benefits are overwhelmingly difficult to quantify. This is because many benefits are diffuse and many take long periods of time to observe (e.g. public health). In the case of San Francisco, where increasing the bike mode share is a planning objective in and of itself, assessment of infrastructure success may be simplified to prevalence of bicycling. However, care must be taken to consider the equity implications for such increases in bicycling (e.g. rise in property value, and willingness of sensitive populations to bicycle).

The general case for investing in bike infrastructure is perhaps justified simply by counts of bicyclists across the city, as illustrated in San Francisco's annual bike count reports (San Francisco Municipal Transportation Agency, 2015). The same could be said for the magnitude of investments (i.e. dollars spent). However, evaluation of *type* and *placement* of infrastructure is more difficult with counts because very little is known about the cause of an increase at a certain intersection. These evaluations benefit from individual behavioral observations in relation to past infrastructure investment, as illustrated in this report.

## Conclusion

Our analysis provides evidence that, on average, San Francisco bicyclists are able to ride on bicycle infrastructure for a greater portion of their routes through the city in 2014 than they were 5 years previously. We come to this conclusion by utilizing volunteered bicycle route GPS data provided by San Francisco bicyclists using a smartphone application developed by SFCTA called CycleTracks as well as by a survey of CycleTracks participants. Given the complexity of bicycle route choice decision-making and the characteristics of the available data, we use a before-and-after quasi-experimental approach and compare four possible explanations for this trend of increased infrastructure use. Ultimately, we suggest that a combination of targeted planning (i.e. planners putting infrastructure where bicyclists currently ride) and route change (i.e. infrastructure placed on alternative routes causes changes in chosen routes) best describe the increased use of San Francisco bicycle infrastructure in an average rider's route, rather than being caused simply by changes in the characteristics of our sample between 2009 and 2014.

There are several possible avenues for improving and extending our work. One area ripe for future analysis is the apparent planning strategy to "connect" outlying parts of the city to the bicycling core as part of a more complete, connected network of bicycle infrastructure. As our sample of bicyclists had only sparse representation in outlying residential neighborhoods, it remains an open question whether this represents a worthwhile investment. Are bicyclists using this infrastructure, rather than other nearby quiet residential streets? And if so, are they using it in sufficient numbers, or are the safety benefits substantial enough to justify the infrastructure investment? A more targeted bicycle count effort in these neighborhoods would more effectively account for the levels and location of bicycling than our crowd-sourced data from the CycleTracks smartphone application.

Another avenue for further exploration is the value San Francisco bicyclists attach to the novel, innovative infrastructure installed since 2010, such as the parking-protected or concrete-curb bicycle lanes. Because these infrastructure types were installed in small volumes around the city, it was difficult to evaluate their value to bicyclists. As more of these types of infrastructure are installed around the city, more robust analyses of their value relative to conventional bicycle lanes, routes, and sharrows will be possible. Again, targeted counts, ideally before and after installation, would also better serve to evaluate their value to bicyclists.

Though route choice modeling comes with its own set of assumptions and potential weaknesses, we also suggest that a discrete choice model of route choice would serve as a valuable complement to the analyses presented in this report. Such an effort could provide a validation check on the route choice model developed previously by SFCTA (Hood et al., 2011), but perhaps more importantly could also be used to evaluate the new, innovative bicycle infrastructure that has been added in the intervening years.

## References


ARB, 2011. Final Regional GHG Emission Targets.

Broach, J., Dill, J., Gliebe, J., 2012. Where do cyclists ride? A route choice model developed with revealed preference GPS data. Transp. Res. Part A Policy Pract. 46, 1730–1740. doi:10.1016/j.tra.2012.07.005

Buehler, R., Pucher, J., 2012a. International Overview: Cycling Trends in Western Europe, North America, and Australia, in: Pucher, J., Buehler, R. (Eds.), City Cycling. MIT Press, Cambridge, MA, pp. 9–30.

Buehler, R., Pucher, J., 2012b. Cycling to work in 90 large American cities: New evidence on the role of bike paths and lanes. Transportation (Amst). 39, 409–432. doi:10.1007/s11116-011-9355-8

Bushell, M.A., Poole, B.W., Zegeer, C. V., Rodriguez, D.A., 2013. Costs for Pedestrian and Bicyclist Infrastructure Improvements: A Resource for Researchers, Engineers, Planners, and the General Public. Chapel Hill, NC.

California Department of Transportation, 2015. Streets and Highways Code.

Charlton, B., Sall, E., Schwartz, M., Hood, J., 2011. Bicycle Route Choice Data Collection using GPS-Enabled Smartphones, in: Transportation Research Board 90th Annual Meeting.

City and County of San Francisco, 2007. San Francisco Charter: Transit-First Policy. doi:10.1017/CBO9781107415324.004

City and County of San Francisco, 2005. Bicycle Plan Policy Framework adoption and related General Plan amendments.

Corey Canapary & Galanis, 2011. San Francisco Bicycling Study Report. San Francisco, CA.

Dill, J., Carr, T., 2003. Bicycle Commuting and Facilities in Major U.S. Cities: If You Build Them, Commuters Will Use Them. Transp. Res. Rec. 1828, 116–123. doi:10.3141/1828-14

Ferenchak, N.N., Marshall, W.E., 2016. Relative (In)Effectiveness of Bicycle Sharrows on Ridership and Safety Outcomes, in: Transportation Research Board 95th Annual Meeting. Washington, D.C.

FHWA, 2009. 2009 National Household Travel Survey. US Department of Transportation.

Figliozzi, M., Blanc, B., 2015. Evaluating the Use of Crowdsourcing as a Data Collection Method for Bicycle Performance Measures and Identification of Facility Improvement Needs.

Furth, P., 2012. Bicycling Infrastructure for Mass Cycling: A Transatlantic Comparison, in: Pucher, J., Buehler, R. (Eds.), City Cycling. MIT Press, Cambridge. MA, pp. 105–140.

