Link21

CONNECT NORTHERN CALIFORNIA

MARKET ANALYSIS REPORT

APPENDIX B: MARKET ANALYSIS TOOLS

March 2022

Link21 is committed to ensuring that information is available in appropriate alternative formats to meet the requirements of persons with disabilities. If you require an alternative version of this file, please email link21dc@link21program.org.

Prepared By:
Link21 Program Management Consultants (PMC)

DRAFT - DELIBERATIVE

INTENTIONALLY LEFT BLANK

ISSUE AND REVISION RECORD

<table>
<tr><th rowspan="2">REVISION</th><th rowspan="2">DATE</th><th rowspan="2">PREPARER(S)</th><th>PMC REVIEWER(S)</th><th>APPROVER</th><th rowspan="2">DESCRIPTION/ NOTES</th></tr>
<tr><th>BART/CCJPA REVIEWER(S)</th><th>BART/CCJPA APPROVER</th></tr>
<tr><td rowspan="2">0</td><td>10/26/2021</td><td>Derek Cheah, Steer</td><td>Elisa Tejedor, Steer</td><td>Sudhish Verma, HNTB</td><td>Working DRAFT submittal</td></tr>
<tr><td>11/29/2021</td><td>---</td><td>Andrew Tang, BART</td><td>Andrew Tang, BART</td><td>Working DRAFT submittal review</td></tr>
<tr><td rowspan="2">1</td><td>03/18/2022</td><td>Derek Cheah, Steer</td><td>Elisa Tejedor, Steer</td><td>Sudhish Verma, HNTB</td><td>DRAFT submittal</td></tr>
<tr><td></td><td></td><td></td><td></td><td>No review requirement</td></tr>
</table>

SHAREPOINT PATH

Records Center/BUSC Record Center/WPA.02/
1-BUSC-PH0-Market_Analysis_Report_Appendix_B-DRAFT

INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

FIGURES

DRAFT - DELIBERATIVE

TABLES

ACRONYMS AND ABBREVIATIONS

ACRONYM/ABBREVIATION	DEFINITION
BART	San Francisco Bay Area Rapid Transit
CCJPA	Capitol Corridor Joint Powers Authority
ACE	Altamont Corridor Express
CalEPA	California Environmental Protection Agency
CMA	congestion management agency
CoC	Communities of Concern
CS	Cambridge Systematics
FTA	Federal Transit Administration
GJT	generalized journey time
GTFS	General Transit Feed Specification
IVT	in-vehicle time
MAST	Market Analysis Spreadsheet Tool
MPO	Metropolitan Planning Organization
MTC	Metropolitan Transportation Commission
MTP/SCS	Metropolitan Transportation Plan/Sustainable Communities Strategy
OD	origin-destination
PDFH	UK Passenger Demand Forecasting Handbook
PMC	Program Management Consultants
SMART	Sonoma-Marin Area Rail Transit
VTPI	Victoria Transport Policy Institute

LINK21 PROGRAM TEAM NAMES

TEAM NAME	TEAM MEMBERS
PMC	The HNTB Team
Program Management Team (PMT)	BART/CCJPA + PMC
Consultants	Consultants supporting program identification/project selection
Link21 Team	PMT + Consultants

INTENTIONALLY LEFT BLANK

MARKET ANALYSIS TOOLS OVERVIEW

This appendix provides the methodology and details that underpin the rail potential analysis discussed at a high level in Chapter 8. Specifically, it discusses:

- **Inputs**: the data and sources which were used in the analysis
- **Hexcells, Clusters, and Hubs**: the different geographies used in the analysis
- **Rail Potential Model Development**: the regression model specification and benchmarking
- **Rail Potential Model Application**: the Market Analysis Spreadsheet Tool (MAST)

Inputs

A wide range of data for hub and cluster characteristics (hubs and clusters are discussed in the following sections) and travel characteristics were collected to enable the testing of variables in the regressions. For hub and cluster characteristics, the following were considered:

- Population and employment for each cluster
- MOSAIC population characteristics to estimate the propensity of the population in each cluster to use transit (MOSAIC data is discussed in more detail in Appendix E)
- Parking information at each cluster including parking costs and parking capacity

For travel characteristics, rail level of service characteristics were considered as well as those from the competing auto mode:

- Rail fares from each of the operators
- Rail journey times
- Rail frequency
- Rail transfers
- Rail crowding
- Auto distance
- Auto travel times (including congestion)
- Auto driving costs (including tolls and per mile costs, such as regular wear-and-tear and gas)

For the analysis, rail costs and levels of service from the rail operators, as listed in Chapter 4 of the *Market Analysis Report*, were used.

The full list of variables considered for the regression model and their information sources is shown in **Table 1**.

DRAFT - DELIBERATIVE

Table 1. List of Variables Considered for the Regression Model

A wide range of variables was tested to consider reasonable factors for rail ridership potential.

VARIABLE TYPE	VARIABLE	SOURCE
Hub and Cluster Characteristics	Population and employment (density and types of jobs)	Metropolitan Planning Organization (MPO) models
	Population characteristics including propensity to use rail	MOSAIC
	Origin and destination parking costs and capacity	Metropolitan Transportation Commission (MTC) Travel Model 1.5
Travel Characteristics	Network distance	MTC highway network
	Drive times (congested)	
	Drive costs (tolls and per mile costs)	
	Rail costs	MTC model and rail operators
	Rail journey times	Rail operators' General Transit Feed Specification (GTFS) feeds
	Rail frequency	
	Rail transfers	
	Rail crowding	Rail operators

Hexcells, Clusters, and Hubs

Hexcells, clusters, and hubs of activity are the geographic units of analysis used to understand travel demand and demographics in the Northern California Megaregion (Megaregion). This section describes them in detail as well as the methodology implemented to define these clusters.

Travel between cluster pairs was used in the rail potential model and MAST to identify unmet rail potential. **Table 2** provides definitions of several key terms.

DRAFT - DELIBERATIVE

Table 2. Key Geography Term Definitions

Various geography levels were defined to conduct the market analysis.

TERM	DEFINITION
Hexcell	Hexagonal areas of 0.5 miles in diameter that collectively cover almost the entire Megaregion.
Cluster	A group of hexcells that can be used to analyze larger geographic areas. In the process, it is defined as a hub plus its surrounding catchment area.
Hub	The central hexcell within a cluster.
Priority Populations	Combination of California Environmental Protection Agency (CalEPA) state designation of priority populations, MTC's Communities of Concern (CoC), and congestion management agency's (CMA) adjusted CoC.[1]

Hexcells

An individual hexcell has a diameter of 0.5 miles and covers an area of approximately 138 acres. There are 107,677 hexcells in the entire Megaregion. However, in many cases these hexcells cover very low-density areas that would not be realistic candidates for rail service. Inclusion of these hexcells adds significant computational burden to the analytical process. As such, a minimum initial threshold was set excluding hexcells from further analysis that do not meet the criterion. **Table 3** shows the analysis impacts of using different population thresholds per hexcell.

Table 3. Hexcell Filtering

Alternative threshold levels present different implications for the market analysis.

POPULATION THRESHOLD PER HEXCELL	IMPLIED EQUIVALENT POPULATION PER ACRE	NUMBER OF HEXCELLS ABOVE THE POPULATION THRESHOLD	PERCENT OF MEGAREGION POPULATION INCLUDED	PERCENT OF MEGAREGION PRIORITY POPULATIONS INCLUDED
No minimum	-	107,677	100%	100%
>10 people	0.07	35,456	99%	99%
>25 people	0.18	25,695	97%	98%
>50 people	0.36	18,953	95%	97%

The threshold of 25 people per hexcell (highlighted in red in Table 3) was selected as it provides significant efficiency benefits (reducing the number of included hexcells by 76%) while having minimal impact on the population covered (97% of the Megaregion's

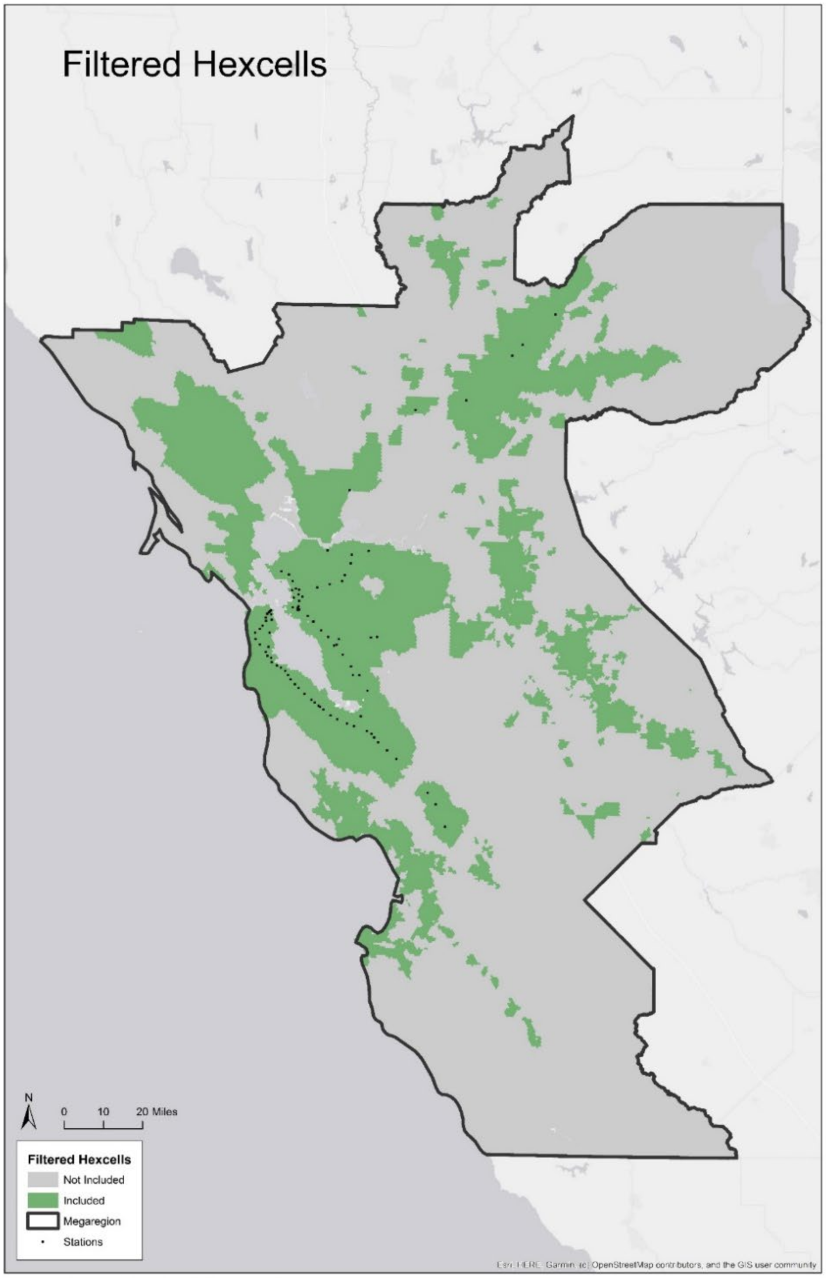
[1] This definition was updated following the market analysis; the revised definition will be incorporated in future phases of work.

DRAFT - DELIBERATIVE

population is included[2]). **Figure 1** shows where the filtered hexcells (as defined in the table above) are within the Megaregion.

Figure 1. Jobs and Population Land Use Density

This figure outlines the areas that meet the initial hexcell threshold criterion.



Source: Program Management Consultants (PMC) analysis of filtered hexcells; note that only existing stations are shown in this map

[2] 95% of the Megaregion's employment is also included.

