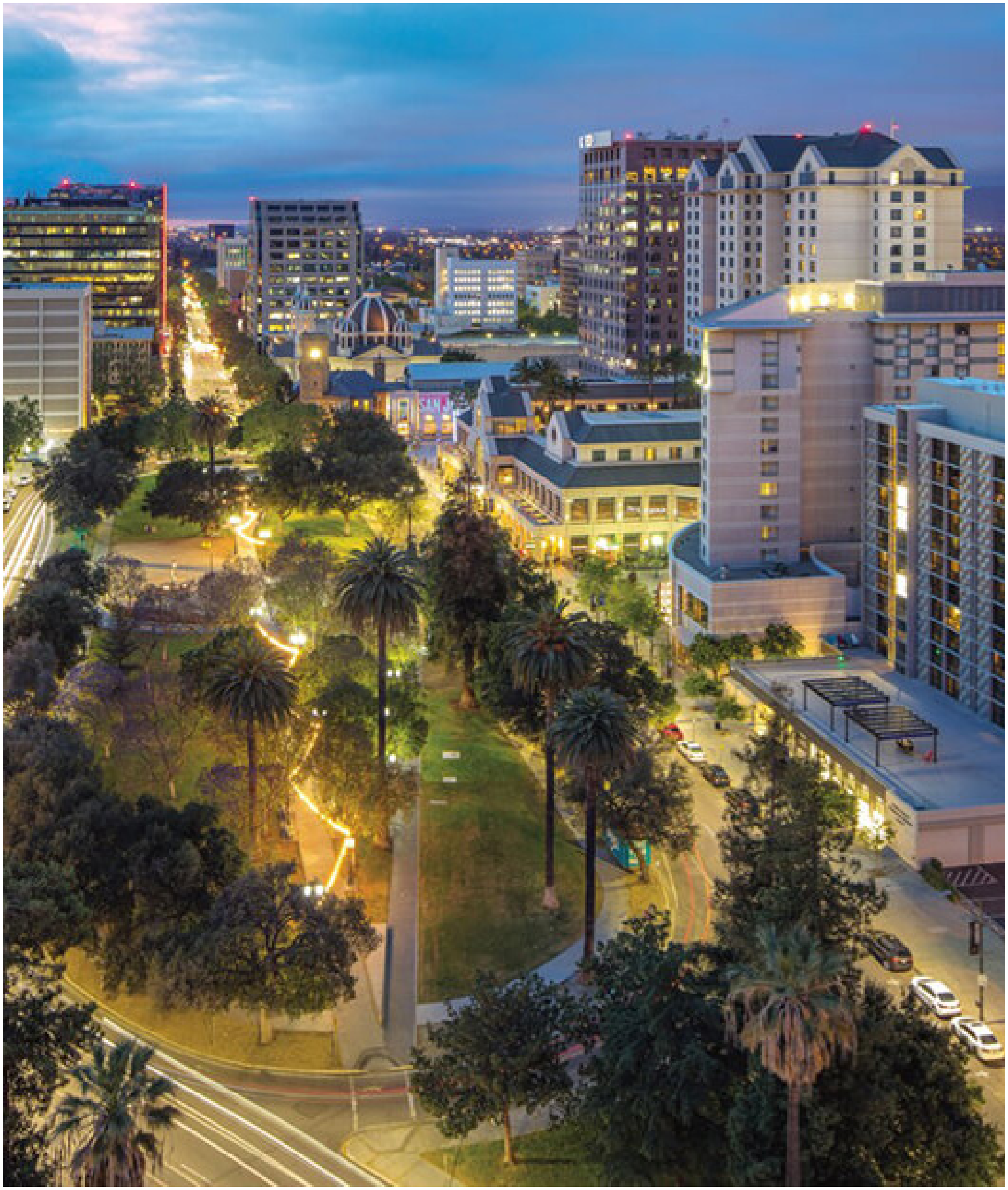

San Jose
Travel Forecasting Report

Table of contents

Introduction 1

Zones 3

Networks 6

Accessibility 7

Vehical Access Model 8

Trip Generation Models 9

Transit Ridership 10

Highway Analysis 11

Introduction

Abstract

A thriving, populous city as San Jose, California, is the ideal environment for exploring transportation implications from working from home trends given the strong presence of the tech sector. Although this industry had pioneered in these practices many years before, it is undeniable how the Covid-19 pandemic accelerated a transition to remote activities when possible, needed, or desired. The analysis expects to forecast not only reductions in Vehicle Miles Traveled (VMT) and/or ridership, but also to see some uneven distribution of travel changes because of how jobs and their associated industries are concentrated across the territory. Additionally, results and implications will be also mindful of inequalities of other kinds, regardless of the location: namely low income, no vehicle ownership, and unemployment / inactivity at the household level.

San Jose Metropolitan Statistical Area (MSA)

San Jose, located in the heart of Santa Clara Valley in California, is the most populous city in Northern California and the third most populous city in the state (according to 2020 Census, 1,013,240 inhabitants, with a decennial growth rate of 7.1%). It is also the largest city by area in Northern California, and the major city of its MSA, the San Jose-Sunnyvale-Santa Clara Metropolitan Statistical Area (MSA), having roughly half of its population.

The metropolitan region is ranked second for the most expensive rent by the National Low Income Housing Coalition. The median home price is $1.3 million and buyers' income is at least $203,000 for affordability. This is strongly related to San Jose's location in Silicon Valley. The high growth in the tech sector in the past decades, driven by large employers such as Apple and Google, contributed to a rise in the average annual salary to $83,960 (as of 2021, this was 1.57x the national average). The tech sector also provided a quick recovery from the recession brought by the pandemic.