Garrard, J., Rissel, C., Bauman, A.E., 2012. Health Benefits of Cycling, in: Pucher, J., Buehler, R. (Eds.), City Cycling. MIT Press, Cambridge, MA, pp. 31–56.

Gordon, R., Tucker, J., 2010. Ruling paves way for San Francisco bike lanes. San Fr. Chron.

Greenfield, J., 2012. Can Chicago reach 30 miles of “green lanes” before the snow flies? [WWW

Document]. Grid Chicago. URL http://gridchicago.com/2012/can-chicago-reach-30-miles-of-green-lanes-before-the-snow-flies/

Heinen, E., van Wee, B., Maat, K., 2010. Commuting by Bicycle: An Overview of the Literature. Transp. Rev. 30, 59–96. doi:10.1080/01441640903187001

Hood, J., Sall, E., Charlton, B., 2011. A GPS-based bicycle route choice model for San Francisco, California. Transp. Lett. Int. J. Transp. Res. 3, 63–75. doi:10.3328/TL.2011.03.01.63-75

Iacono, M., Krizek, K.J., El-Geneidy, A., 2010. Measuring non-motorized accessibility: issues, alternatives, and execution. J. Transp. Geogr. 18, 133–140. doi:10.1016/j.jtrangeo.2009.02.002

Ibeas, A., Bordagaray, M., Dios, J. De, 2013. Modelling the effects of pro bicycle infrastructure and policies: towards sustainable urban mobility. Working Paper.

James, S., 2011. A Wobbling Effort to Accommodate Cyclists. New York Times.

Krizek, K., 2006. Two Approaches to Valuing Some of Bicycle Facilities' Presumed Benefits. J. Am. Plan. Assoc. 72, 309–320. doi:10.1080/01944360608976753

Krizek, K.J., Barnes, G., Pointexter, G., Mogush, P., Thompson, K., Levinson, D., Tilahun, N., Loutzenheiser, D., Kidson, D., Hunter, W., Tharpe, D., Gillenwater, Z., Killingsworth, R., 2006. NCHRP Report 552: Guidelines for analysis of investments in bicycle facilities. Transportation Research Board.

Larsen, J., Patterson, Z., El-Geneidy, A.M., 2013. Build it. But where? The use of geographic information systems in identifying locations for new cycling infrastructure. Int. J. Sustain. Transp. 7, 299–317.

Li, Y., Huang, Q., Kerber, M., Zhang, L., Guibas, L., 2013. Large-Scale Joint Map Matching of GPS Traces. Proc. 21st ACM SIGSPATIAL Int. Conf. Adv. Geogr. Inf. Syst. 214–223. doi:10.1145/2525314.2525333

Madsen, T., Davis, B., Del Chiaro, B., 2010. Clean Cars in California: Four Decades of Progress in the Unfinished Battle to Clean Up Our Air.

McNeil, N., Monsere M, C., Dill, J., 2014. The Influence of Bike Lane Buffer Types on Perceived Comfort and Safety of Bicyclists and Potential Bicyclists, in: Transportation Research Board 94th Annual Meeting.

Mobility Investment Priorities, 2014. Cycle Tracks. Austin, TX.

National Association of City Transportation Officials, 2014. Urban Bikeway Design Guide, Second Edition, 2nd ed. Washington, D.C.

Pucher, J., Dill, J., Handy, S., 2010. Infrastructure, programs, and policies to increase bicycling: An international review. Prev. Med. (Baltim). 50, S106–S125. doi:10.1016/j.ypmed.2009.07.028

Sall, E., 2013. Truth in Data. Use in Data. Faith in Analysts.

San Francisco Municipal Transportation Agency, 2015. Annual Bicycle Count Survey 2014. San

Francisco, CA.

San Francisco Municipal Transportation Agency, 2013a. SFMTA 2013 Bicycle Count. San Francisco, CA.

San Francisco Municipal Transportation Agency, 2013b. SFMTA Strategic Plan. San Francisco, CA.

San Francisco Municipal Transportation Agency, 2013c. SFMTA Bicycle Strategy. San Francisco, CA.

San Francisco Municipal Transportation Agency, 2012. 2012 San Francisco State of Cycling Report. San Francisco, CA.

San Francisco Municipal Transportation Agency, 2009. San Francisco Bicycle Plan. San Francisco, CA.

San Francisco Municipal Transportation Agency, 2008. City of San Francisco 2008 Bicycle Count Report. San Francisco, CA.

San Francisco Municipal Transportation Agency Bicycle Program Staff, 2010. January 2010 SFMTA Bicycle Program Report to the San Francisco Bicycle Advisory Committee (BAC). San Francisco, CA.

Schuessler, N., Axhausen, K.W., 2009. Map-matching of GPS traces on high-resolution navigation networks using the Multiple Hypothesis Technique (MHT), October.

Sperling, D., Gordon, D., 2008. Two Billion Cars: Driving Toward Sustainability. Oxford University Press, New York City.

Superior Court of California - County of San Francisco, 2006. Order Granting Petition for Peremptory Writ of Mandate.

Tilahun, N.Y., Levinson, D.M., Krizek, K.J., 2007. Trails, lanes, or traffic: Valuing bicycle facilities with an adaptive stated preference survey. Transp. Res. Part A Policy Pract. 41, 287–301. doi:10.1016/j.tra.2006.09.007

Winters, M., Teschke, K., 2010. Route Preferences Among Adults in the Near Market for Bicycling: Findings of the Cycling in Cities Study. Am. J. Heal. Promot. 25, 40–47.

Winters, M., Teschke, K., Grant, M., Setton, E.M., Brauer, M., 2011. How Far Out of the Way Will We Travel? Transp. Res. Rec. J. Transp. Res. Board 2190, 1–10. doi:10.3141/2190-01

Wolf, J., Bachman, W., Oliveira, M., Auld, J., Mohammadian, A. (Kouros), Vovsha, P., Zmud, J., 2014a. Applying GPS Data to Understand Travel Behavior, Volume II: Guidelines. Washington, D.C.