Clusters

One of the key purposes of the market analysis is to understand the rail/transit ridership potential between origin-destination (OD) pairs across the Megaregion in order to identify those that may benefit most from additional investment. To do this, appropriate origins and destinations needed to be identified.

Various geographies exist throughout the Megaregion that could have been used (e.g., census tracts, transportation analysis zones (TAZ), county boundaries, etc.). However, each of these vary significantly in size with boundaries often determined through historic precedence, which did not lend itself well to more consistent analysis across the Megaregion. As such, a combination of hexcells and clusters was used as the units of measurement for our analysis.

The remainder of this section identifies the cluster methodology used to identify the process by which hexcells were grouped into clusters. The main considerations are presented in **Table 4**. The output is a set of clusters that collectively represent the locations across the Megaregion considered as part of the OD potential analysis.

Table 4. Criteria for Cluster Development

Various requirements were considered when developing the cluster geographic system.

REQUIREMENT	DEFINITION
Assess the potential of possible station locations and scenarios	The market analysis needs to understand the potential of possible station locations both in isolation and in combination with other locations. As such, the existence of a station/hub in a given geography should not prohibit the testing of a station/hub in any other location (within reason).[3] Therefore, some of the clusters overlap (i.e., the catchment areas of some of the hubs overlap).
Flexible approach	It is possible that new locations worthy of assessment are identified at a later stage of the work. Any methodology needs to accommodate this without the need for the work to be entirely redone.
Minimize bias	There should be a clear and consistent process for identifying cluster locations whereby the results are not overly subject to professional judgment.
Repeatability	The Link21 Program (Link21) is expected to take place over several decades. Over this time, it is possible that new data/forecasts become available. While this does not necessarily mean that any process needs to be repeated, the approach should not prohibit this.

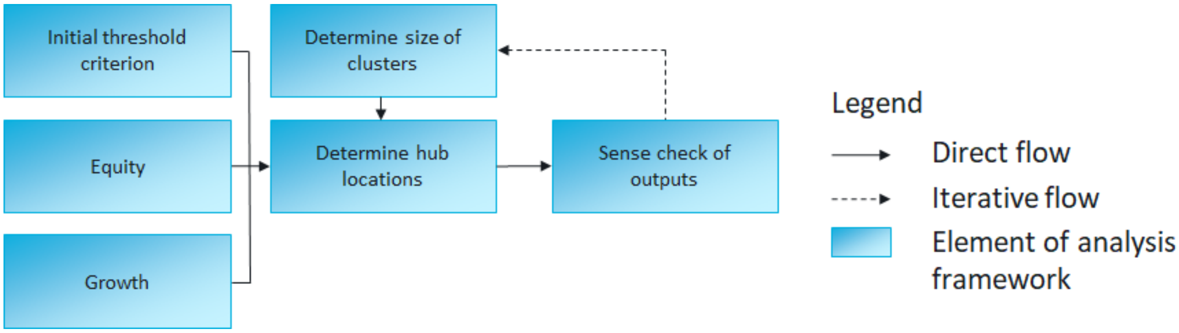
DRAFT - DELIBERATIVE

[3] Otherwise, every individual hexcell would be assessed, which would negate the purpose of clustering.

Cluster Development

Figure 2 provides an overview of the analytical framework used to develop the clusters.

Figure 2. Overview of the Clustering Methodology



The first step in the process was to determine the size of the clusters through analysis of catchments for existing stations.

Consideration was given to three items that each feed into the determination of the hub locations:

1. **Initial threshold criterion** is used to filter the data to only consider hexcells with a minimum level of density to aid in calculation efficiency.
2. **Equity** allows the consideration of impacts on priority populations.
3. **Growth** allows simultaneous identification of hub locations that satisfy criteria in both the base and future years.

The hub locations were then determined by considering relative levels of density and proximity. Finally, the locations were subject to various sense checks, including an iteration through the prior analytical steps if required.

Each of these steps are discussed in the following section.

DETERMINE THE SIZE OF CLUSTERS

The size of clusters was determined by assessing the catchments of existing stations, thereby finding a reasonable catchment within which most trips were expected for a given possible station to originate.

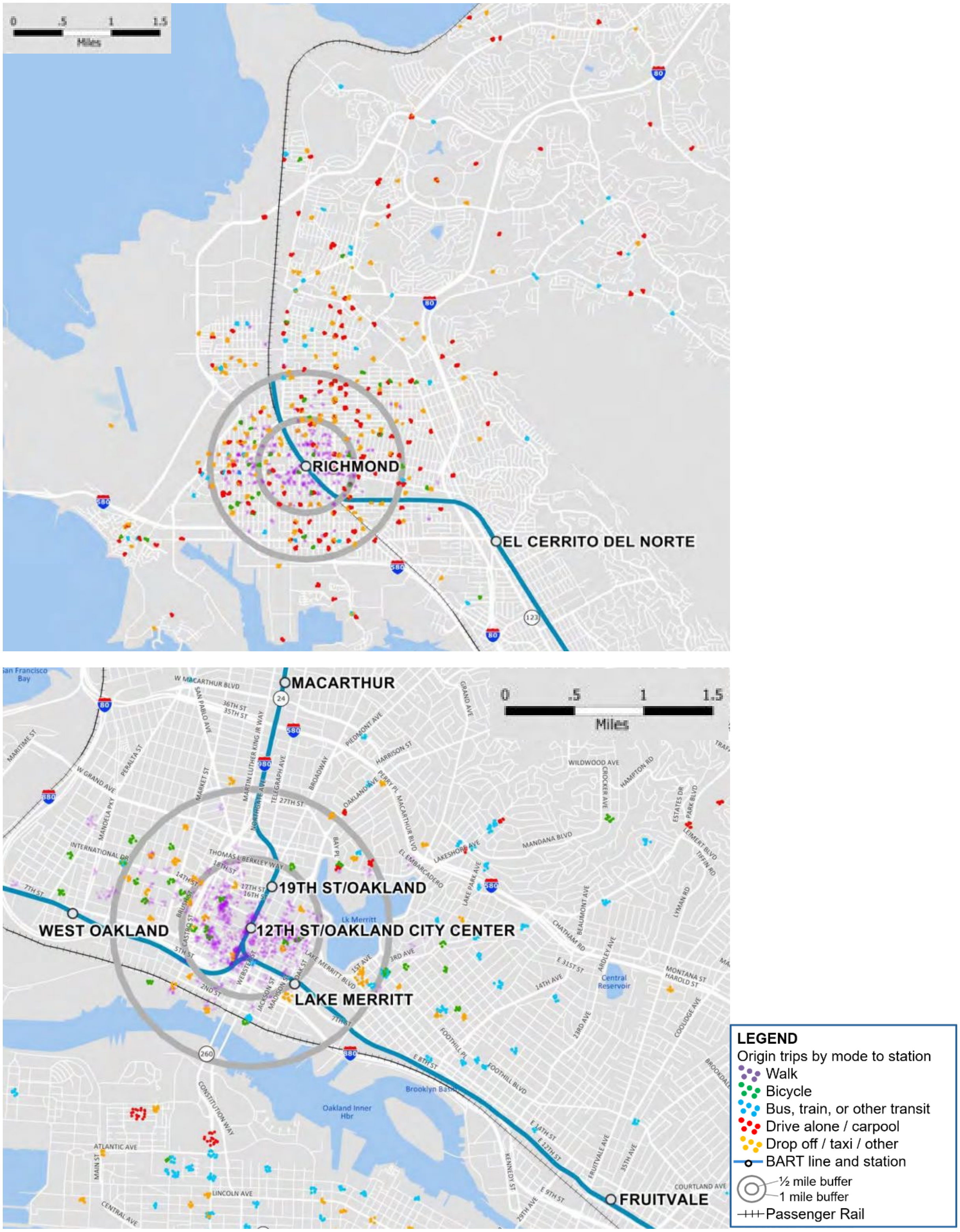
The core sources of data for this analysis were the 2015 San Francisco Bay Area Rapid Transit (BART) Station Profile Survey and the 2016 Capitol Corridor OD Survey. These surveys provided information on where people using each station originated from, as well as the mode they used to access the station.

Figure 3 provides examples of the data extracted from the BART 2015 Station Profile Survey showing that station catchments vary according to the following criteria:

- Land use density around the station
- Distance of the rail trip
- Spacing of stations

Figure 3. Examples of BART Station Profile Survey Data

A downtown Oakland station has very different catchment characteristics from an end-of-line BART station.

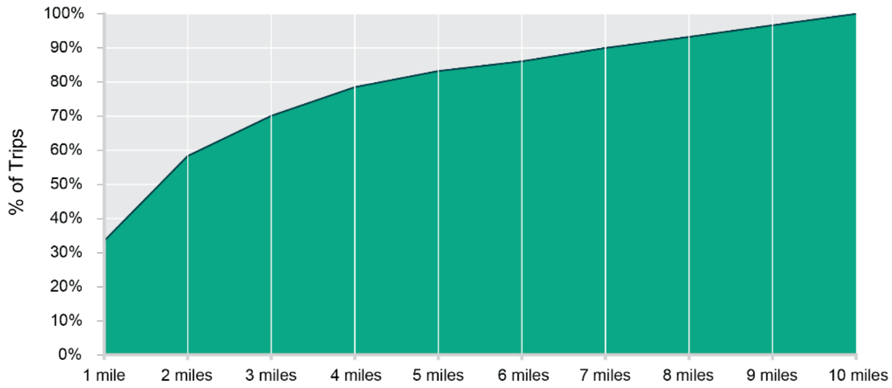


Source: BART 2015 Station Profile Survey

Given the requirement for possible stations to be assessed in isolation (i.e., excluding the effects of other nearby stations), the first two criteria were focused on. **Figures 4, 5,** and **6** provide examples of the analysis undertaken.

Figure 4. Access Distances to BART Stations

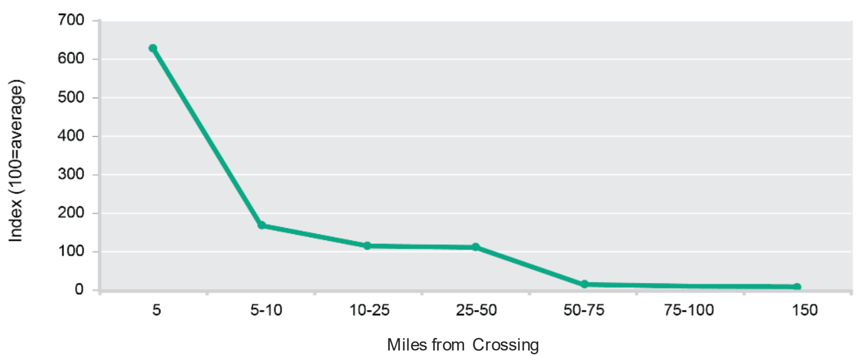
For end-of-line BART stations, approximately 80% of trips are typically made from within 5 miles of a station.



Source: PMC analysis of BART Station Profile Survey data

Figure 5. Trip Rates by Distance from the BART Transbay Tube

Transbay trip rates[4] decline with distance from the Transbay Tube.