As for transportation, the average commute time is 26-28 minutes for most of San Jose with the southern tip being the highest of 33 minutes or more. The central city of San Jose has the highest rate of use of transit and lowest use of personal vehicles for commuting of all the MSA zip codes. Despite the presence of biking and pedestrian activities in the central city, the city is still highly dependent on automobiles for commuting with over 80% of car usage. The average vehicle ownership varies from 1.8 to 2.2 with higher ownership of 2.3 cars per household or greater in certain areas to the southeast of the city and concentrated areas around Rockspring with no vehicle ownership. The proportion of commuters who cycle to work ranges from 0.1% to 1.5%, constituting a small share. Lastly, most of the transit service is concentrated in between the central city and Santa Clara along the Airport.

Methodology

Using census tracts, the number of employees (labor supply) and jobs (labor demand) were considered according to broad categories of industries: 'basic jobs', 'service jobs', and 'retail jobs'. Employees and jobs were added up since LEHD OnTheMap data suggested the share of employees working at their same tract (and thus implying no interzonal travel) was negligible, so theoretically possible overlaps have been disregarded in practice.

Service jobs were considered, as opposed to retail jobs or basic jobs, as suitable for working from home. This definition is a simplification, since belonging to an industry does not imply by itself the type of task that is performed (which is what actually would define suitability for remoteness). The assumption is that in quantitative terms, the effects are canceled (the service jobs that are not suitable for WFH compensate basic and retail jobs that are). It is acknowledged, however, that spatial distribution might experience some noise because of this.

Under a hypothesis of upcoming paradigms of working from home on a twice-a-week basis, employees would not commute to their jobs during 40% of the week. Considering the object of this travel forecasting exercise is an average weekday, this is equivalent to reducing 40% of the workforce + positions in the industry any given day (assuming WFH choice is uniform along the week, e.g., ignoring any biases towards Fridays). It must be acknowledged, however, that this reduction is not only related to the pandemic's aftermath: probably some of these workers were already remote once a week, for example, meaning that some percentage points of 40% have not been recently added. Demographic data also accounts for employees who worked from home as a default, and this information is taken into account as well prior to defining the baseline and the alternative scenario.