Wolf, J., Bachman, W., Oliveira, M.G.S., Auld, J., Mohammadian, A. (Kouros), Vovsha, P., 2014b. Applying GPS Data to Understand Travel Behavior, Volume I: Background, Methods, and Tests. Washington, D.C.

Yang, J., Kang, S., Chon, K., 2005. The map matching algorithm of GPS data with relatively long polling time intervals. J. East. Asia Soc. Transp. Stud. 6, 2561–2573.

doi:10.11175/easts.6.2561

Zorn, L., Sall, E., Bomberg, M., 2012. Completing the Cycle: Incorporating CycleTracks into SF-CHAMP, in: 4th TRB Conference on Innovations in Travel Modeling.

# Appendix A – Detailed Data Cleaning Methods

## Data Process Methodology

### *Section 1: Processing CycleTracks Data (Database Development)*

The data collected from the CycleTracks application was converted into csv (comma separated value) files, and processed using the statistical programming language R. In this process we created two databases: a "Global Points Table" that contains the GPS point data for each recorded bike trip, and a "Global Records Table" that contains the corresponding user information. We aggregated the data collected from phase 1 (beginning November of 2009, and ending March of 2010) and phase 2 (beginning November of 2013, and ending March of 2014) within these databases, and joined them through the 'trip ID' primary key--a common field that assigns a unique integer value for each recorded bike trip.

*Creating the Global Records Table*

The global records table contains all of the unique trip IDs recorded during the study period (even those that were recorded outside of the two collection phases), along with the corresponding user information. This was created using the following process:

1. Categorize the points csv file (named as d.p12.p) in their respective phase by creating a new field, "phase". This is done by iterating through each GPS point's recorded date, where trips recorded between November of 2009 and March of 2010, are demarcated as "phase 1", otherwise, they will be in "phase 2".
2. Create a vector containing all the duplicated trip IDs in the points table (named as repeats_points). Create another table from the main points csv file (d.p12.p), with all of the duplicate trip IDs removed (using repeats_points vector, now creating the points_norepeats table).
3. Merge the points table that now have no duplicate trip IDs (points_norepeats) with the trip records table (d.p12.r), using trip_ID as the primary key. This will create a global records table (global_records) that has all the unique trip IDs, and corresponding user information (including the user ID). Create and populate a "phase" field inside the new global records table.
4. Merge the global records table (global_records) with the survey participant records table (d.sp.r) to collect the trip IDs that were not recorded in either of the phase 1 or phase 2 collection periods. These trip IDs will be tagged as "neither" in the "phase" field. Remove any duplicate trip IDs that were collected from the merge.
5. The final global records table will contain all of the unique trip IDs recorded during the study period, along with their corresponding user information.

*Creating the Global Points Table*

The global points table contains all of the bike trips recorded during the study period, characterized as a series of GPS points collected at regular intervals during the individual trip. This was created using the following process:

1. Subset a vector (neither_tid) that contains all of the trip IDs from the global records table (global_records), tagged as "neither" in the "phase" field.
2. Subset a table (neither_sp_p) that contains all of the GPS points with corresponding trip IDs tagged as "neither". This table will collect all of the GPS points that were collected outside of the two collection phases.
3. Create all the fields found in the main points table (d.p21.p) inside the prior subsetted table (neither_sp_p). This will allow us to perform an rbind (right merge), which requires the two tables to have the same number of fields.
4. Perform an rbind (right merge) between the main points table and the subsetted table (neither_sp_p). This will create the global gps table that contains all of the GPS points ever recorded (including the GPS points recorded outside the two collection phases).
5. Perform a merge between the global gps table (global_gps) and global records table (using trip id as the primary key) to collect the user ID column, which will correspond to all of the trip IDs.
6. The final global GPS table will now have all of the bike routes recorded during the study period, along with a corresponding user ID.

## *Section 2: San Francisco Bike Facilities Network*

*Identifying Bike Facilities*

We first subjected the base road network, provided by the SFCTA to a topology validation--a process that identifies errors in the network's geometries. Upon ensuring connectivity throughout the SFCTA roads layer, all designated bike facilities (bike lanes, bike routes, and bike paths) were manually identified using data provided by the SFMTA bikeway network. There were, however, occasional inconsistencies in the placement of particular bike facility links (Figure 1). Therefore, only the bike facilities that had a corresponding SFCTA network link were identified and demarcated.

*Categorizing Bike Infrastructure and Characteristics*

The following fields were created inside the SFCTA road network's attribute table:

1. BIKE_CLASS: a value from 0 to 3, where 0 indicates a non-bike facility, 1 as a bike path, 2 as a bike lane, and 3 as a bike route.
2. YEAR_INSTALLED: indicates the year a particular link is designated as a bike facility
3. SURFACE_TREATMENT: indicates additional treatments on the bike facility surface,

including “Green Paint” and “Green Sharrow”.

4. INNOVATIVE: “Separated Bikeway”, “Green Wave”, “Buffered Bike Lane”, “Bike Boulevard”, “Back-In Angle Parking”
5. SHARROW: a value of 0 or 1, where 1 indicates the presence of a sharrow.
6. FACILITY_TYPE: indicates the bike facility, including “Bike Path”, “Bike Lane”, and “Bike Route”.
7. BARRIER: “Safe-Hit Posts”, “Parking”, “Concrete Curbs”
8. COMPLETE: a value of <null> or 1, where 1 indicates that the link was manually reviewed
9. NOTES: extra notes regarding the specific link

The additional fields were populated with the data recorded in the SFMTA bikeway network dataset. Each recognized bike facility link in the SFCTA roads layer was cross-referenced with its corresponding SFMTA bikeway link, and the data for each added field was manually transferred onto the SFCTA attribute table. Several updated networks were validated using the Google Maps Street View, noting the most recent features of the selected bikeway in place. This added validation process also led to identifying novel bikeway features unrecognized by the SFMTA layer, and were subsequently added onto the updated network.