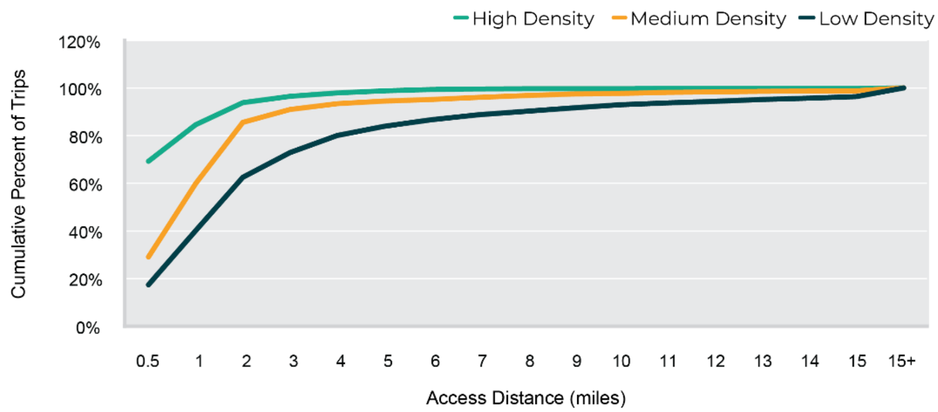
Source: PMC analysis of MTC Travel Model outputs

Note: The middle of the existing Transbay Tube was used as the point definition of the Transbay Crossing.

[4] Trip rates are defined as trips per capita, (i.e., trips divided by population).

Figure 6. Access Distance to BART Stations by Station Density

There are more trips closer to BART stations where there is higher density at the station.



Source: PMC analysis of BART Station Profile data; note that these density cutoffs are defined below.

From the analysis, three density thresholds were identified:

- **High** (>72 people + jobs per acre): Determined based on catchments for BART stations in high-density areas, such as downtown San Francisco and Oakland.
- **Medium** (29-72 people + jobs per acre): Determined based on catchment for stations across the rest of Oakland (excluding downtown), Berkeley, and for areas in San Francisco outside of downtown.
- **Low** (23-29 people + jobs per acre): Lower density areas with the lower bound of 23 based on the Federal Transit Administration's (FTA) definition of high population density.[5]

Three distance thresholds for accessing the Transbay Corridor from households were also determined:

- **Near** (people living <7 miles from the middle of the BART Transbay Tube): There is a sharp drop off in transit trip rates between 5 and 10 miles from the middle of the BART Transbay Tube (Figure 5).
- **Medium** (people living between 7 and 50 miles from the middle of the BART Transbay Tube): 50 miles roughly corresponds to BART's service area.
- **Far** (people living >50 miles from the middle of the BART Transbay Tube): Encompasses the remainder of the Megaregion.

DRAFT - DELIBERATIVE

[5] FTA Guidelines for Land Use and Economic Development Effects for New Starts and Small Starts Projects, 2013. Note that the high population density definition means that a "significant amount of land in station areas is available for new development or redevelopment at transit-supportive densities." Using high population density seemed to be an appropriate minimum expectation for the low density cutoff given the transit-supportive density in the Megaregion.

Based on these density and distance thresholds, nine catchment sizes were created that incorporate the differences observed in existing catchment sizes. The assumed catchments are in **Table 5**.

Table 5. Cluster Radius Guidelines

This matrix of density and distance bands sets the rules for cluster radius based on its hub properties.

	NEAR DISTANCE <7 MILES ACCESS	MEDIUM DISTANCE 7-50 MILES ACCESS	LONG DISTANCE >50 MILES ACCESS
High Density >72 population + jobs per acre	1 mile	2 miles	5 miles
Medium Density 28-72 population + jobs per acre	2 miles	3 miles	7 miles
Low Density 23-29 population + jobs per acre	5 miles	7 miles	10 miles

The smallest catchment radius is 1 mile for high density areas near the bay crossing. The largest catchment radius is 10 miles for low density areas far from the bay crossing.

For future year thresholds, the distances remain unchanged, but the density thresholds are increased by the average growth forecast for the Megaregion (17% between the base and future years of 2015 and 2040). The thresholds for the future year are:

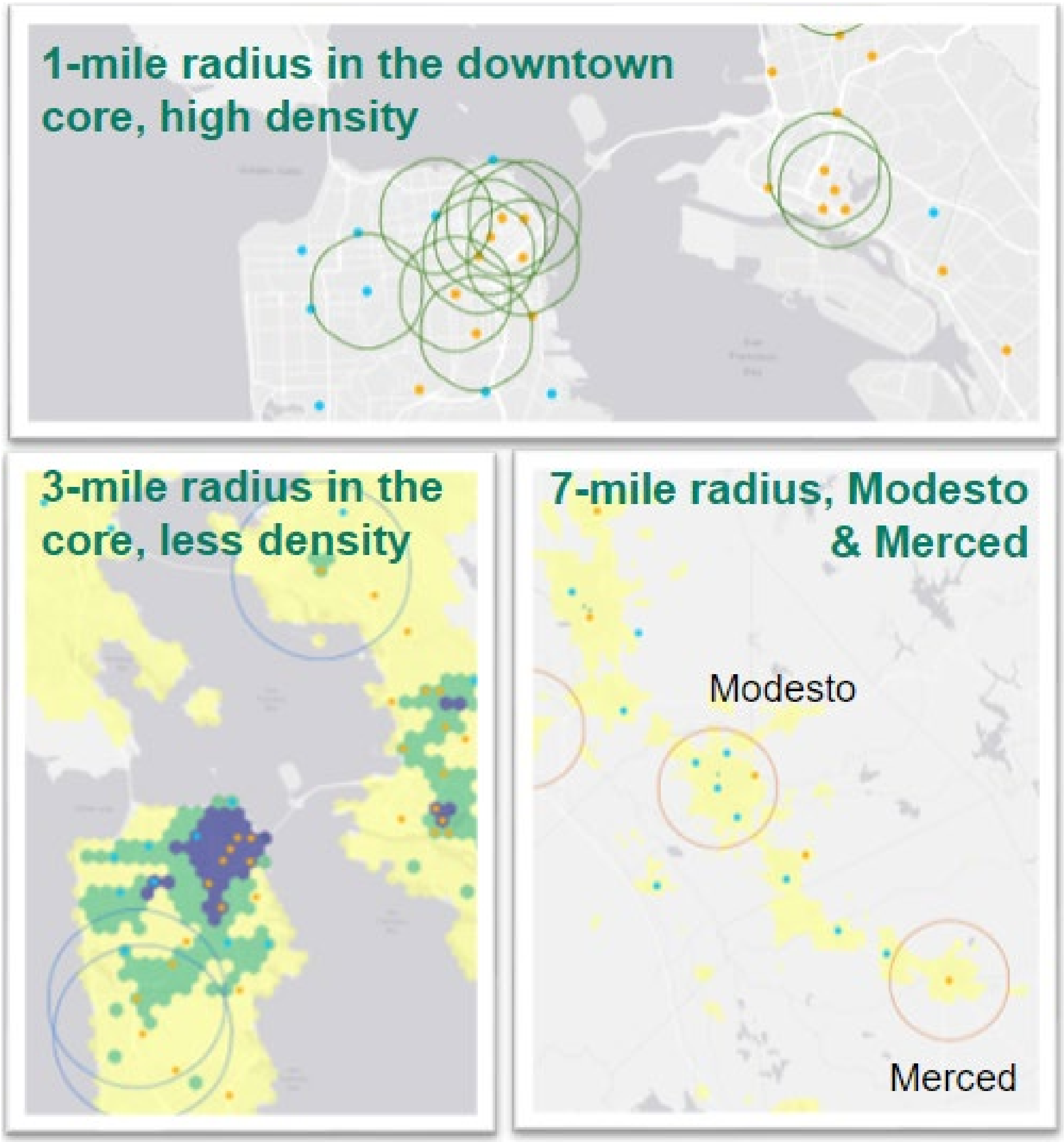
- **High** (>84 people + jobs per acre)
- **Medium** (34-84 people + jobs per acre)
- **Low** (27-34 people + jobs per acre)

Figure 7 shows examples of different cluster sizes throughout the Megaregion.

DRAFT - DELIBERATIVE

Figure 7. Cluster Sizes

Cluster sizes vary throughout the Megaregion.



INITIAL THRESHOLD CRITERION

An individual hexcell has a diameter of 0.5 miles and covers an area of approximately 138 acres. There are 107,677 hexcells in the entire Megaregion. For increased efficiency, a minimum initial threshold was set excluding hexcells that did not meet the minimum population criterion from further analysis. This process was described earlier in the Hexcells section.

The selection of hubs from the array of hexcells is described in the Hub Selection section.

DRAFT - DELIBERATIVE

FUTURE YEAR ASSUMPTIONS

The underlying population and employment data on which the clustering methodology is based comes from each of the four MPO models that cover the Megaregion. Assumptions regarding future year growth[6] also come from these models, specifically the following scenarios:

- MTC: Plan Bay Area 2040
- Sacramento Area Council of Governments (SACOG): 2040 Metropolitan Transportation Plan/Sustainable Communities Strategy (MTP/SCS)
- Three Counties: 2042 Regional Transportation Plan (RTP)
- Association of Monterey Bay Area Governments (AMBAG): 2040 MTP/SCS

When determining the hub locations, density thresholds were used to identify appropriate locations. These thresholds were initially defined in relation to the base year. Future year thresholds were developed by applying the average overall growth in population and employment across the Megaregion to the base year thresholds. A hexcell was determined to exceed the overall threshold if it exceeded the threshold in either the base or future years.

EQUITY CONSIDERATIONS

Given Link21's commitment to equity — specifically partnering with priority populations and residents to maximize benefits and minimize burdens for communities that, historically and currently, suffer and experience negative impacts from transportation infrastructure projects — it is important to consider priority populations explicitly throughout the clustering framework, and to identify impacts on priority populations separately from the general population as part of subsequent analyses.

To meet these dual needs, hexcells were identified as either being priority or general population[7]. This designation was based upon a combination of MTC's identification of CoCs, CalEPA's identification of state priority populations, and CMA adjustments.

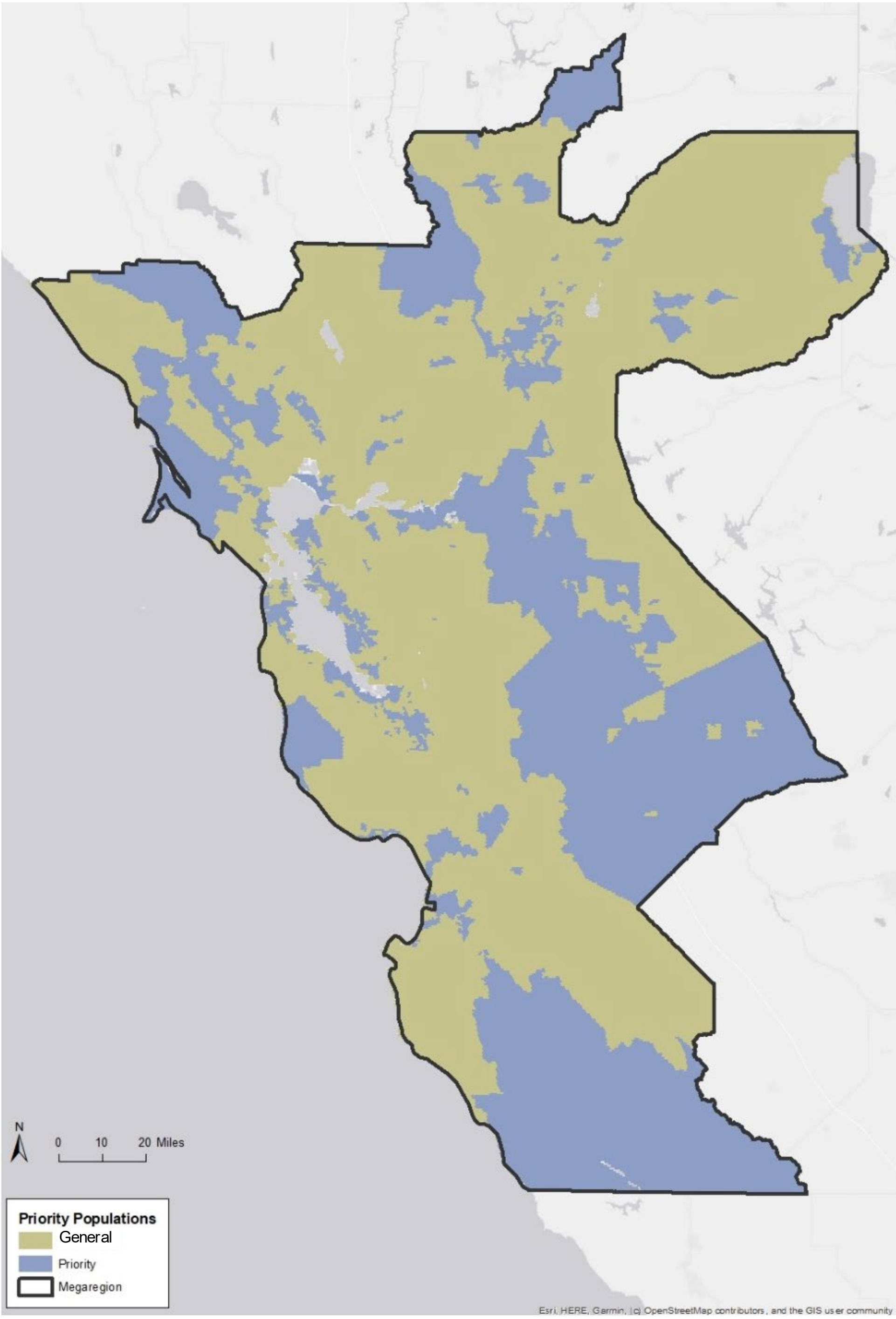
Identified priority populations hexcells were both tracked (for the purposes of monitoring/reporting on the impacts of alternatives on priority populations) and subject to different density threshold criteria when determining the hub locations.[8]