An additional implication of the exercise is that for commuting purposes, jobs and employees suitable for WFH disappear, but projected behavior for these workers will not be equivalent to that of an unemployed, and might imply other kinds of travel. These workers are not completely deleted, but just unobservable in terms of commuting, and reinterpreted with regards to other activities.

Zones

Summary

The target zones, defined as the census tracts belonging to the San Jose MSA (Metropolitan Statistical Area) have been analyzed for the suggested topics: mainly household, income, vehicle ownership, and employment-related census variables. On top of this diagnosis, the proposed alternative scenario accounts for deepened working from home trends as compared to a pre-pandemic baseline (sources are from 2019, namely the 2019 US ACS 5-yr. estimates). After defining assumptions about the nature and the future of those trends, the phenomenon is expected to reduce the network's number of commuters, *a priori* without affecting employment (at least directly), which would seem unprecedented at a large scale. It was known that their expected impact would be uneven both spatially and across industries, which in turn would correlate to income groups and other socioeconomic dimensions.

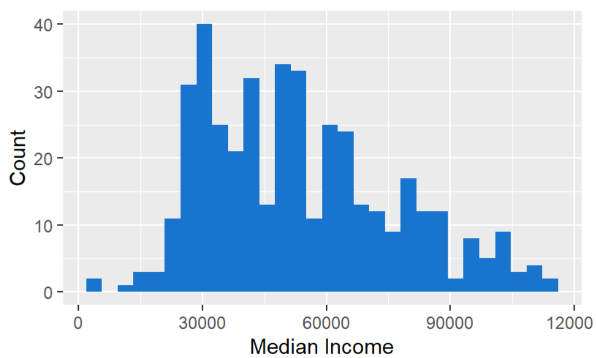

Figure 1 - Distribution of the median income across tracts.

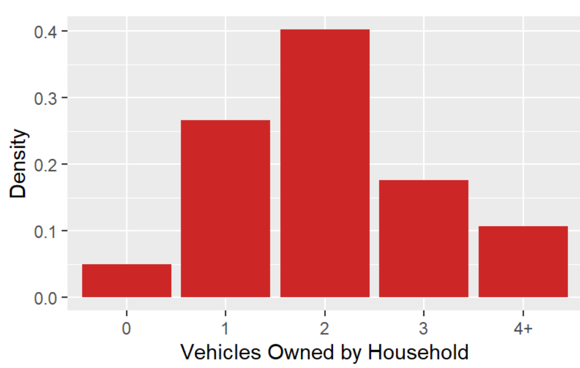

Figure 2 - Distribution of vehicle ownership across tracts.

ACS variables

The scope of the analysis of census data was related to the broad income (Figures 1 and 3) and vehicle ownership distributions (Figure 2), while also showing the spatial distribution of the most vulnerable tail: where households not owning a vehicle (Figure 5), not featuring any workers (Figure 6), or earning less than 10 thousand dollars a year (Figure 7) are concentrated. Summarized statistics of how tracts are distributed according to these vulnerability indicators in a more precise way are also provided (Figure 4).

LEHD variables

As for labor-related variables, census block-level data was collapsed into census tract-level data, which was in turn joined to its ACS counterpart. A dual approach of labor supply (employees at each location and from each

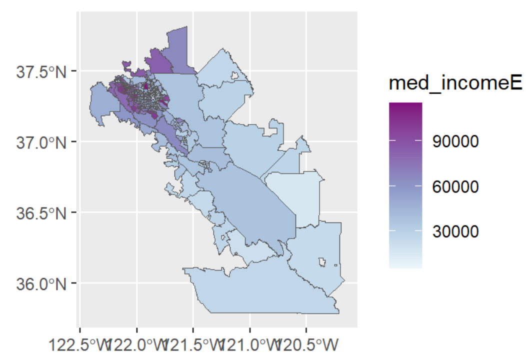

Figure 3 - Spatial distribution of median income by tract.