*Assumptions and Extrapolations*

1. The SFCTA base road network digitized a majority of two-way roads as two separate links stacked on top of each other. In an effort to differentiate between the stacked links, an additional “direction” field was produced for the entire network, which provided a compass direction for each link. This was important for identifying two-way roads that featured a different bikeway facility depending on the direction.
2. There were certain cases where a single SFCTA network link existed for a particular bike facility, while the SFMTA recognized the same link as two different bike facilities. This prompted us to subjectively select the “best” facility to represent the chosen link, using local contextual knowledge on ridership choice. For example, the SFMTA recognizes the northwest-bound bikeway along the Embarcadero Waterfront as a bike lane and bike path. The bike path, assuming it’s referring to the Ferry Building Promenade, is not an ideal method of traversing the region via bike due to the promenade being a multi-use path for pedestrians and bicyclists. These series of links were therefore characterized as bike lanes to capture bicyclist behavior in this local context.

## *Section 3: Data Cleaning and Map Matching Thresholds and Parameters*

**Data Cleaning Parameters**

*# maxCalculatedSpeed (mps) set at reasonable speed because of good accuracy (except when 0)*

maxCalculatedSpeed <- 16

```
# maxCalculatedAccel (mps^2) set at reasonable speed because of good accuracy (except when 0)
maxCalculatedAccel <- 1

# maxCalculatedCrowSpeed (mps) set really high because of poor accuracy
maxCalculatedCrowSpeed <- 50

# maxSpeedDifference (mps) set really high because of poor accuracy
maxSpeedDifference <- 40

# maxDistance (m) set low based on empirical histogram
maxDistance <- 150

# maxhAccuracy (m) set at level to discard very noisy data
maxhAccuracy <- 200
```

**Map Matching Parameters**

```
# pAddedSearchRadius (m): used to expand the search radius of the hAccuracy in case
# no links are within the GPS point
pAddedSearchRadius = 30

# pTopologyTolerance (m): snapping tolerance where it is assumed links connect
pTopologyTolerance = 1

# pSparsityThreshold (gps points / km): sparsity tolerance to consider the
# there to be enough GPS data to accurately estimate a trip
pSparsityTolerance = 4

# pTripDistanceTolerance (m): distance tolerance to a string of GPS points
# per trip. Calculated as the diagonal of the bounding box (Envelope)
```

pTripDistanceTolerance = 200

*# pStartDistance (m): distance to consider a gps point close enough to a link*
*# to select that link as the start link. This ensures that when a sub network is*
*# created (when multiple disjoint network clusters are created because GPS points*
*# jump around), only the GPS points close to the largest subnetwork are considered*
*# viable starting points.*
pStartDistance = 50

# Appendix B – CycleTracks Survey Instrument

# 2014 CycleTracks Survey

## Introduction

**Page description:**

ID 236

Welcome to the 2014 CycleTracks Survey! Thank you for taking the time to help us. This survey, combined with the routes you've provided through the CycleTracks app, will help the San Francisco County Transportation Authority (SFCTA) to better understand the infrastructure that you prefer. This in turn helps SFCTA prioritize projects that are most beneficial to bicyclists.

This survey takes 15-20 minutes to complete. As an incentive for participation, you will be entered in a raffle to win a *$100 debit gift card*. Due to the small number of participants in this study, you have an excellent chance of winning this prize!

**By completing this survey you agree that the SFCTA can combine your survey results with your CycleTracks records and provide this anonymous data to the UC Davis research team.**

## Your Bicycle

**Page description:**

DATA Shortname / Alias: **UsualBikeType** Variable name: **UsualBikeType**

ID 37

1. What type of bicycle do you usually ride for transportation (not for exercise)?

- Road
- City/Traditional
- Mountain
- Cargo
- Recumbent
- Hybrid
- Cruiser
- Folding
- Fixed-gear
- Touring
- Not Sure
- Other

DATA Shortname / Alias: **MotoBike** Variable name: **MotoBike**

ID 39

2. Is your usual bicycle motorized?

- No
- Yes, electric
- Yes, gas/diesel

DATA Shortname / Alias: **FacilityComfort**

ID 40

3. In general, how comfortable would you be riding a bicycle in the following kinds of streets in daylight and good weather? (Move your mouse over the "definition" label for the definition of 'protected bicycle lane')

| | Comfortable | Uncomfortable, but I'd ride there anyway | Uncomfortable, and I wouldn't ride on it |
|---|---|---|---|
| An off-street bicycle path | ○ | ○ | ○ |
| A quiet residential street with a 25 mph speed limit | ○ | ○ | ○ |
| A **two-lane** (one lane in either direction) local street with a 35 mph speed limit and on-street parking, ***without a bicycle lane*** | ○ | ○ | ○ |
| A **two-lane** (one lane in either direction) local street with a 35 mph speed limit and on-street parking, ***with a bicycle lane*** | ○ | ○ | ○ |
| A **two-lane** (one lane in either direction) local street with a 35 mph speed limit and on-street parking, ***with a protected bicycle lane*** **(definition)** | ○ | ○ | ○ |
| A **four-lane** (two lanes in either direction) street with a 35 mph speed limit and on-street parking, ***without a bicycle lane*** | ○ | ○ | ○ |
| A **four-lane** (two lanes in either direction) street with a 35 mph speed limit and on-street parking, ***with a bicycle lane*** | ○ | ○ | ○ |
| A **four-lane** (two lanes in either direction) street with 35 mph speed limit and on-street parking ***with a protected bicycle lane*** **(definition)** | ○ | ○ | ○ |

## Purposes

**Page description:**

DATA Shortname / Alias: **BicycleLengthPurpose**

ID 140

4. How many years of experience do you have riding your bicycle for the following purposes? (Move your mouse over the "definition" label to reveal the purpose definitions)

| | More than five years | Between one and five years | Less than one year | Never (or Not Applicable) |
|---|---|---|---|---|
| Commute (definition) | ○ | ○ | ○ | ○ |
| School (definition) | ○ | ○ | ○ | ○ |
| Work-related (definition) | ○ | ○ | ○ | ○ |
| Exercise (definition) | ○ | ○ | ○ | ○ |
| Social (definition) | ○ | ○ | ○ | ○ |
| Shopping (definition) | ○ | ○ | ○ | ○ |
| Errand (definition) | ○ | ○ | ○ | ○ |
| Any other purpose (definition) | ○ | ○ | ○ | ○ |