Figure 8 shows the location of priority populations hexcells within the Megaregion.

---

[6] Alternative future growth scenarios are considered as part of a wider analysis, but they are not used in the identification of clusters.
[7] Link21 has since updated its definition of priority populations.
[8] This is discussed in the Hub Selection section under "Identify hexcells that exceed the threshold criteria." Thresholds were halved for priority population hexcells.

Figure 8. Priority Populations in the Megaregion



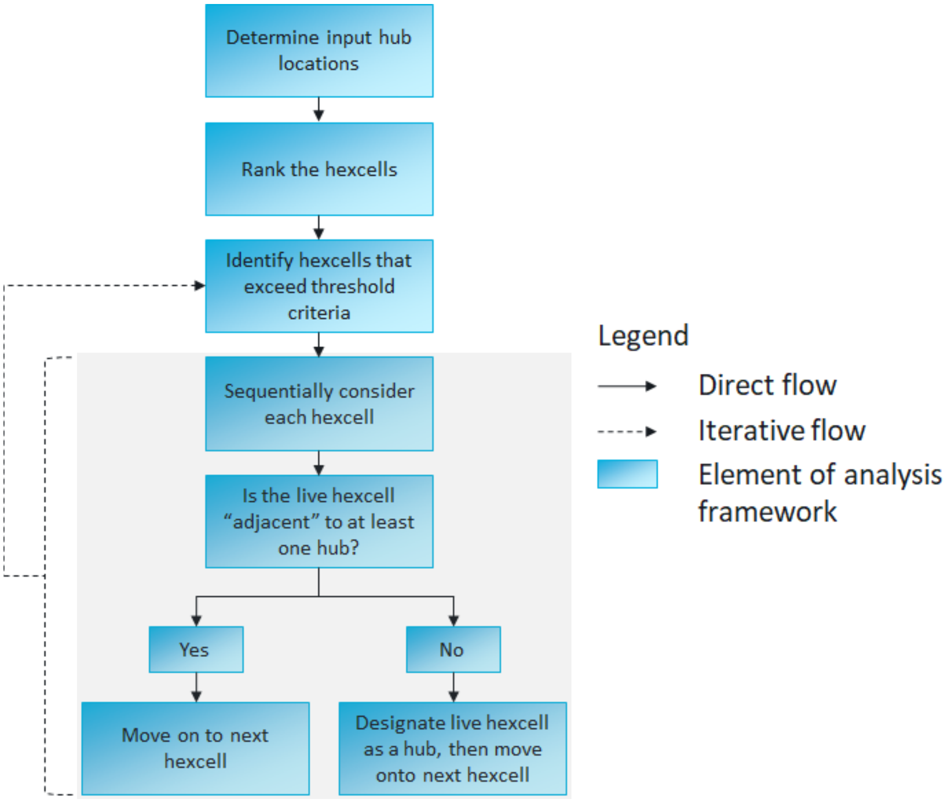
Source: PMC analysis of priority populations hexcells

DRAFT - DELIBERATIVE

Hub Selection

Figure 9 is an overview of the process used to select hubs.

Figure 9. Overview of Hub Selection Methodology



DRAFT - DELIBERATIVE

Each step of the process for identifying hub locations is described as follows:

- **Determine input hub locations:** Some hexcells were determined up front to be hubs. These are hexcells that align with the locations of existing or planned stations for any of the following services: BART, San Joaquins, Caltrain, Capitol Corridor, Sonoma-Marin Area Rail Transit (SMART), Altamont Corridor Express (ACE), and underground light rail San Francisco Municipal Railway (Muni) stations.
- **Rank the hexcells:** Each remaining hexcell was ranked according to the sum of its population and employment in both the base and future year.
  - Example: If Hexcell A has a base year population/employment of 50/100 and future year population/employment of 75/125, then its cumulative size would be 350. Similarly, if Hexcell B has a base year population/employment of 75/75 and future year population/employment of 125/125, then its cumulative size would be 400. Therefore, Hexcell B would be ranked higher than Hexcell A.
- **Identify hexcells that exceed the threshold criteria**: Each hexcell was then compared against the high-density threshold criteria (Table 3).

  - For general population hexcells, the population and employment in each year was compared to the density thresholds. If at least one individual value of the hexcell (i.e., the population or employment in the base or future year) exceeded the thresholds, then the hexcell was determined to exceed the threshold criteria.
    - Example: If Hexcell A has a total base year population and employment of 60 per acre and future year population and employment of 80 per acre, then the hexcell would not pass the thresholds (of 72 and 84) in either year.
  - For priority populations hexcells, the same check was used except that the threshold values were halved (or equivalently, the population/employment values are doubled). Thus, priority populations hexcells were more likely to pass the threshold criteria than general population hexcells.
    - Example: If Hexcell A has a total base year population and employment of 60 per acre and future year population and employment of 80 per acre, then the hexcell exceeds the priority populations thresholds (of 36 and 42 – half the threshold for general population hexcells).
- **Sequentially consider each hexcell**: Each hexcell that exceeded the threshold criteria was considered in accordance with their ranking. The hexcell with the highest ranking was considered first, followed by the hexcell with the next highest ranking, and so on:
  - The distance of the considered hexcell to each currently identified hub was calculated.
  - Adjacency criteria were set at half the cluster radii shown in Table 5.
    - Example: If Hexcell A is high density and 60 miles from the bay crossing (Far distance), then the adjacency criteria would be 2.5 miles (half of 5 miles).
  - If the considered hexcell was within the adjacency criteria of one or more currently identified hubs, then it did not meet the requirements to be an additional hub. If the considered hexcell was outside the adjacency criteria of all currently identified hubs, then the considered hexcell was designated a new hub.
  - Once this determination was made for the considered hexcell, the process was repeated for the next hexcell that exceeded the threshold criteria according to the hexcell ranking.
- **Repeat the process for the subsequent threshold criteria**: This process was then repeated for the medium-density threshold criteria and the low-density threshold criteria (shown in Table 5). Hexcells with fewer than 23 population + jobs per acre were not considered.

At the end of the process, all hexcells that met any of the threshold criteria were either identified as hubs or were within the adjacency criteria of at least one hub.

Hubs

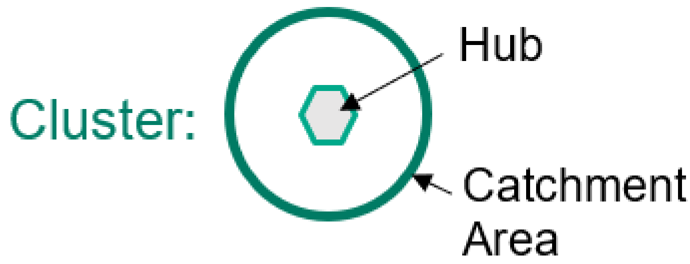
Each identified hub was classified as one of the following:

- Existing rail station
- Planned rail station
- New hub with high population or job density either today or forecast in the future as identified by the procedure in Figure 9.

Figure 10 shows the spatial relationship between a hub and a cluster. The hub is the hexcell at the center of a cluster with the cluster comprising the catchment area surrounding the hub.

Figure 10. Spatial Relationship between a Hub and a Cluster

Each cluster is comprised of a hub and a catchment area.



A total of 202 hubs were identified in the Megaregion: 122 existing stations, 30 planned stations, and 50 new hubs as a result of high population and employment density. Hubs were more closely located in areas that align with existing stations; are closer to the bay; or have higher population/employment densities, particularly for priority populations (e.g., around San Francisco and, to a lesser extent, Sacramento); and less closely located in areas further from the bay or with lower population/employment densities (e.g., in the San Joaquin Valley). This pattern of hubs was considered reasonable and is illustrated in **Figure 11** and **Figure 12**.

As shown in **Table 6**, the coverage of the clusters ensures that the vast majority of population and jobs are within a cluster, while all major population/employment centers are either directly covered by or are immediately adjacent to a hub.

Table 6. Coverage of the Clusters

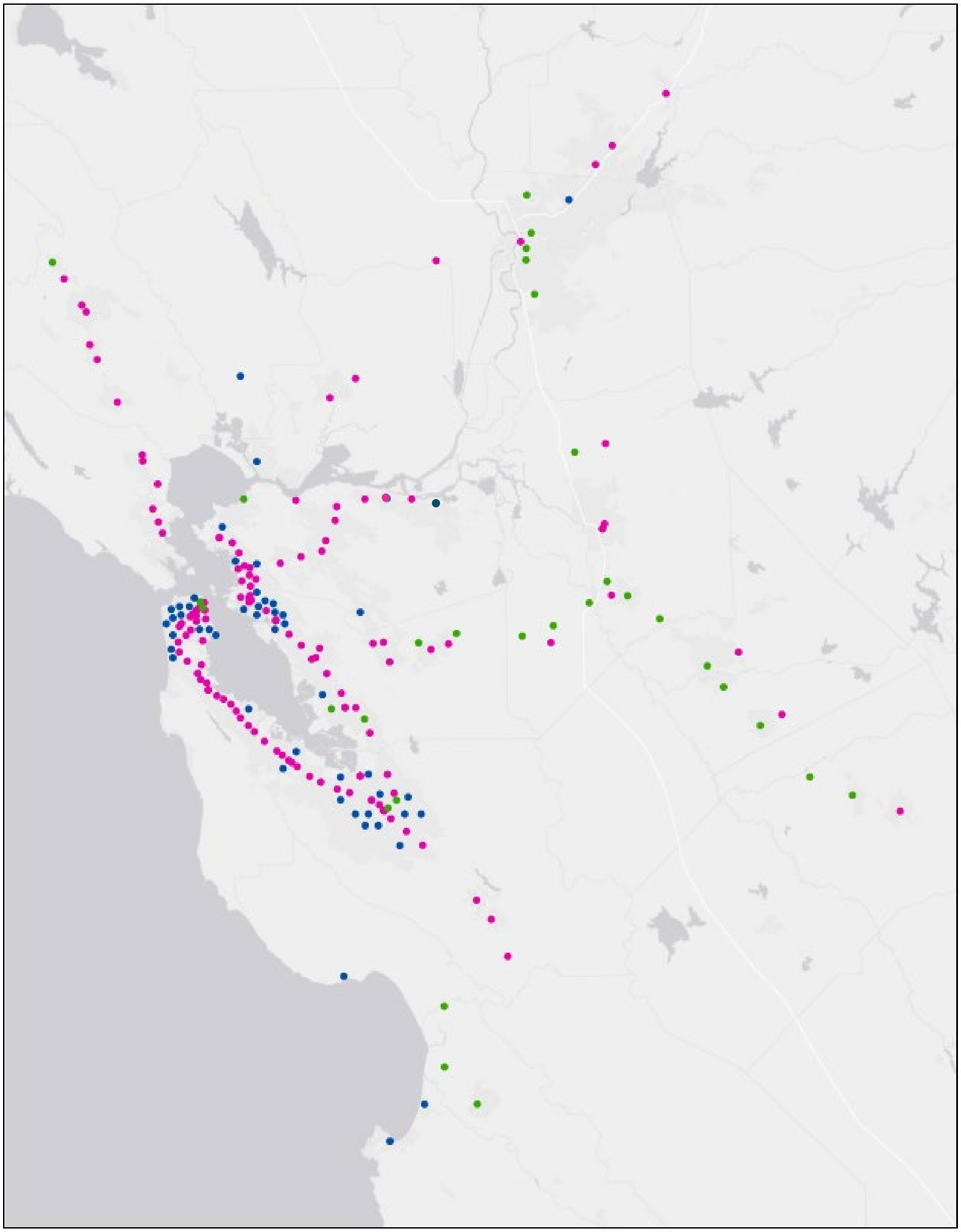
Over 90% of current and future population and employment are within a 5-mile catchment area from a hub.