Vulnerability Ratios

Variable	N	Mean	Std. Dev.	Min	Pctl. 25	Pctl. 50	Pctl. 75	Max
ratio_no_veh	417	0.049	0.049	0	0.02	0.034	0.064	0.387
ratio_no_worker	417	0.19	0.074	0.028	0.143	0.176	0.231	0.551
ratio_lt_10k	417	0.029	0.025	0	0.012	0.024	0.038	0.198

Figure 4 - Summary statistics for the whole MSA, focusing on the distribution of the most vulnerable population across the census tracts.

industry) and labor demand (job positions at each location and from each industry) was conducted. One challenge at this point was the possibility of positions belonging to local residents, since summing employees and jobs per tract might had overestimated the metric because of their potential overlap. At the other end, if the overlap had been perfect (all local workers work within their tracts, and/or all local jobs are undertaken by tract residents), the correct approach would have been to compute the maximum value between labor demand and labor supply counts. As if it were not enough, an infinite continuum of intermediate Venn diagram combinations could have been the most accurate answer instead, requiring either additional (but extremely sensitive) assumptions, or further evidence.

By using LEHD OnTheMap for observing a small but representative random sample of the defined zones, it held true that the proportion of workers who happened to live at the same tract, or of residents who happened to work at the same tract, was completely negligible. The simplification of having purely independent supply and demand was then a feasible and compelling option: for modeling purposes, jobs generated at a census tract would attract trips from elsewhere, and residents from the tract having an occupation would similarly generate outbound trips according to their magnitude. This assumption also made working from home trends more visible, given that they would clearly divert from the norm (the only workers who do not travel to other defined zones).

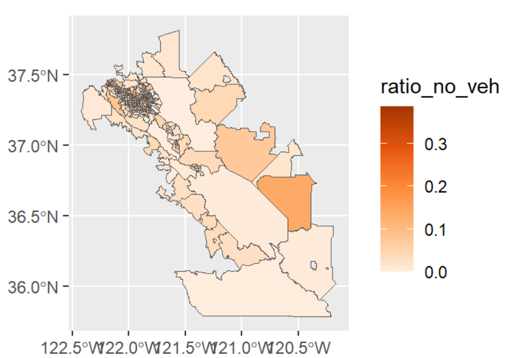

Figure 5 - Share of households not owning any vehicles. One outer census tract draws some attention, but the ratio seems to be higher near Downtown San Jose.

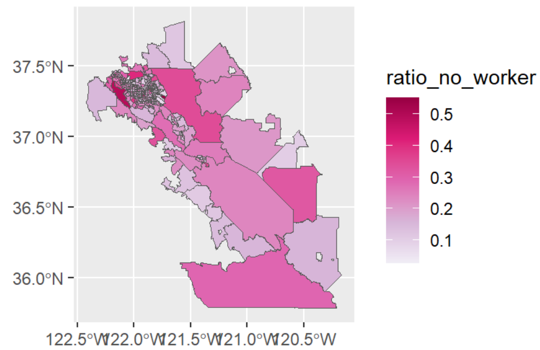

Figure 6 - Share of households not having any workers. Spatial correlation is not that clear, and senior households could offer an explanation for high values while not necessarily being problematic.

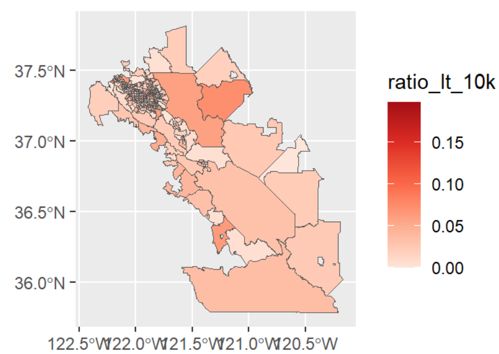

Figure 7 - Share of households with an annual income under USD 10k. Some centric and peripheral tracts show the highest values, whereas the furthest outskirts are not particularly poorer from this perspective.