DATA Shortname / Alias: **BicycleFreqPurpose**

ID 11

5. Please choose the category that best describes how often you ride your bicycle for each travel purpose within San Francisco.

| | Daily | Several times per week | Several times per month | Once per month or less | Never (or Not Applicable) |
|---|---|---|---|---|---|
| Commute (definition) | ○ | ○ | ○ | ○ | ○ |
| School (definition) | ○ | ○ | ○ | ○ | ○ |
| Work-related (definition) | ○ | ○ | ○ | ○ | ○ |
| Exercise (definition) | ○ | ○ | ○ | ○ | ○ |
| Social (definition) | ○ | ○ | ○ | ○ | ○ |
| Shopping (definition) | ○ | ○ | ○ | ○ | ○ |
| Errand (definition) | ○ | ○ | ○ | ○ | ○ |
| Any other purpose (definition) | ○ | ○ | ○ | ○ | ○ |

DATA Shortname / Alias: **BikeUnavailable** Variable name: **BikeUnavailable**

ID 193

6. If you could not ride your bicycle during commuting hours (e.g. your bicycle was broken, stolen, in the shop, etc.), how would you travel to work? (Select "I do not ride my bicycle for my commute" if you do not ride your bicycle to work or to school)

- Walk
- Skate or skateboard
- Motorcycle or scooter
- Drive alone in a car (or other vehicle)
- Carpool
- Bus (MUNI, AC Transit, etc.)
- Light rail, streetcar, or cable car
- BART
- Caltrain
- Taxi (includes Uber, Lyft, Sidecar, and other taxi and car-sharing services)
- Bike share
- Borrowing a bike
- I would postpone traveling to work, by working from home, fixing my bike, or calling in sick.
- **I do not ride my bicycle for my commute**

## Routes

**Page description:**

ID 238

**Think about the <u>place</u> you ride your bicycle to most often in San Francisco and keep this destination in mind for the following questions:**

DATA Shortname / Alias: **FavePlacePurpose** Variable name: **FavePlacePurpose**

ID 178

7. What activity do you do at this **<u>place</u>**?

- Commute (definition)
- School (definition)
- Work-related (definition)
- Exercise (definition)
- Social (definition)
- Shopping (definition)
- Errand (definition)
- Other

DATA Shortname / Alias: **BicycleAttire** Variable name: **BicycleAttire**

ID 255

8. On a normal day (sunny, light fog, 65 degrees F), which of the following best describes your attire when bicycling to this **<u>place</u>**?

- Everyday clothes (i.e. the clothes I was already wearing).
- Bicycle-specific clothes (e.g. athletic, alternative "riding" clothes).

DATA Shortname / Alias: **FavePlaceUsualRoute** Variable name: **FavePlaceUsualRoute**

ID 179

9. When bicycling to this **place**, do you have a specific route that you use most of the time? *

- Yes
- No

LOGIC Dynamically shown if "When bicycling to this **place**, do you have a specific route that you use most of the time?" = Yes

DATA Shortname / Alias: **UsualRouteDuration**

ID 23

10. Approximately how long have you used your usual route to this **place**?

Year(s)

Month(s)

LOGIC Dynamically shown if "When bicycling to this **place**, do you have a specific route that you use most of the time?" = Yes

DATA Shortname / Alias: **DivertFreq** Variable name: **DivertFreq**

ID 25

11. About how often do you use other routes to go to this **place**?

- fewer than 1 in 10 trips
- 1 in 10 trips
- 2 in 10 trips
- 3 in 10 trips
- 4 in 10 trips
- 5 in 10 trips

LOGIC Dynamically shown if "When bicycling to this **place**, do you have a specific route that you use most of the time?" = Yes

DATA Shortname / Alias: **WhyDivert**

ID 29

12. Why do you use other routes to this **place**? Please select all that apply:

- ☐ Weather
- ☐ Convenience
- ☐ Additional stops on the trip
- ☐ Personal safety
- ☐ Traffic safety
- ☐ Traffic congestion
- ☐ Social reasons
- ☐ For fun or novelty
- ☐ Pavement quality
- ☐ Construction activities
- ☐ Aesthetics
- ☐ Not pressed for time
- ☐ Pressed for time
- ☐ Not applicable (I always use the same route)
- ☐ Other ____________

LOGIC Dynamically shown if "When bicycling to this **place**, do you have a specific route that you use most of the time?" = Yes

DATA Shortname / Alias: **FavePlaceOtherRoutes** Variable name: **FavePlaceOtherRoutes**

ID 181

13. Approximately how many routes do you use to get to this **place**?

- 1
- 2 to 3
- 4 to 5
- 6 to 10
- more than 10

LOGIC Dynamically shown if "When bicycling to this **place**, do you have a specific route that you use most of the time?" = Yes

ID 280

14. Since you started regularly bicycling to this **place**, have any of the following bicycle infrastructure changes or bicycle amenities been **added to the route(s)** you use to get to this **place**? Please select all that apply:

Sharrow (shared lane marking)

Green Sharrow

Green Sharrow

Bicycle Lane

Bicycle Lane

Green Bicycle Lane

Green Bicycle Lane

Buffered Bicycle Lane

Buffered Bicycle Lane

Protected Bicycle Lane with Soft Hit Posts

Parking Protected Bicycle Lane

Physically Separated Bicycle Lane

Off-Street Bicycle Path

Off Street Bicycle Path

Green Wave Signal Timing

Green Wave Signal Timing

Bicycle Box (space to queue in front of cars at red lights)

Bay Area Bike Share

Bay Area Bike Share

Unrestricted Bicycle Space on BART

Bicycle Rack

Bicycle Rack

Bicycle Parking Corral

Bicycle Parking Corral

Bicycle Locker

Bicycle Locker

Transit Bicycle Parking Station

Indoor Bicycle Parking

Indoor Bicycle Parking

Workplace Shower Facility

Workplace Shower Facility

None of the Above

None of the Above

LOGIC Dynamically shown if "When bicycling to this **place**, do you have a specific route that you use most of the time?" = Yes