	PERCENT OF POPULATION		PERCENT OF EMPLOYMENT		PERCENT OF HEXCELLS
	Base	**Future**	**Base**	**Future**	
All Hexcells					
Directly within a hub	3.2%	9.3%	3.3%	9.0%	0.6%
Within 0.5 miles of a hub	17.9%	29.9%	17.8%	29.2%	4.1%
Within 5 miles of a hub	90.9%	94.2%	90.2%	94.0%	63.8%
Further than 5 miles from a hub	9.1%	5.8%	9.8%	6.0%	36.2%
Priority Populations Hexcells Only					
Directly within a hub	6.0%	11.6%	6.0%	11.8%	1.4%
Within 0.5 miles of a hub	26.6%	34.6%	26.0%	34.5%	7.6%
Within 5 miles of a hub	95.5%	96.8%	94.6%	96.8%	71.5%
Further than 5 miles from a hub	4.5%	3.2%	5.4%	3.2%	28.5%

Note that hubs represent areas of market opportunity for Link21, but they are not necessarily the locations of future Link21 stations. They identify broader geographic areas that may benefit from enhanced service; the market analysis does not seek to determine the precise details of where stations may be located.

Figure 11. Hubs Across the Megaregion

There are 202 hubs across the Megaregion.

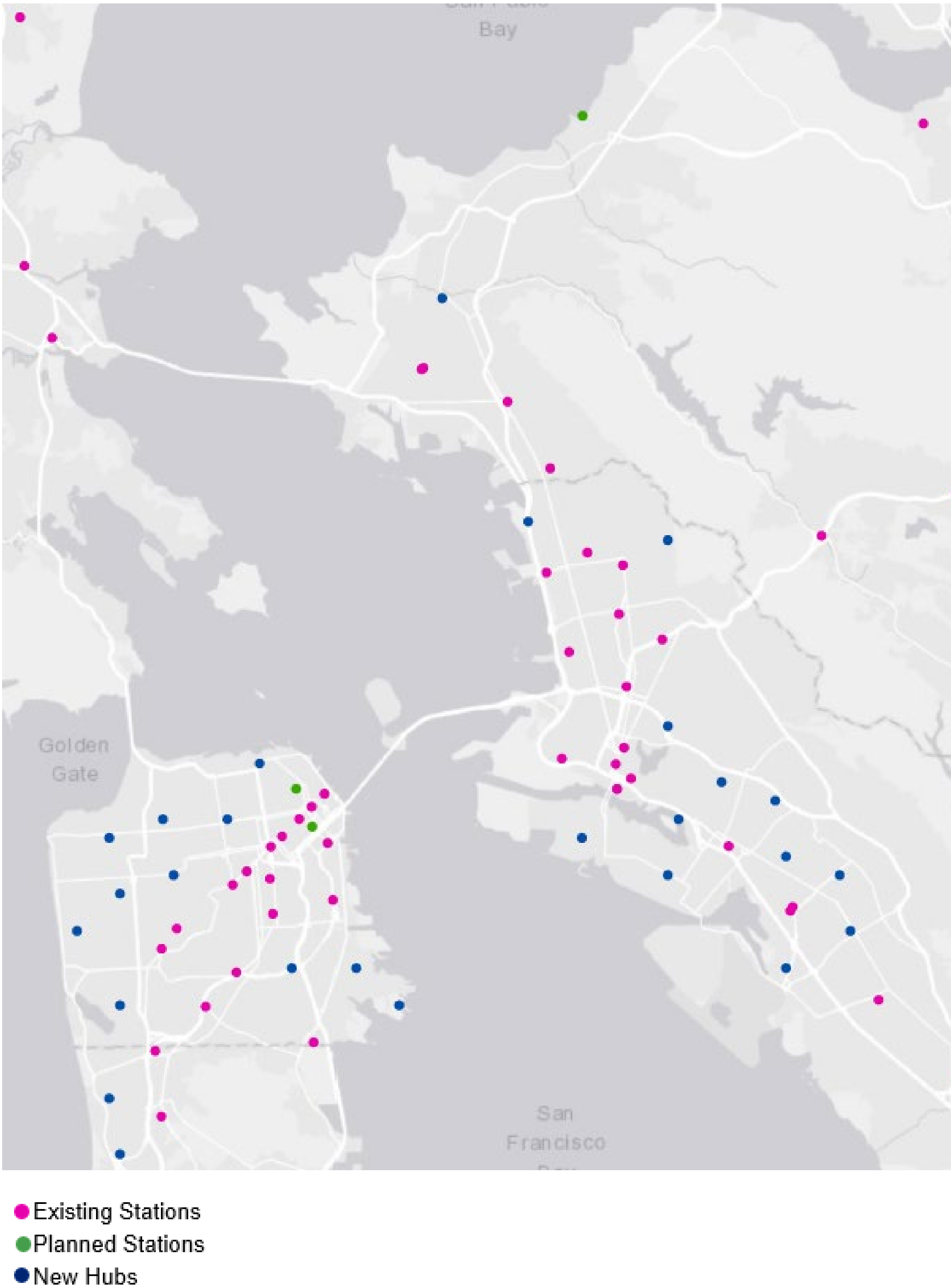


Existing Stations
Planned Stations
New Hubs

Most hubs around the San Francisco Bay are located much closer together.

Figure 12. Core San Francisco and East Bay Hubs

In the core San Francisco and East Bay areas, hubs are numerous and close together.



DRAFT - DELIBERATIVE

Rail Potential Model Development

Rail potential models identify factors driving high rail ridership in the Megaregion and help identify the conditions that enable high rail ridership.

Overview

The rail potential model is integral for addressing the main objective of the market analysis — providing a list of corridors in the Megaregion with high rail potential that can be served by Link21. A rail potential model was estimated to identify the conditions that could enable high rail ridership in the Megaregion. This model estimated rail potential between cluster pairs (i.e., a cluster pair represented OD travel from one cluster to another cluster).

The core purpose of this model was to identify factors that enable high rail ridership, not to forecast ridership for these corridors. This was not meant to be a ridership forecasting exercise and any models or estimates discussed should not be interpreted as ridership estimates.

Model

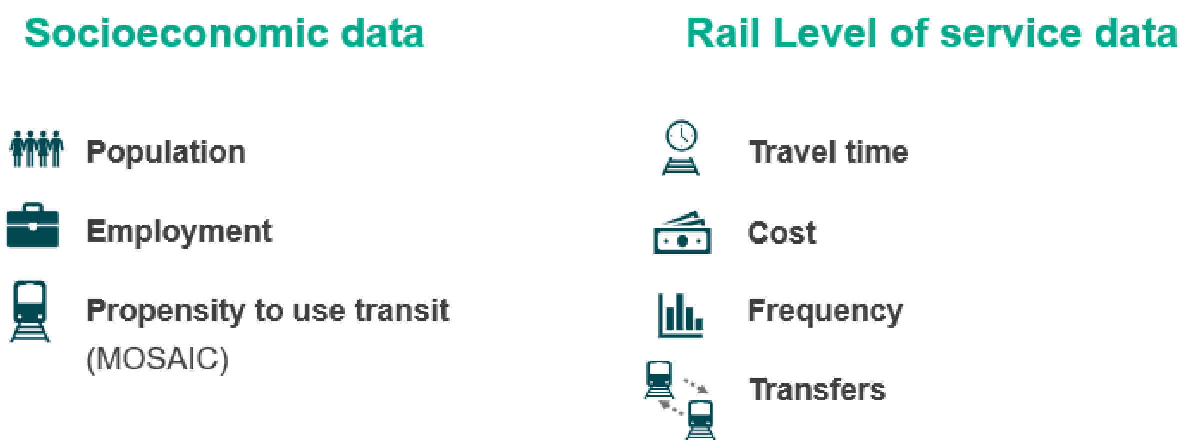
The model was estimated by fitting a regression to observed unconstrained[9] rail ridership data for base year 2015 by cluster pairs, which were later grouped into rail corridors. The regression helped identify the parameters that drive rail potential by estimating rail potential as a function of key factors, including hub and cluster characteristics and travel characteristics.

The regression models applied in the MAST contained socioeconomic characteristics of the cluster, such as population and employment density and parking costs, as well as rail level of service characteristics, such as travel time, cost, frequency, and transfers.

[9] See section entitled “Effects of crowding on rail demand” for details on how this was developed.

Figure 13. Model Variables

The variables with the highest impact on existing rail demand included socioeconomic data and trip characteristics (rail level of service data).



In addition to the trip characteristics, the MAST adjusted the estimated rail potential according to the estimated level of crowding experienced between cluster pairs. Therefore, crowding impacts were estimated within the MAST (discussed in the next section).

While testing different model specifications, an outlier analysis was conducted to better understand cluster pairs that did not strongly align with the regression results for most cluster-cluster pairs. This resulted in the usage of markers to account for the special characteristics of these markets (e.g., very long distance trips and very short distance trips have lower demand). These included:

- **Transbay marker:** This definition is based on which county the origin and destination are in. One end must be in San Francisco or San Mateo counties and one end must be in the East as defined below.
  - East: Alameda, Contra Costa, El Dorado, Napa, Placer, Sacramento, San Joaquin, Solano, Sutter, Yolo, and Yuba counties
  - Other counties — Stanislaus, Merced, Santa Clara, San Benito, Santa Cruz, Monterey, Marin, and Sonoma — were not included in the East because most trips between these counties and San Francisco or San Mateo counties do not use the Transbay Corridor
- **Off-peak marker:** This recognizes that peak and off-peak travel patterns are different.
- **Multiple distance markers:** This includes both long distance (>=30 and >=90 miles) and short distance (<=2 and <=3 miles) markers. Note that the trip distance bins are additive (e.g., a trip that is <=2 miles has both the values from <=2 miles and <=3 miles.)
- **Parking cost markers:** When parking costs increase the cost of an auto trip, there is higher rail potential than in similar clusters without parking costs. The presence of

parking costs is not assumed to change in the future year baseline. The models performed worse when this coefficient was not used.