WFH assumptions

Among the employees and jobs available at the LEHD database, industries suitable to working from home practices were defined. These were: Information, Finance and Insurance, Professional, Scientific, and Technical Services, Management of Companies and Enterprises, Other Services [except Public Administration], and Public Administration. Reasons for this selection were intuitive, based on stereotypical positions and tasks performed by each sector.

The future trend for these compatible jobs was set in the following way: given that trips will be estimated for the average weekday, and that a two days per week *WFH* regime seems an upcoming norm, not commuting on two out of five weekdays is equivalent to reducing the number of commuters on the average day by the same proportion: 40% less workers would travel to their jobs and back. Employees who had declared having been already working from their homes as the default regime in 2019 were deducted before applying this subtraction.

Alternative scenario: Implications

Before zooming into bias and skewness brought by the uneven distribution of the involved industries, a first implication to be mindful of is that the baseline only includes data about those who worked from home permanently (who were removed), and overlooks any possibly existing weekly regime (e.g., flexible Fridays where *WFH* is an option). For that reason, the number of commuters in the sectors suitable for working from home is set to drop by 40% with regards to the total number of workers who did not work from home permanently. However, part of this percentage was not commuting even before the pandemic. For example, if the average worker had been favored by this regime once a week, the current twice-a-week assumption would mean 40% of the workers are not on the streets commuting, but only half of them (each weekday is equal to 20 percentage points) are derived from the most recent trends, whereas the other half had been removed from commuting networks in the last decades due to technology and modern habits, and it would be inaccurate to attribute them to post-pandemic customs.

After clarifying the initial interpretations, an additional choropleth (Figure 8) shows where employees and positions that could disregard commuting are more concentrated (those suitable for *WFH* as a percentage of the total number of employees and positions altogether). On top of the correlation with income and its distributive impact, suitability was found to also be concentrated near Downtown San Jose. This poses some limits to hypothetical overall social benefits derived from reducing long-distance trips: travel time and fuel saved, or avoided externalities like emissions, might not be as large as expected.

Finally, a weighted scatter plot (Figure 9) shows a positive correlation between census tracts' ability to work from home and their median incomes. This would not necessarily harm the poorest census tracts, but the reduction of any trip saturation would be rather limited to the richer ones. What is an additional concern is reflected by the weights: thicker points indicate higher percentages of households not owning a vehicle. Not surprisingly, most of them are located at the bottom of the income distribution, where *WFM* is not able to help in mitigating the demand for mobility. Nevertheless, the potential alleviation of traffic, transit, and transportation in general for the richest sectors thanks to post-pandemic working from home trends should make public resources available to be instead invested in the poorer communities who most need it.

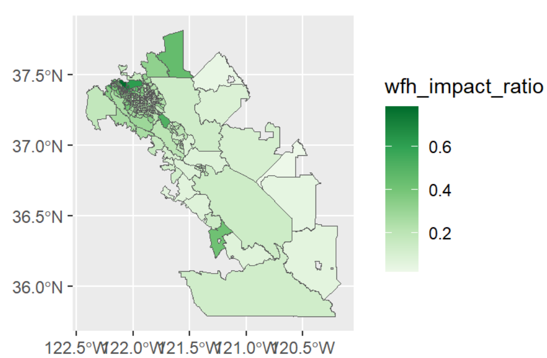

Figure 8 - Ratio of jobs that would be suitable for working from home according to this paper's definition. Recall Figure 3 and how similar their highlighted census tracts are. Additionally, Downtown San Jose concentrates most of the jobs suitable for the analyzed regime.

Figure 9 - WFH suitability versus median income. The positive correlation is undeniable, and the large circles at the bottom (tracts with many households not owning vehicles and cannot work from home) suggest possible scopes for future policymaking.

Networks

Accessibility

Vehical Access Model

Trip Generation Models

Transit Ridership

Highway Analysis