ID 291

15. Since you started regularly bicycling to this **place**, have any of the following bicycle infrastructure changes or bicycle amenities **caused you to permanently change your usual route** to this **place**? Please select all that apply:

- Sharrow (shared lane marking)
- Green Sharrow
- Bicycle Lane
- Green Bicycle Lane
- Buffered Bicycle Lane
- Protected Bicycle Lane with Soft Hit Posts
- Parking Protected Bicycle Lane
- Physically Separated Bicycle Lane
- Off Street Bicycle Path
- Green Wave Signal Timing
- Bicycle Box (space to queue in front of cars at red lights)
- Bay Area Bike Share
- Unrestricted Bicycle Space on BART
- Bicycle Rack
- Bicycle Parking Corral
- Bicycle Locker
- Transit Bicycle Parking Station
- Indoor Bicycle Parking
- Workplace Shower Facility
- None of the Above

LOGIC Dynamically shown if "When bicycling to this **place**, do you have a specific route that you use most of the time?" = No

DATA Shortname / Alias: **FavePlaceWhyNoUsualRoute**

ID 258

16. Why do you not have a usual route to this **place**? Please select all that apply:

- ☐ Weather
- ☐ Convenience
- ☐ Additional stops on the trip
- ☐ Personal safety
- ☐ Traffic safety
- ☐ Traffic congestion
- ☐ Social reasons
- ☐ For fun or novelty
- ☐ Not familiar with the options available
- ☐ Have not found the best route
- ☐ Pavement quality
- ☐ Construction activities
- ☐ Aesthetics
- ☐ Not pressed for time
- ☐ Pressed for time
- ☐ No reason
- ☐ Other

LOGIC Dynamically shown if "When bicycling to this **place**, do you have a specific route that you use most of the time?" = No

ID 293

17. Have any of the following bicycle infrastructure changes or bicycle amenities been **added to the routes** you use to get to this **place**? Please select all that apply:

Sharrow (shared lane marking)

Green Sharrow

Green Sharrow

Bicycle Lane

Bicycle Lane

Green Bicycle Lane

Green Bicycle Lane

Buffered Bicycle Lane

Buffered Bicycle Lane

Protected Bicycle Lane with Soft Hit Posts

Parking Protected Bicycle Lane

Physically Separated Bicycle Lane

Off-Street Bicycle Path

Off Street Bicycle Path

Green Wave Signal Timing

Green Wave Signal Timing

Bicycle Box (space to queue in front of cars at red lights)

Bay Area Bike Share

Bay Area Bike Share

Unrestricted Bicycle Space on BART

Bicycle Rack

Bicycle Rack

Bicycle Parking Corral

Bicycle Parking Corral

Bicycle Locker

Bicycle Locker

Transit Bicycle Parking Station

Indoor Bicycle Parking

Indoor Bicycle Parking

Workplace Shower Facility

Workplace Shower Facility

None of the Above

None of the Above

LOGIC Dynamically shown if "When bicycling to this **place**, do you have a specific route that you use most of the time?" = No

ID 294

18. *In general*, have any of the following bicycle infrastructure changes or bicycle amenities caused you to **change which routes you use** to get to this **place**? Please select all that apply:

| | | |
|---|---|---|
| Sharrow (shared lane marking) | Green Sharrow<br>Green Sharrow | Bicycle Lane<br>Bicycle Lane |
| Green Bicycle Lane<br>Green Bicycle Lane | Buffered Bicycle Lane<br>Buffered Bicycle Lane | Protected Bicycle Lane with Soft Hit Posts |
| Parking Protected Bicycle Lane | Physically Separated Bicycle Lane | Off-Street Bicycle Path<br>Off Street Bicycle Path |
| Green Wave Signal Timing<br>Green Wave Signal Timing | Bicycle Box (space to queue in front of cars at red lights) | Bay Area Bike Share<br>Bay Area Bike Share |
| Unrestricted Bicycle Space on BART | Bicycle Rack<br>Bicycle Rack | Bicycle Parking Corral<br>Bicycle Parking Corral |
| Bicycle Locker<br>Bicycle Locker | Transit Bicycle Parking Station | Indoor Bicycle Parking<br>Indoor Bicycle Parking |
| Workplace Shower Facility<br>Workplace Shower Facility | None of the Above<br>None of the Above | |

LOGIC Dynamically shown if "When bicycling to this **place**, do you have a specific route that you use most of the time?" = Yes or "When bicycling to this **place**, do you have a specific route that you use most of the time?" = No

ID 295

19. ***In general***, have any of the following bicycle infrastructure changes or bicycle amenities caused you to **change where you ride your bicycle in San Francisco**? Please select all that apply:

| | | |
|---|---|---|
| Sharrow (shared lane marking) | Green Sharrow | Bicycle Lane |
| Green Bicycle Lane | Buffered Bicycle Lane | Protected Bicycle Lane with Soft Hit Posts |
| Parking Protected Bicycle Lane | Physically Separated Bicycle Lane | Off Street Bicycle Path |
| Green Wave Signal Timing | Bicycle Box (space to queue in front of cars at red lights) | Bay Area Bike Share |
| Unrestricted Bicycle Space on BART | Bicycle Rack | Bicycle Parking Corral |
| Bicycle Locker | Transit Bicycle Parking Station | Indoor Bicycle Parking |
| Workplace Shower Facility | None of the Above | |

**Live and Work**

**Page description:**

DATA Shortname / Alias: **EmploymentStatus** Variable name: **EmploymentStatus**

ID 7

20. What is your current employment status?

- ○ Full-time
- ○ Part-time
- ○ Non-employed
- ○ Retired

ID 281

21. The following questions ask how long you have lived, worked, or attended school in San Francisco:

| | Not applicable | Year(s) | Month(s) |
|---|---|---|---|
| Approximately how long have you currently been living in San Francisco? | ☐ | | |
| Approximately how long have you currently been working in San Francisco? | ☐ | | |
| Approximately how long have you currently been attending school in San Francisco? | ☐ | | |