- **BART end-of-line marker:** This represents BART stations that have larger catchment areas than average. Many of these stations are towards the end of a line, and some are terminus stations. The full list of stations that receive this marker are Walnut Creek, Richmond, El Cerrito del Norte, Pleasant Hill/Contra Costa Centre, West Dublin/Pleasanton, Dublin/Pleasanton, and Antioch. This is consistent with BART's understanding of behavior at those stations.

The rail potential model is presented in **Table 7**. All coefficients have the expected sign, and values are considered reasonable (i.e., broadly appropriate order of magnitude).

Table 7. Rail Potential Model Coefficients

VARIABLE	COEFF.	NOTES
Population/employment	1.00	If population or employment increases by 10%, forecast trips should increase by broadly 10%.[10]
Rail journey time (minutes)	-0.83	If rail journey time decreases by 10%, forecast trips should increase by 8.3% — this is in line with most benchmarks.
Rail cost (dollars)	-0.35	If rail cost decreases by 10%, forecast trips should increase by 3.5% — this is in line with (or slightly more sensitive than) most benchmarks.
Rail frequency (trains per hour)	0.23	If rail frequency is increased by 10%, forecast trips should increase by broadly 2.3%.
Rail transfers (binary indicator of rail-rail transfers)	-0.52	Penalty varies depending on trip length. Equivalent to ~25 minutes for a 35-mile trip, ~50 minutes for an 85-mile trip.
Transbay marker[11]	1.25	Transbay trips have higher demand.
Off-peak marker	-0.40	Off-peak trips have lower demand.
Long distance marker (>=30 miles)	-0.96	Long distance trips have lower demand.
Longer distance marker (>=90 miles)	-2.40	Very long distance trips have lower demand.
Short distance marker (<=3 miles)	-1.37	Short distance trips have lower demand.

[10] This selected variable was asserted by taking it outside of the regression itself.
[11] While this coefficient has a very high positive value, models tested with the transbay variable performed better than models tested without the transbay variable.

VARIABLE	COEFF.	NOTES
Shorter distance marker (<=2miles)	-1.28	Very short distance trips have lower demand.
Parking cost marker (parking cost exists)	0.62	Trips with auto parking costs (e.g., downtown San Francisco) have higher demand
BART end of line marker	0.83	Trips that have an extended Park and Ride/catchment area have higher demand
Widely spaced stations (>=5 miles from other stations)	1.12	Widely spaced stations have higher demand
Constant	-15.36	*model constant*
R-squared	0.51	*measure of model fit*
Total Error Percentage	-8.6%	*measure of model fit*

Benchmarking

The overall model had an R-squared[12] of 0.51. While this R-squared would generally be considered low, the rail potential models were used for ranking and not demand forecasting. The overall model error was relatively low at <9% (**Figure 14**), and it had reasonable performance across a range of geographies and levels of service. Additionally, this model covered a wide geography and various rail services and resulted in an R-squared higher than the MTC model.[13]

The main focuses of the rail potential model benchmarking were the:

- Ability of the model to reasonably match observed demand
- Reasonableness of the behavioral responses within the model
- Reasonableness of the model when used for validation

When comparing against actual observed rail trips, the broad trend of the projected rail trips was reasonable as seen in **Figure 15** and also when disaggregated by distance as seen in **Figure 16**. Note that noise is expected with outliers considering the wide range of levels of service (LOS) and trips.

[12] R-squared is a statistical measure of model fit, showing how close the modeled data are to the fitted regression line. A higher R-squared shows better model fit (1.0 = perfect model fit).

[13] The MTC model is the model mentioned in the BART Ridership Model (BRM) Development Report prepared by Fehr & Peers, November 2019. The MTC's current regional travel demand model is an activity-based model known as Travel Model One. Note that there are various R-squareds (0.29 to 0.46) corresponding to the model's performance for 2015-2019 conditions (each year) in the AM peak, PM peak, and off-peak periods.

Figure 14. Model Estimation Errors

The overall forecast demand is 391,000 trips compared with the estimated actual demand of 427,000, which is an underestimation of 8.6%. The underestimation is higher in the off-peak than the peak period.

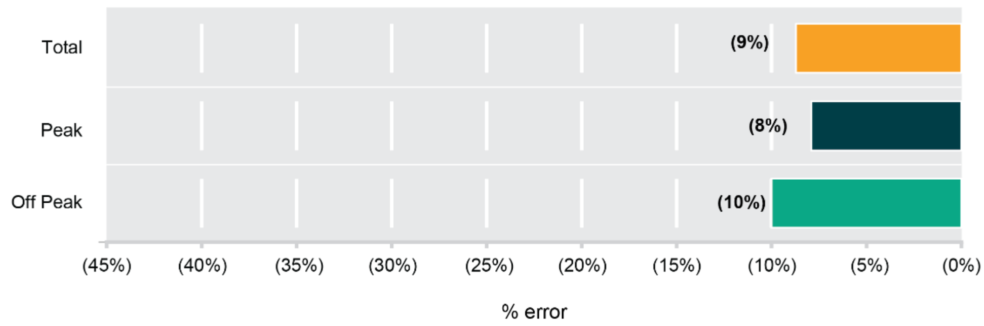
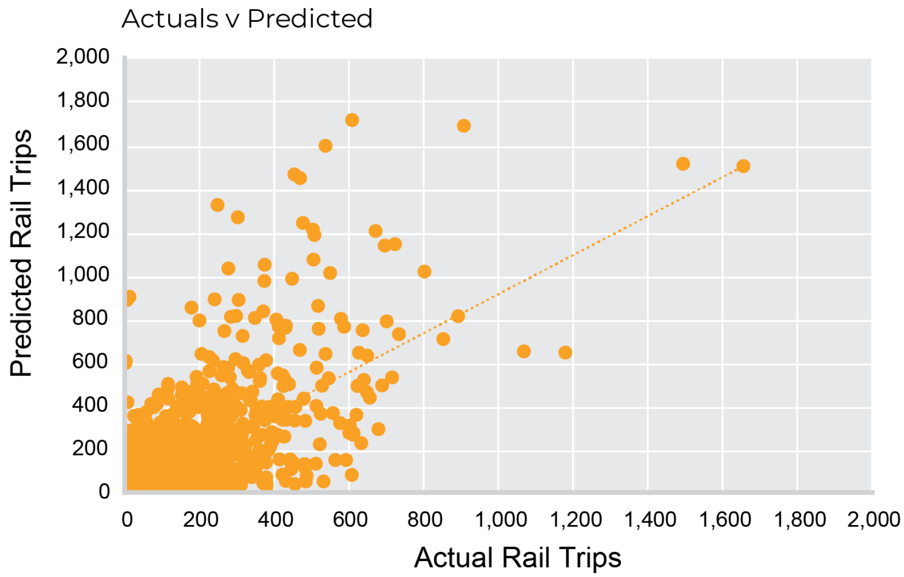


Figure 15. Actual vs. Predicted Rail Trips

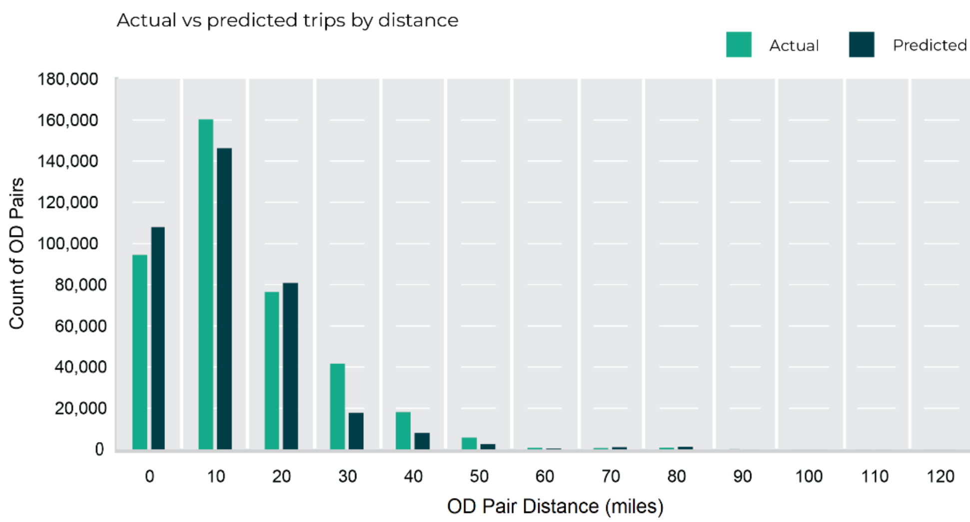
The broad trend of projected rail trips vs. actual observed rail trips is reasonable.



DRAFT - DELIBERATIVE

Figure 16. Actual vs. Predicted Trips by Distance

When grouping cluster pairs by distance, the broad trend of the projected rail trips is reasonable.



Comparisons between observed and estimated cluster pair trips were also conducted to ensure reasonableness of the model. Some examples are:

- **Embarcadero – Fruitvale** base year actual trips (from OD trip tables[14]): 2,489 peak hour (including both AM and PM) trips, estimated trips: 2,347
- **Powell Street – Lafayette** base year actual trips (from OD trip tables): 73 peak hour (including both AM and PM) trips, estimated trips: 96

Key variable parameters were benchmarked against existing literature and other studies to ensure the model responded reasonably to different changes. These included:

- **PDFH** (UK Passenger Demand Forecasting Handbook version 5.0): this is a well-respected guide for rail demand forecasting used in the UK, dated April 2013
- **VTPI** (Victoria Transport Policy Institute in British Columbia, Canada): this institute published *Transit Price Elasticities and Cross-Elasticities*, dated July 21, 2011
- **CS** (Cambridge Systematics) Responses: this is a memorandum regarding information requested in "Section 3.2 Validation and Documentation" of the *Independent Peer Review of the California High-Speed Rail Ridership and Revenue Forecasting Process, 2005-10, Draft Report for Internal Review* (February 7, 2011), dated June 8, 2011
- **Tinbergen** (Tinbergen Institute Discussion Paper): this paper discusses *Intermodal Competition in the London-Paris Passenger Market: High-Speed Rail and Air Transport*, dated 2009

[14] Trip table development is discussed in detail in Appendix D.

- **U.S. intercity rail operator** [confidential model documentation]

Table 8 shows the comparison of the rail potential model parameters against these industry benchmarks.

Table 8. Comparison of Rail Potential Model Parameters

The rail potential model compares well against industry benchmarks.