## Bicycle Parking

**Page description:**

DATA Shortname / Alias: **BicycleParkingPreferences**

ID 239

22. We'd like to ask about your attitudes and preferences with respect to **day-time bicycle parking in San Francisco**. There are no right or wrong

answers; we want only your true opinions. To what extent do you agree or disagree with the following statements?

| | Strongly disagree | Disagree | Neutral | Agree | Strongly agree |
|---|---|---|---|---|---|
| I think about where I will park my bicycle before I depart. | ○ | ○ | ○ | ○ | ○ |
| I choose my destination based on the presence of bicycle parking. | ○ | ○ | ○ | ○ | ○ |
| Bicycle parking influences whether I will ride my bicycle or choose a different means of transportation. | ○ | ○ | ○ | ○ | ○ |
| I feel comfortable parking in non-designated bicycle parking locations (e.g. tree, railing, sign-post, etc.). | ○ | ○ | ○ | ○ | ○ |
| I feel comfortable parking in bicycle racks. | ○ | ○ | ○ | ○ | ○ |
| I feel comfortable parking in a bicycle locker or cage (i.e. bicycle is completely enclosed). | ○ | ○ | ○ | ○ | ○ |
| There are usually no bicycle racks nearby where I want to park. | ○ | ○ | ○ | ○ | ○ |
| Bicycle racks nearby are usually full where I want to park. | ○ | ○ | ○ | ○ | ○ |
| I usually end up parking more than one city block from my ultimate destination. | ○ | ○ | ○ | ○ | ○ |
| I need to use a heavy duty bicycle lock to protect my bicycle from being stolen when parking in public. | ○ | ○ | ○ | ○ | ○ |
| I worry about my bicycle being stolen when I park my bicycle outside in my neighborhood. | ○ | ○ | ○ | ○ | ○ |
| I worry about my bicycle being stolen when I park my bicycle outside in other neighborhoods. | ○ | ○ | ○ | ○ | ○ |

**Page description:**

DATA Shortname / Alias: **TravelAttitudes**

ID 79

23. We'd like to ask about your attitudes and preferences with respect to travel and the environment. There are no right or wrong answers; we want only your true opinions. To what extent do you agree or disagree with the following statements?

| | Strongly disagree | Disagree | Neutral | Agree | Strongly agree |
|---|---|---|---|---|---|
| Travel time is generally wasted time. | ○ | ○ | ○ | ○ | ○ |
| I like riding a bicycle. | ○ | ○ | ○ | ○ | ○ |
| Environmental concerns affect the choices I make about my daily travel. | ○ | ○ | ○ | ○ | ○ |
| I ride my bicycle for transportation as often as I can. | ○ | ○ | ○ | ○ | ○ |
| I like to vary my route to get to the same place. | ○ | ○ | ○ | ○ | ○ |
| I need a car to do many of the things I like to do. | ○ | ○ | ○ | ○ | ○ |
| I ride my bicycle because I can reliably estimate how long it will take me to get to my destination. | ○ | ○ | ○ | ○ | ○ |
| I like driving. | ○ | ○ | ○ | ○ | ○ |
| I feel safe riding my bicycle in San Francisco traffic. | ○ | ○ | ○ | ○ | ○ |
| I ride my bicycle because the transit system is overcrowded. | ○ | ○ | ○ | ○ | ○ |
| My route choices are influenced by personal safety concerns (e.g. crime). | ○ | ○ | ○ | ○ | ○ |
| I like using public transit. | ○ | ○ | ○ | ○ | ○ |
| I ride my bicycle because it saves me money. | ○ | ○ | ○ | ○ | ○ |
| Bicycling is generally more convenient than taking transit. | ○ | ○ | ○ | ○ | ○ |
| Bicycle traffic laws are adequately enforced. | ○ | ○ | ○ | ○ | ○ |
| My personal health affects the choices I make about my daily travel. | ○ | ○ | ○ | ○ | ○ |

## Lifestyle

**Page description:**

DATA Shortname / Alias: **FriendsCycle**

ID 206

24. How often do you bicycle to meet up with your friends within San Francisco for the following purposes?

| | Several times a week | Once a week | A few times per month | Once a month | Less than once a month | Never |
|---|---|---|---|---|---|---|
| Going to a coffee shop | ○ | ○ | ○ | ○ | ○ | ○ |
| Going to a bar | ○ | ○ | ○ | ○ | ○ | ○ |
| Going to the park | ○ | ○ | ○ | ○ | ○ | ○ |
| Going to any other type of social gathering | ○ | ○ | ○ | ○ | ○ | ○ |

DATA Shortname / Alias: **CoffeePref** Variable name: **CoffeePref**

ID 102

25. If you had a choice between the following coffee options, which would you choose?

- ○ Starbucks
- ○ Peet's
- ○ Ritual Roasters / Blue Bottle / Philz / Four Barrel
- ○ None. I don't like coffee

DATA Shortname / Alias: **PhysicallyActive** Variable name: **PhysicallyActive**

ID 119

26. How much time **per week** are you physically active? By physically active we mean moderate-intensity aerobic activity such as brisk walking, jogging, bicycling, etc.

- Less than 1 hour
- Between 1 and 2 hours
- Between 2 and 3 hours
- Between 3 and 7 hours
- More than 7 hours

## General Information

**Page description:**

DATA Shortname / Alias: **CycleTracks0910** Variable name: **CycleTracks0910**

ID 104

27. Did you log a trip with CycleTracks in 2009 or 2010? *

- Yes
- No

LOGIC Dynamically shown if "Did you log a trip with CycleTracks in 2009 or 2010?" = Yes

DATA Shortname / Alias: **Email0910** Variable name: **Email0910**

ID 105

28. If you provided an email to SFCTA in 2009/10, what was it? (We are only using your email to match the trips you took back then to the trips you are taking now. This kind of before-and-after comparison provides the most robust information.)