VARIABLE	PARAMETER VALUE	PDFH BENCHMARKING	OTHER BENCHMARKING
Population/ employment	1.00	1.0 (population), 1.3 (employment)	***
Rail journey time	-0.83	-0.9 to -1.35 (generalized journey time [GJT])	-0.63 to -1.20 [U.S. rail operator], -0.97 to -1.05 [CS]
Rail cost	-0.35	-0.40 to -1.2 (lower elasticity for shorter trip distances)	-0.15 to -0.6 (peak) -0.3 to -1.0 (off-peak) [VTPI] -0.51 to -1.20 [U.S. intercity rail operator]
Rail frequency	0.23	Included in GJT	0.43-0.68 [London-Paris high-speed rail]; 0.25-0.49 [U.S. intercity rail operator]
Rail transfers	-0.52	Included in GJT	

****This is consistent with guidance from the PDFH on forecasting growth in rail demand resulting from external factors. Elasticities are provided for various movement types (major city-major city, city-rural, etc.). The elasticity to population is typically close to 1 (unless it is deemed not to play a role due to other factors). Employment elasticities are typically 1 or slightly above 1. This does not account for changing demographics within the population.*

Rail Potential Model Application

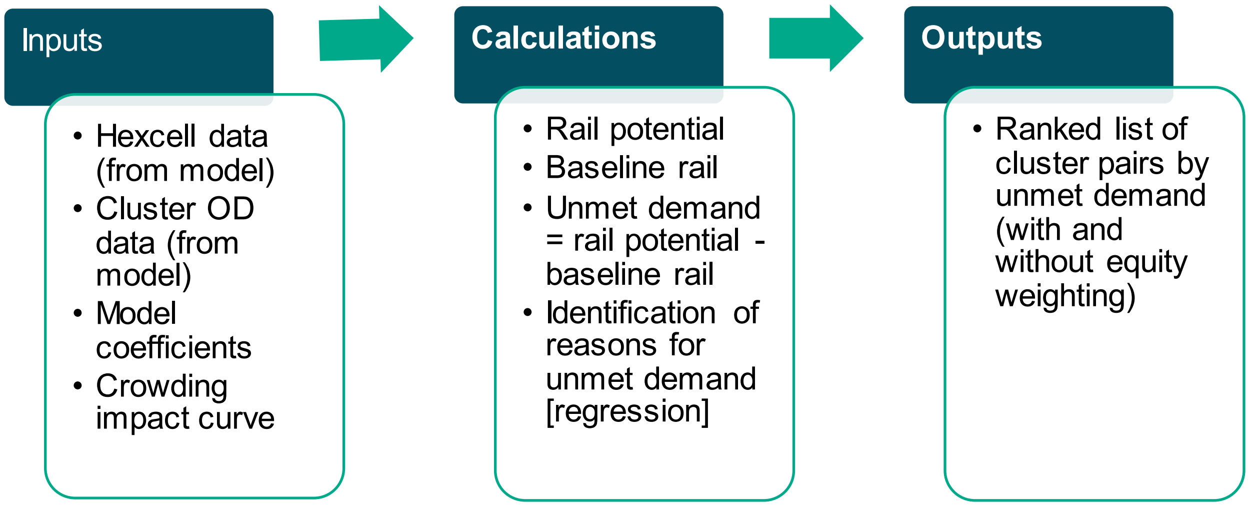
The MAST applies the estimated regressions to all cluster pairs, identifies the drivers of unmet demand, and produces a ranked list of OD pairs by unmet demand, which supports the development of Link21 program concepts. For each cluster pair, this process involves starting with observed OD volumes from the rail operators (with transit propensity applied from MOSAIC); then the regression factors as listed in Table 7 are applied to determine the total cluster pair rail potential.

The final outputs of the MAST are a ranked list of cluster pairs by rail potential that are used for corridor development. The ranking included equity weights to account for priority populations, as defined by Link21 (discussed earlier in this appendix under **Equity considerations**). This ranking process is described in more detail in Appendix C.

Figure 17 describes the inputs, calculations, and outputs for the MAST.

Figure 17. Inputs, Calculations, and Outputs for the MAST

The MAST uses the rail potential model to identify high levels of unmet rail potential in the Megaregion.



To determine unmet rail potential, the MAST applied the rail potential models to account for:

- Effect of crowding on rail potential
- Future year growth
- Good service rail potential

Each of these is described in more detail in the following sections. The grouping of OD pairs into rail corridors is discussed in Appendix C.

Effect of Crowding on Rail Potential

A crowding curve was estimated using a combination of observed data on cluster pairs with capacity constraints and benchmarking from wider evidence and other sources for capture rates[15] (including PDFH and Metrolinx). As such, the model assumed crowding reduced the number of people who use a new service and was used to incorporate crowding impacts in the analysis. The crowding impact was estimated using similar methodology to MTC's *Plan Bay Area 2050*. The curve in **Figure 18** comes from Metrolinx evidence for when the load factor is greater than 70%; no impact was assumed for crowding levels under 70% to align better with both the PDFH and observed evidence.

[15] Capture rates refer to the percentage of passengers who choose to board the train given the level of crowding.

Figure 18. Crowding Curve

When train load factor[16] increases, a factor is applied to in-vehicle time (IVT) that reduces the attractiveness of rail.

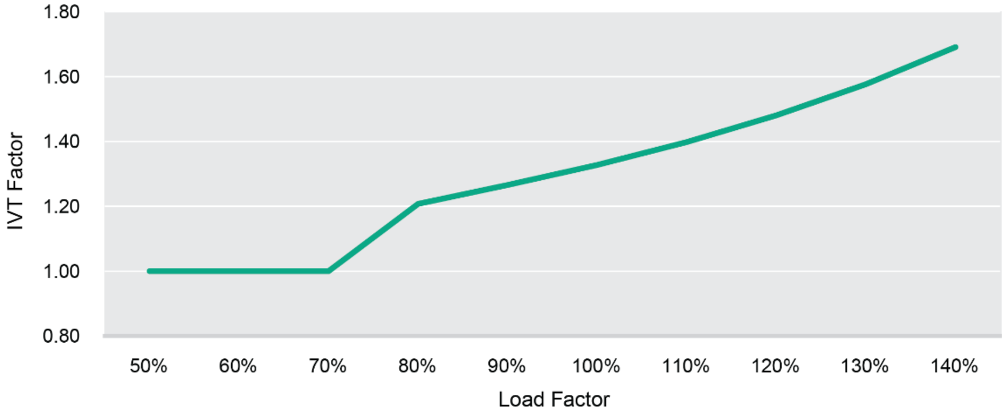


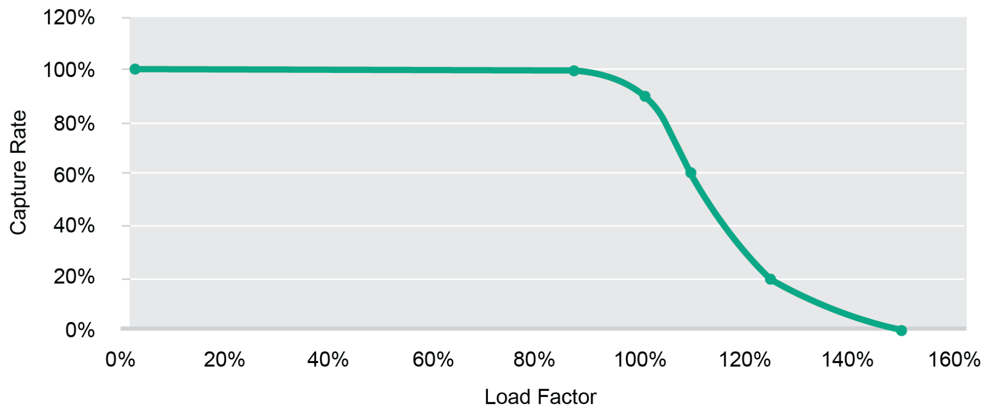
Figure 19 shows the impact of these IVT factors when applied within the model framework. As load factor increases, the IVT factor applied increases; therefore, the capture rate for incremental demand decreases.

This process is first applied to observed ridership in reverse (i.e., to constrained observed demand) in order to estimate unconstrained demand. It is this estimated unconstrained demand that is used within the regression analysis in order to best estimate the effects of variables, such as journey time or frequency, on unconstrained demand. Once unconstrained demand is estimated for all cluster pairs, this same process is applied again in order to convert back to constrained demand.

[16] Load factor is calculated as the number of riders divided by the number of seats.

Figure 19. Train Capture Rate

When train load factor increases, we need to acknowledge that people will choose to not ride.



The model's response to crowding is considered reasonable, as shown in **Table 9**. Crowding affects the transbay market more than the non-transbay market. Without the future baseline improvements, significant rail potential is projected to be unmet due to crowding. Unmet potential in the future is reduced under baseline conditions because the baseline contains projects that increase transbay capacity.

The future baseline includes several transit projects included in the regional transportation plans adopted by MPOs. The following major projects are included in the future baseline:

- South Bay Connect
- Caltrain Enhanced Growth Scenario
- Valley Link
- Valley Rail
- California High-Speed Rail
- CalMod Electrification
- San Francisco Central Subway Phase 3
- SMART Expansion to Larkspur and Windsor
- Caltrain Downtown Extension (DTX)
- BART to Silicon Valley Phase II
- BART Core Capacity Project
- BART Irvington Station
- ACE 8 Trains per Day
- Rail Extension to Monterey County

DRAFT - DELIBERATIVE

Table 9. Model Response to Crowding

	ALL CLUSTER-CLUSTER PAIRS	TRANSBAY CLUSTER-CLUSTER PAIRS	NON-TRANSBAY CLUSTER-CLUSTER PAIRS
Existing year	-7.4%	-12.6%	-2.7%
Future year (existing level of service)	-16.2%	-24.1%	-9.1%
Future year baseline	-7.1%	-11.8%	-1.5%

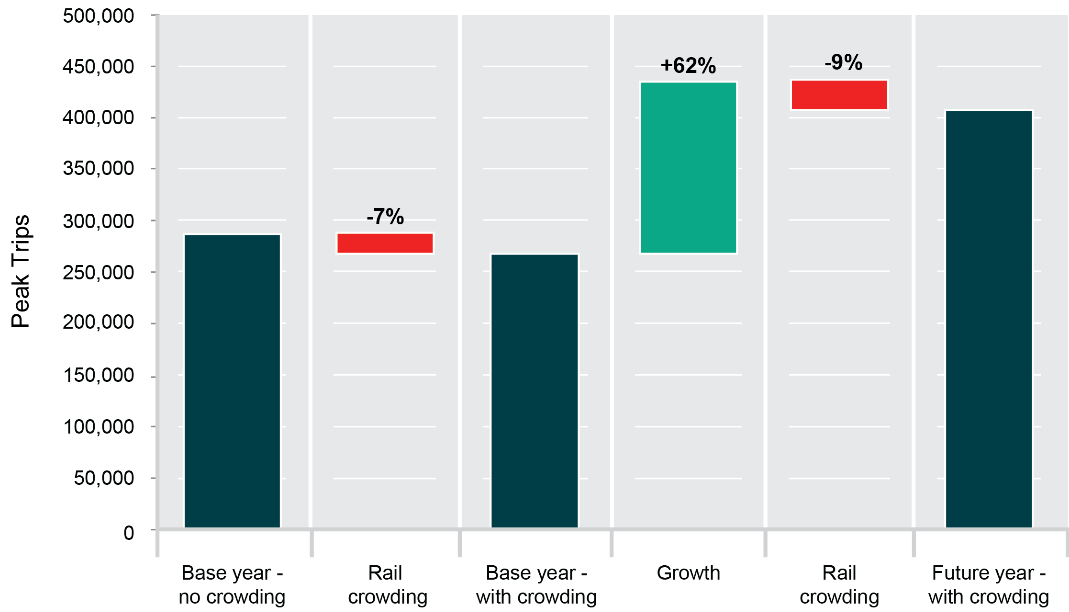
Future Growth

Growth includes population and employment growth to which the model's response is considered to be reasonable. Rail potential grows slightly faster than population and employment growth (33% rail potential growth vs. 28% population and employment growth) because clusters with rail service are expected to grow faster than clusters without rail service.

In addition to population and employment driven impacts, growth also includes new cluster pairs with new service that were added to the future year baseline and existing pairs with enhanced service. On average, the rail journey time went down by 12%, and rail frequency went up by 41%. Additionally, 5% of trips that previously needed a transfer could be direct in the future year.

Figure 20 shows the impact on peak trips across all OD pairs.