DATA Shortname / Alias: **BikeAdvocacy**

ID 122

29. Are you a member of the San Francisco Bicycle Coalition or another bicycle advocacy group? Select all that apply:

- ☐ San Francisco Bicycle Coalition
- ☐ No
- ☐ Other ______

DATA Shortname / Alias: **YearBorn** Variable name: **YearBorn**

ID 108

30. In what year were you born (e.g. 1960)?

______

DATA Shortname / Alias: **Sex** Variable name: **Sex**

ID 107

31. What is your sex?

- ○ Female
- ○ Male

DATA Shortname / Alias: **HHAges**

ID 109

32. Please indicate how many OTHER members of your household are in each age category:

Age under six

Age 6 to 15

Age 16 to 17

Age 18 to 64

Age 65 or older

DATA Shortname / Alias: **BikeWithKids**

ID 120

33. If you have children under the age of 16, for what purposes do you bicycle with them?

- [ ] To school
- [ ] To a park
- [ ] To a friend's house
- [ ] To go shopping
- [ ] For recreation
- [ ] None
- [ ] Other

## General Information

**Page description:**

Shortname / Alias: **Education** Variable name: **Education**

ID 111

34. What is your highest completed level of education?

- No formal education
- Grade school or junior high school
- High school diploma or equivalent
- Associates degree or technical school certificates
- Four-year bachelor’s degree
- Graduate degree(s)

Shortname / Alias: **PctVehAvailable**

ID 32

35. About what percent of the time is a motor vehicle (cars, motorcycles, small trucks) available to you when you want it?

| | 0% | 20% | 40% | 60% | 80% | 100% |
|---|---|---|---|---|---|---|
| Personal motor vehicle | ○ | ○ | ○ | ○ | ○ | ○ |
| Carshare motor vehicle | ○ | ○ | ○ | ○ | ○ | ○ |

DATA Shortname / Alias: **NumHHVeh** Variable name: **NumHHVeh**

ID 124

36. How many *personal* motor vehicles (cars, motorcycles, small trucks) does your household have?

- 0
- 1
- 2
- 3
- 4
- 5+

DATA Shortname / Alias: **CrossStreets** Variable name: **CrossStreets**

ID 106

37. What intersection is nearest to your home? (This information will be kept confidential)

Your street

Nearest cross-street

ID 298

38. **ZIP Code??? City??**

- Option 1
- Option 2

DATA Shortname / Alias: **RentOwn** Variable name: **RentOwn**

ID 113

39. Do you own or rent your home? *

- Rent
- Own
- Neither (e.g. couch surfing)

LOGIC Dynamically shown if "Do you own or rent your home?" = Rent or "Do you own or rent your home?" = Neither (e.g. couch surfing)

DATA Shortname / Alias: **LiveFamilyRent** Variable name: **LiveFamilyRent**

ID 271

40. Do you live with family, a partner, or others with whom you share an income? *

- Yes
- No

LOGIC Dynamically shown if "Do you own or rent your home?" = Own

DATA Shortname / Alias: **LiveFamilyOwn** Variable name: **LiveFamilyOwn**

ID 275

41. Do you live with family, a partner, or others with whom you share an income? *

- Yes
- No

LOGIC Dynamically shown if "Do you live with family, a partner, or others with whom you share an income?" = No or "Do you live with family, a partner, or others with whom you share an income?" = No

DATA Shortname / Alias: **PersonalIncome** Variable name: **PersonalIncome**

ID 114

42. Please check the category that contains your own approximate annual income before taxes.

- Less Than $10,000
- $10,000 - $25,000
- $25,000 - $50,000
- $50,000 - $75,000
- $75,000 - $100,000
- $100,000 - $150,000
- $150,000 - $200,000
- Greater than $200,000

LOGIC Dynamically shown if "Do you live with family, a partner, or others with whom you share an income?" = Yes or "Do you live with family, a partner, or others with whom you share an income?" = Yes

DATA Shortname / Alias: **HHIncome** Variable name: **HHIncome**

ID 272

43. Please check the category that contains your approximate annual household income before taxes.

- Less Than $10,000
- $10,000 - $25,000
- $25,000 - $50,000
- $50,000 - $75,000
- $75,000 - $100,000
- $100,000 - $150,000
- $150,000 - $200,000
- Greater than $200,000

LOGIC Dynamically shown if "Do you live with family, a partner, or others with whom you share an income?" = No

DATA Shortname / Alias: **RentPercentPersonal** Variable name: **RentPercentPersonal**

ID 115

44. About what percent of your monthly income do you spend on rent?

0% 100%

50%

LOGIC Dynamically shown if "Do you live with family, a partner, or others with whom you share an income?" = Yes

DATA Shortname / Alias: **RentPercentHH** Variable name: **RentPercentHH**

ID 273

45. About what percent of your household's monthly income do you spend on rent?

0% 100%
50%

LOGIC Dynamically shown if "Do you live with family, a partner, or others with whom you share an income?" = No

DATA Shortname / Alias: **MortgagePercentPersonal** Variable name: **MortgagePercentPersonal**

ID 116

46. About what percent of your monthly income do you spend on mortgage payments?

0% 100%
50%

LOGIC Dynamically shown if "Do you live with family, a partner, or others with whom you share an income?" = Yes

DATA Shortname / Alias: **MortgagePercentHH** Variable name: **MortgagePercentHH**

ID 274

47. About what percent of your household's monthly income do you spend on mortgage payments?

0% 100%
50%

DATA Shortname / Alias: **Race**

ID 117

48. What is your race or ethnicity? Please select all that apply:

- ☐ Black or African-American
- ☐ Asian
- ☐ Pacific-Islander or Native Hawaiian
- ☐ Hispanic
- ☐ White
- ☐ American Indian or Alaskan Native
- ☐ Other

## Thank You!

ID 1

Thank you for completing this survey!

We know your time is valuable. The results of this survey will be used to help SFCTA improve San Francisco's transportation system.