Figure 20. Future Year AM and PM Peak Trip Impacts for all OD Pairs



DRAFT - DELIBERATIVE

Crowding reduces peak demand by 7% and 9% in the base and future years, respectively. The total growth impacts (from population growth, employment growth, and baseline service improvements) cause a 62% increase in peak rail potential.[17]

Good Rail Service

When evaluating unmet rail potential in the Megaregion, the analysis should avoid bias toward any particular service, geographic characteristics, or limited existing service levels. Therefore, the concept of “good rail service” was used to identify a theoretical upper bound on rail potential, which in turn was used to identify market opportunities for Link21. Once good market opportunities were identified, the notion of “good rail service” was no longer used to evaluate Link21 alternatives.

WHAT IS GOOD RAIL SERVICE?

Good rail service is not intended to represent actual realistic services that could be run between stations. Rather, it is intended to provide a consistent basis on which to assess the relative ridership potential across cluster pairs. Importantly, this means that good rail service has nothing to do with the service that runs today or with the service that realistically could be run between two stations.

For example, good rail service would typically be seen as fast, frequent, affordable, direct, and with plenty of available seats. In many cases, some, or all, of these might not be plausible; there may, for example, be physical or network barriers that make providing a direct service impractical.

Therefore, good rail service is explicitly a theoretical concept, not a practical one (at least not for all cluster pairs).

HOW LINK21 DEFINES GOOD RAIL SERVICE

Good rail service is not considered to be a uniform concept. In particular, it is something that is different for longer/shorter distance trips:

- For longer distance trips, a competitive travel time and sufficient operating hours/frequency to enable reasonable flexibility regarding travel time choice are likely to be critical factors.
- For shorter distance trips, a high frequency, high capacity, and direct service (i.e., without the need to transfer) are likely to be critical factors.

[17] This 62% increase in peak period rail potential includes baseline service improvements alongside population and employment growth, and it is factored down due to rail crowding (as shown in Figure 20). By comparison, the previously reported 33% growth in rail potential applies for average daily trips, and it is based solely on population and employment growth.

To account for these differences, the definition of good rail service is segmented on a trip length[18] basis to account for trips better served by a good level of regional rail service versus a good level of urban rail/BART service as defined below.

The segmentation uses the following criteria:

- Trips lower than 30 minutes → urban rail
- Trips higher than 90 minutes → regional rail
- Trips between 30-90 minutes → test both urban and regional rail and take the maximum demand estimate of the two

On this basis, the definition of good rail service for each variable is outlined as follows:

- **Rail travel time** is a function of the operating speed of each service and road network distance between cluster pairs.
  - Good rail service should provide a rail travel time that is competitive relative to auto for any given cluster pair. It is the lower of:
    - The rail journey time based on the assumed operating speed (defined for both urban rail and regional rail service below), and
    - 120% of the calculated auto journey time
  - For urban rail, good rail service is defined as an average rail operating speed (including station stops) of 40 mph. Note that this broadly aligns with the highest operating speeds currently in service and is slightly higher than the existing average BART operating speeds of 30-35 mph.[19]
  - For regional rail service, good rail service is defined as an average rail operating speed (including station stops) of 50 mph. This is slightly faster than any existing average operating speeds (currently the maximum average speed is approximately 40 mph on Capitol Corridor) and has been set to better align with average longer distance speeds by auto.
- **Rail cost** is a function of rail costs per mile.
  - Good rail service should be competitively priced relative to auto for any given cluster pair. It is the lower of:
    - For urban rail, \$2 + \$0.10 per mile (minimum cost of \$2.10); for regional rail, \$0.25 per mile
    - 150% of the calculated auto cost
- **Rail frequency** is defined by trains per hour (tph).
  - Good rail service is defined separately for the peak and off-peak periods as well as for urban rail and regional rail.

[18] Trip length includes both access and egress time.
[19] BART System Facts webpage

DRAFT - DELIBERATIVE

- During peak periods, good rail service is defined for urban rail as a maximum of 8 tph or the tph following implementation of the BART Core Capacity project, and as 4 tph for regional rail. This compares to expected service levels on BART of broadly 6 tph on average following the Core Capacity project and broadly 1-2 tph for regional rail.[20]
- During off-peak periods, good rail service is defined as 5 tph for urban rail and 2 tph for regional rail. This is broadly in line with expected BART service of 5 tph on average following the Core Capacity project, but quite a bit higher than the existing regional rail service of broadly 0.5 tph on average.

- **Rail transfers** indicates whether the trip involves transfers.
  - The definition of good rail service assumes that trips are direct (i.e., that no transfers are required).
  - In practice it is not viable to provide direct service between all cluster pairs. However, since good rail service is a theoretical construct used in estimating theoretical rail potential, it is appropriate to assume no transfers are required for any cluster pair.
- **Rail crowding** is defined as the average crowding level across all segments of a rail trip, e.g., if a trip from Station A to Station C has the intermediate Station B, the crowding level for a trip from Station A to C would be the average of the crowding level on the A to B segment and the B to C segment
  - Good rail service assumes there is no crowding impact.

There are various other factors not included within the regression models that could be considered for service to be considered good. These were not incorporated into the regression models as they were not expected to be main drivers that were significantly affecting rail potential. These may need to be considered during the upcoming detailed travel demand modeling work. These include, but are not limited to:

- Availability of parking at the station for those driving to the station
- Ease of access/appropriate wayfinding for those walking/cycling to the station
- Lighting, security, and amenities at the station
- Upkeep/cleanliness of the trains
- Onboard facilities (restrooms, food/beverage, Wi-Fi, etc.)
- Reliability of train service
- Ease of egress from the destination station

Table 10 summarizes the definition of each good service rail characteristic for each service type (i.e., urban rail and regional rail).

DRAFT - DELIBERATIVE

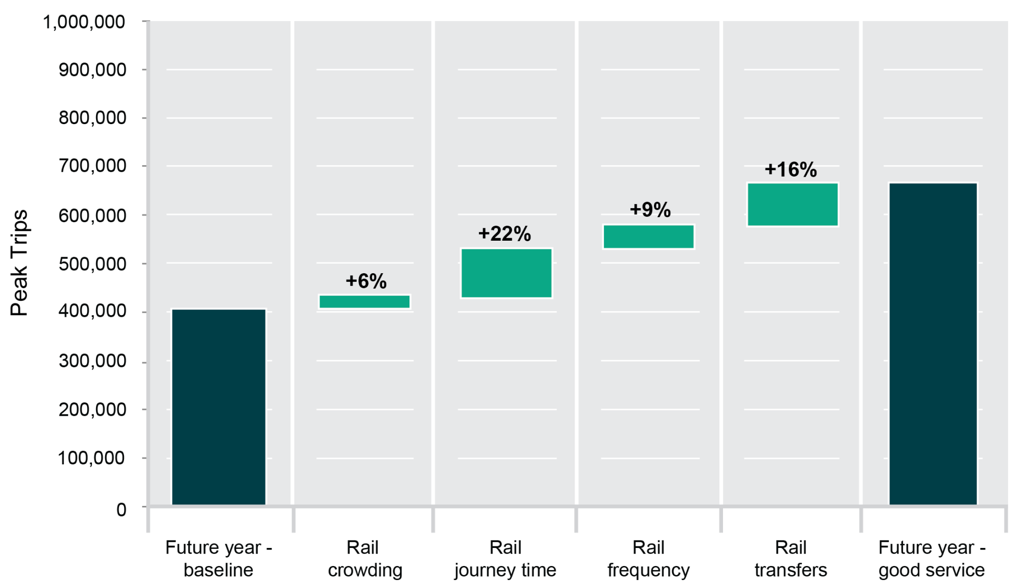
[20] There are multiple different regional rail services with a range of stopping patterns. As such, the actual tph can vary quite significantly by station.

Table 10. Definition of Good Rail Service

RAIL SERVICE CHARACTERISTIC	URBAN RAIL	REGIONAL RAIL
Rail travel time	Whichever is lower: ▫ Rail travel time based on 40 mph average speed ▫ 120% of auto travel time	Whichever is lower: ▫ Rail travel time based on 50 mph average speed ▫ 120% of auto travel time
Rail cost	Whichever is lower: ▫ $2 + $0.10 per mile (minimum of $2.10) ▫ 150% of auto cost	Whichever is lower: ▫ $0.25 per mile ▫ 150% of auto cost
Peak rail frequency	Whichever is higher: ▫ 8 tph ▫ tph following BART Core Capacity project	Peak: 4 tph
Off-peak rail frequency	5 tph	2 tph
Rail transfers	No transfers	No transfers
Rail crowding	No impacts from crowding	No impacts from crowding

Figure 21 shows how rail potential grows if good rail service is assumed for all OD pairs.

Figure 21. Good Rail Service Impacts for AM and PM Peak Trips for all OD Pairs



DRAFT - DELIBERATIVE

The largest benefits come from improved rail journey times (+22%) and replacing transfers with direct service (+16%).

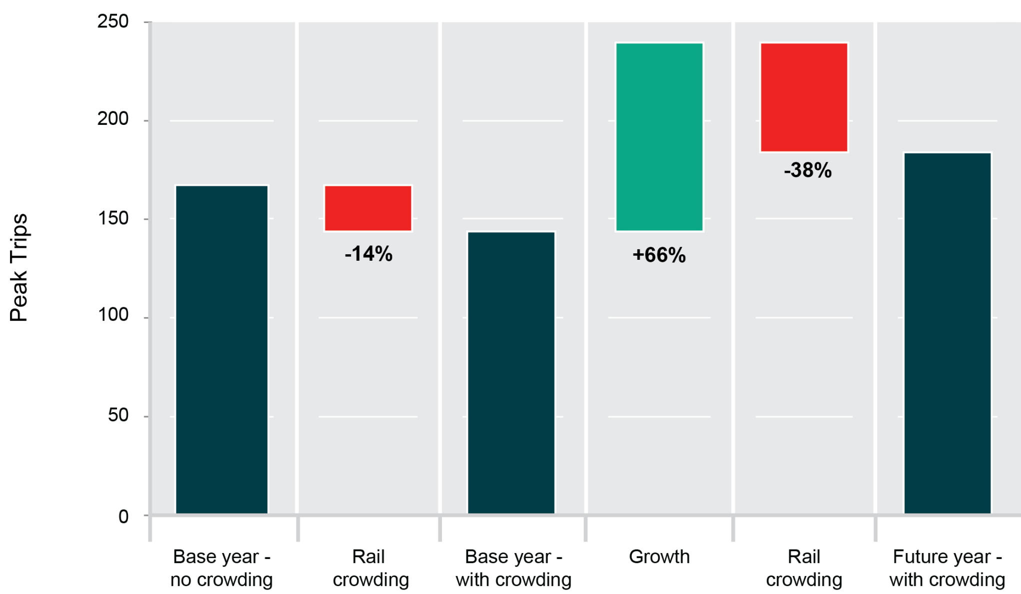
SPECIFIC EXAMPLE: EMBARCADERO TO EMERYVILLE CLUSTER PAIR

Examining Embarcadero to Emeryville illustrates the changes for this specific cluster pair.

In **Figure 22**, the following impacts are shown:

- Crowding impacts in the base year reduced rail potential by 14%.
- Growth in population and jobs in addition to service improvements increased rail potential by 66%.
- Crowding in the future year reduced rail potential by 38%, a larger impact due to high levels of growth.

Figure 22. Peak Period Impacts for the Embarcadero to Emeryville Cluster Pair



DRAFT - DELIBERATIVE

Figure 23 shows the impacts of assuming good rail service:

- Relieving crowding increased rail potential by 29%
- Improving rail journey time by 76% increased rail potential by 225%
  - Travel time improved from 81 minutes in the baseline to 20 minutes under good rail service.
- Increasing frequency from approximately 1.33/hour to 8/hour increased the rail potential by 52%.
- Removing transfers from this trip increased rail potential by 68%.

Figure 23. Good Service Impacts for the Embarcadero to Emeryville Cluster Pair